THE LAW OF DATA SCRAPING

A REVIEW OF UK LAW ON TEXT AND DATA MINING

CREATe Working Paper 2021/2

SHEONA BURROW

The Law of Data Scraping: A review of UK law on text and data mining

Sheona Burrow*

Abstract

Data is perceived to be a key asset in the digital economy. Many governments have been keen to promote and exploit data driven economies. Data scraping is a widely used technique that automatically extracts information from different (often online) sources, whilst data mining is the machine reading of data to identify useful information not immediately obvious on human reading. In 2014, the UK implemented a limited exception to copyright law for text and data mining (TDM). However, copyright is only one layer of legal protection available to 'data' and the protection of data has been the subject of a long-running tension between property based rights and concurrent protection for data owners in liability rules arising through competition and contract law. Maintaining an appropriate balance between protecting rightholders and users has remained problematic. This paper summarises the legal protection available in the UK for different types of data, and the (limited) interpretation of that protection by the UK courts. The analysis is situated in a review of the academic literature. Ultimately this paper will conclude that the layered protection for data is confusing for end users, and that the case law on the protection and exceptions available to those seeking to engage in TDM limited and fact dependent.

* Sheona Burrow is a (part time) postdoctoral research fellow at CREATe, the UK Copyright and Creative Economy Centre, University of Glasgow, and a qualified solicitor in Scotland, England and Wales. The research has been supported by the ESRC Urban Big Data Centre at the University of Glasgow (grant numbers ES/S007105/1 and ES/L011921/1.

Introduction

Initially addressed as a 'challenge of technology,'[1] electronic data is now embraced as a key 'asset' in the digital economy[2] and 'big data' has been identified as the 'next frontier for innovations, competition and productivity.[3] Inter-governmental discussions look at promoting the 'free flow of information' in the digital economy[4] and the 'digital single market.'[5] Yet, challenges of availability and accessibility remain[6] and the sheer volume of data now collected (whether through individual use of applications and services or by businesses operating in the field of data collection[7]) may have led to a 'data rich but information poor'[8] situation, where data analysis tools are desperately needed to turn data into social or commercial value. This balance between effective use of information and incentives for collection and exploitation of data has not been achieved with clarity and this is apparent when considering the ability to 'mine' data for useful information.

Text and data mining (TDM) are one of those tools, whereby machine reading of large amounts of data can identify patterns, trends and other useful information not obvious through human reading.[9] Typically text mining can be employed on structured or unstructured collections of text

---

[1] European Commission, Green Paper on copyright and the challenge of technology Copyright Issues requiring immediate action dated 7 June 1988 Catalogue number CB-CO-88-267-EN-C ISSN 0254-1475.

[2] See e.g. OECD, Data in the Digital Age, OECD Publishing, Paris, 2019 https://www.oecd.org/going-digital/data-in-the-digital-age.pdf.

[3] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung Byers, A., Big Data: The Next Frontier for Innovation, Competition and Productivity, McKinsey Global Institute, 1 May 2011 - https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation#. See also Intellectual Property Office, Eight Great Technologies: Big Data (2014) available at https://www.gov.uk/government/publications/eight-great-technologies-big-data.

[4] OECD, The OECD Ministerial Declaration on the Digital Economy: Innovation, Growth and Social Prosperity (the Cancun Declaration), OECD, Paris, 2016 https://www.oecd.org/internet/Digital-Economy-MinisterialDeclaration-2016.pdf. This has led to a series of recommendations on research data, public sector information and health data. See OECD (2006), Recommendation of the Council concerning Access to Research Data from Public Funding, OECD (2008), Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information and OECD (2016) Recommendation of the Council on Health Data Governance.

[5] European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, The Digital Agenda for Europe – Driving European Growth Digitally, COM (2012) 784 final,18th December 2012 and European Commission (2015) Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A digital single market strategy for Europe, COM (2015) 192 final, 6 May 2015.

[6] OECD, Data in the Digital Age, OECD Publishing, Paris, 2019 https://www.oecd.org/going-digital/data-in-the-digital-age.pdf p3.

[7] *ibid* p2.

[8] Han, J., Kamber, M., Pei, J., Datamining: concept and techniques, 3rd edn. (Morgan Kaufmann, Waltham, USA, 2011), p5.

[9] For explanation and discussion of the techniques and value of TDM see e.g. Hargreaves, I., Guibault, L., Handke, C., Valcke, P., & Martens, B. (2014). Standardisation in the area of innovation and technological development, notably in the field of Text and Data Mining: report from the expert group. (Studies and reports). Publications Office of the European Union. https://doi.org/10.2777/71122, Brook, M., Murray-

whereas data mining is employed on databases.[10] JISC identified six broad categories of use: systematic reviews of literature, development of new hypotheses, testing of hypotheses, reusable representations and enhancing usability of the research base.[11] This has potential to realise both economic and social potential in the UK.[12] According to Borghi:

> the processing of data contained in a large collection of scientific papers in a particular medical field could suggest a possible association between a gene and a disease, or between a drug and an adverse event, without this connection being explicitly identified or mentioned in any of the papers.[13]

One of the issues identified in JISC's 2012 report on the value and benefits of TDM in the UK[14] is a lack of clarity around the legal protection of data and the permitted uses of data.[15] The tension between accessibility, lawful use and the protection of the economic position of those who collate or add value to information has been the subject of academic debate for decades.[16] The protection of 'information' and the investment in obtaining and exploiting information has potentially always been somewhat unclear, with shifting ground on the 'originality

---

Rust, P. and Oppenheim, C., "The Social, Political and Legal Aspects of Text and Data Mining (TDM)" (2014) D-Lib Magazine, 20(11/12), doi:10.1045/november2014-brook accessed at https://openaccess.city.ac.uk/id/eprint/4784/1/D-Lib%20Magazine.pdf and Cocoru, D., & Boehm, M., "An analytical review of text and data mining practices and approaches in Europe", 2016, Open Forum Europe available at https://www.creative-destruction.org/publication/ofe-tdm/.

[10] Hearst, M., "What is Text Mining?" Online. Available: http://people.ischool.berkeley.edu/~hearst/text-mining.html last accessed 14 December 2020.

[11] JISC, *The Value and Benefit of Text Mining to UK Further and Higher Education*. Digital Infrastructure (2012) Available at: http://bit.ly/jisc-textm Programme: Digital Infrastructure www.jisc.ac.uk/whatwedo/programmes/di_directions.aspx pp15-16.

[12] *Ibid*.

[13] See Borghi, M., 'Text and Data Mining', Copyright User - https://www.copyrightuser.org/understand/exceptions/text-data-mining/.

[14] JISC (2012) *supra* note 11, pp15-16.

[15] *ibid* p21.

[16] See e.g. Reichman, J.H., & Samuelson, P., "Intellectual Property Rights in Data" (1997) *Vanderbilt Law Review* Vol. 50, p51, Wolken, J.C., "Just the Facts, Ma'am – A Case for Uniform Federal Regulation of Information Databases in the New Information Age", (1998) *Syracuse Law Review*, Vol. 48, p1263, Maurer, S., Hugenholtz, P., Onsrud, H., "Europe's Database Experiment" (2001), *Science*, Vol. 294, pp789-790, Westkamp, G., "Protecting Databases Under US and European Law – Methodical Approaches to the Protection of Investments Between Unfair Competition and Intellectual Property Concepts", (2003) *IIC* Vol. 34 pp772-778, Dietrich, N., Guibault, L., Margoni, T., Siewicz, K., Spindler, G., & Wiebe, A., 'Safe to Be Open: Study on the Protection of Research Data and Recommendations for Access and Usage' (2013) Göttingen, Germany: Universitätsverlag Göttingen, available at: http://eprints.gla.ac.uk/129335 and Drexl, J. Designing Competitive Markets for Industrial Data: Between Propertisation and Access (2016) *Max Planck Institute for Innovation and Competition Research Paper No. 16-13* available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2862975. For those interested, a list of the key contributions to this literature is appended in Appendix 2.

requirements'[17] and the protection of compilations by copyright[18] leading to a gap that the EU attempted to fill with the sui generis database right in 1996,[19] alongside the increasing tendency towards protection for personal information[20] that culminated in the General Data Protection Regulation 2016.[21] Recently the debate has moved away from the deficiencies of the sui generis database right to the need for a 'data producers right',[22] and whether this swings the balance too far from lawful use.[23] Discussion of the 'press publisher's right' also discusses the status of information under copyright law.[24]

[17] See e.g. Ginsburg, J., "Creation and Commercial Value: Copyright Protection of Works of Information", (1990) *Columbia Law Review*, Vol. 90 (7), p1865, Ginsburg, J., "No Sweat Copyright and Other Protection of Works of Information and *Feist v Rural Telephone*", (1992) *Columbia Law Review*, Vol. 92(2), p338-388, Sheils, P.T., & Penchina, R., "What's all the Fuss About Feist – The Sky Is Not Falling on the Intellectual Property Rights of Online Database Proprietors," (1992) *University of Dayton Law Review*, Vol.17, p563, Cerina, P., "The Originality Requirement in the Protection of Databases in Europe and the United States," (1993) *IIC* p579.

[18] Hicks, J.B., "Copyright and Computer Databases: Is Traditional Compilation Law Adequate?" (1987) Vol. 65 *Texas Law Review* p993.

[19] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases.

[20] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and the UK Data Protection Act 1998.

[21] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC ("GDPR").

[22] Hoeren, T., "Big Data and the Ownership in Data: Recent Developments in Europe,"(2014) *European Intellectual Property Review* pp751-754, De Franceschi, A. & Lehmann, M., "Data as tradable commodity and new measures for their protection," (2015) *Italian Law Journal,* Vol. 1, p51-72, Wiebe, A., "Protection of industrial data – a new property right for the digital economy?", (2017) *Journal of Intellectual Property Law & Practice,* Vol.12(1), p62–71, and Stepanov, I., "Introducing a property right over data in the EU: the data producer's right – an evaluation," (2019) *International Review of Law, Computers and Technology* Vol. 34(1), p65 – 86.

[23] See e.g. Hugenholtz, P., "Abuse of Database Right: Sole-source information banks under the EU Database Directive" Paper presented at Conference 'Antitrust, Patent and Copyright', École des Mines/UC Berkeley, Paris, January 15-16, 2004 available at https://www.ivir.nl/publicaties/download/abuseofdatabaseright.pdf , Zech, H., "A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data", (2016) *Journal of Intellectual Property Law & Practice* Vol. 11, pp460-470, Barbero, M., Cocoru, D., Graux, H., Hillebrand, A., Linz, F., Osimo, D., Siede, A., Wauters, P., "Study on emerging issues of data ownership, interoperability, (re-)usability and access to data, and liability", 25 April 2018. Available at https://ec.europa.eu/digital-single-market/en/news/study-emerging-issues-data-ownership-interoperability-re-usability-and-access-data-and Taylor, L., "The Ethics of Big Data as a Public Good: Which Public? Whose Good?." *LawArXiv*, 19 June 2018. Web. https://doi.org/10.1098/rsta.2016.0126 and Ducato, R., & Strowel, A., "Limitations to text and data mining and consumer empowerment: making the case for a right to "machine legibility"' (2019) *IIC* Vol. 50(6), p649-684 and Tombal, T., "Economic dependence and data access" (2020) IIC Vol. 51(1) p70-98.

[24] With reference to the Berne Convention Article 10(1). See also Höppner, T., Kretschmer, M., Xalabarder, R., "CREATe public lectures on the proposed EU right for press publishers", (2017) *European Intellectual Property Review* Vol. 39(10), pp.607-622.

The decision to place protection of information firmly within the realm of property based rights[25] was always controversial.[26] Such protection co-exists alongside contractual and unfair competition rights and remedies.[27] The result is a patchwork of applicable legal protections and exceptions.[28] Critics have suggested this inhibits the efficiency of the market in data[29] and can make engaging with data for analytics purposes daunting. JISC identified 'copyright-law driven restrictions' on text mining as an inhibitor in the take up of TDM in the UK terming this as a 'market

[25] For more on data property see Hugenholtz, P.B., "Data Property: Unwelcome Guest in the House of IP", in J. Reda (Ed.), Better Regulation for Copyright: Academics meet Policy Makers: Wed 6 Sept 2017 15:00-18:30: University of Southampton, The Greens|EFA in the European Parliament (pp. 65-77). https://juliareda.eu/events/better-regulation-for-copyright/ available at https://www.ivir.nl/publicaties/download/Data_property_Muenster.pdf; Hoeren, T. & Bitter, P., "Data ownership is dead: long live data ownership" (2018) *European Intellectual Property Review* Vol. 40(6), p347-348; Geiger, C., "The future of copyright in Europe: striking a fair balance between protection and access to information" (2010) *Intellectual Property Quarterly,* Vol. 1, p1-14

[26] See Metaxas, G., "Protection of databases: quietly steering in the wrong direction?" (1990) *European Intellectual Property Review*, Vol. 12(7), pp227-234, Rosler, D. B., "The European Union's Proposed Directive for the Legal Protection of Databases: A New Threat to the Free Flow of Information" (1995) *High Technology Law Journal*, Vol. 10, p105, Westkamp, G., "Protecting Databases Under US and European Law – Methodical Approaches to the Protection of Investments Between Unfair Competition and Intellectual Property Concepts", (2003) *IIC* Vol. 34, p772-778. For more generally on the tension between property and liability rules see Calabresi, G., & Melamed, A.D., "Property Rules, Liability Rules and Inalienability: One view of the Cathedral" (1972) *Harvard Law Review* Vol. 85, p1089. See also Cooper, E. and Burrow, S. "Photographic Copyright and the Intellectual Property Enterprise Court in Historical Perspective," (2018), *Legal Studies*, Vol. 39(1), pp. 143-165.

[27] Sanks, T., "Database Protection National and International Attempts to Provide Legal Protection for Databases," (1998) *Florida.State University Law Review* Vol. 25, p991, Westkamp, G., "Protecting Databases Under US and European Law – Methodical Approaches to the Protection of Investments Between Unfair Competition and Intellectual Property Concepts", (2003) *IIC* , Vol. 34, pp772-778, and Drexl, J., "Connected Devices – An Unfair Competition Law Approach to Data Access Rights of Users" (Max Planck Institute for Innovation & Competition Research Paper, No. 20-22), (2020) Available at SSRN: https://ssrn.com/abstract=3746163

[28] Triaille, J.P., de Meeûs d'Argenteuil, J., & de Francquen, A., "Study on the legal framework of text and data mining", 2014, Publications Office of the European Union DOI : 10.2780/1475

[29] Handke C, Guibault L, Vallbé J-J., "Is Europe falling behind in data mining? Copyright's impact on data mining in academic research" in: Schmidt B, Dobreva M (eds) *New avenues for electronic publishing in the age of infinite collections and citizen science: scale, openness* and trust (IOS Press, Amsterdam, Netherlands, 2015). Even the European Commission itself expressed doubt about the effectiveness of the SGDR - European Commission, *First evaluation of Directive 96/9/EC on the legal protection of databases*, 12 December 2005 - although this has been tempered in the second evaluations European Commission, *Evaluation of Directive 96/9/EC on the legal protection of databases*, 25th April 2018. See also Austin, W.L., "A Thoughtful and Practical Analysis of Database Protection under Copyright Law, and a Critique of Sui Generis Protection" (1997) *Journal of Technology Law & Policy*, Vol. 3, p35, Reichman, J.H., & Uhlir, P.H., "Database Protection at the Crossroads: Recent Developments and Their Impact on Science and Technology" (1999) *Berkeley Technology Law Journal*, Vol. 14, p793, Maurer, S., Hugenholtz, P., Onsrud, H., "Europe's Database Experiment" (2001) *Science*, Vol. 294, pp789-790, Derclaye, E., "Do sections 3 and 3A of the CDPA violate the Database Directive? A closer look at the definition of a database in the U.K. and its compatibility with European law" (2002) *European Intellectual Property Review*, Vol.24(10), pp466-474, Lipton, J., "Databases as intellectual property: new legal approaches", (2003), *European Intellectual Property Review*, Vol.25(3), pp139-145.

failure'.[30] The Hargreaves *Review on Intellectual Property and Growth* in 2011[31] was the starting gun for a series of legal reforms in the UK,[32] some of which sought to clarify the law on the protection and use of information to fully realise potential[33] and there was specific inclusion of a copyright exception for TDM.[34] However, by enacting a copyright only exception, the UK government has addressed only one aspect of the patchwork of legal protection, despite academics urging the government to consider something more broad ranging and functional[35] suggesting it is time for European countries to consider 'fair use'.[36]

The goal of this paper is to summarise the legal protection currently available in the UK for different types of data, and the exceptions to those protections that allow for data use, particularly in the field of TDM or web or data 'scraping.' This paper will not consider normative questions around the legal protection of data, readers are directed to Reichman and Samuelson's comprehensive 1997 paper on *Intellectual Property Rights in Data*[37] for an historical overview of these aspects and to the literature listed in Appendix 2 for further academic commentary of the normative aspects of protection. JISC's 2012 Report *The Value and Benefit*

---

[30] JISC (2012) *The Value and Benefit of Text Mining to UK Further and Higher Education*. Digital Infrastructure. Available at: http://bit.ly/jisc-textm Programme: Digital Infrastructure www.jisc.ac.uk/whatwedo/programmes/di_directions.aspx pp21.

[31] Hargreaves, I., *Review of Intellectual Property and Growth,* Intellectual Property Office, May 2011 ("Hargreaves Review").

[32] Copyright and Rights in Performances (Quotation and Parody) Regulations 2014, the Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014, the Copyright and Rights in Performances (Disability) Regulations 2014, and the Copyright and Rights in Performances (Personal Copies for Private Use) Regulations 2014.

[33] Hargreaves Review, p41.

[34] Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014/1372 reg.3(2) inserting section 29A into the Copyright Designs and Patents Act 1988.

[35] Margoni, T., & Dore, G., "Why we need a text and data mining exception (but it is not enough)", 10th edition of the Language Resources and Evaluation Conference (LREC 2016), Portorož (Slovenia), 23-28 May 2016 available at https://zenodo.org/record/248048#.WXdf2oiGNEY and Margoni T., & Kretschmer M. (2018) The text and data mining exception in the proposal for a Directive on Copyright in the Digital Single Market: why it is not what EU copyright law needs. https://www.create.ac.uk/blog/2018/04/25/why-tdm-exception-copyright-directive-digital-single-market-not-what-eu-copyright-needs/ and Rosati, E., "An EU text and data mining exception for the few: would it make sense?" (2018) *Journal of Intellectual Property Law & Practice*, Vol. 13(6), pp429-430, Geiger, C., Frosio, G., & Bulayenko, O., "Text and data mining in the proposed copyright reform: making the EU ready for an age of big data? Legal analysis and policy recommendations (Legislative Comment)" (2018) *IIC,* Vol., 49(7), pp814-844.

[36] See Margoni, T., & Dore, G., "Why we need a text and data mining exception (but it is not enough)," 1 April 2016. 10th edition of the Language Resources and Evaluation Conference (LREC 2016), Portorož (Slovenia), 23-28 May 2016 available at https://zenodo.org/record/248048#.WXdf2oiGNEY and Aplin., T., & Bently, L., *Global Mandatory Fair Use: The Nature and Scope of the Right to Quote Copyright Works* (Cambridge: Cambridge University Press, 2020). The position taken by US academics is that TDM falls within the scope of fair use – see Carroll, M., "Copyright and the Progress of Science: Why Text and Data Mining is Lawful" (2019) *University of California Davis Law Review*, Vol.53, p893 and Sag, M., "The New Legal Landscape for Text Mining and Machine Learning" (2019) *Journal of the Copyright Society of the USA*, Vol.66, p3.

[37] Reichman, J.H. & Samuelson, P. "Intellectual Property Rights in Data" (1997) *Vanderbilt Law Review* Vol. 50, p51.

of Text Mining to UK Further and Higher Education[38] provides further background about TDM in the UK and the non-legal challenges relating to TDM, which are not discussed in this paper.

After setting out the current understanding of TDM in part 1, this paper will describe the legislation in the UK that affords legal protection to different types of data and summarise the exceptions under this legislation in part 2. The paper will then look at the interpretation of the legal protection and exceptions in the UK, and where relevant EU, courts. This paper will conclude that the layered protection of data is confusing for end users and that UK case law is both limited and fact dependent, offering little insight into developing general approaches to TDM in any industry.

1. Text and Data Mining (TDM)

Han et al suggest that data mining can 'be viewed as a result of the natural evolution of information technology'.[39] Similar terms such as 'knowledge mining from data', 'knowledge extraction', 'data/pattern analysis', 'data archaeology', 'data dredging' or 'knowledge discovery from data (KDD).'[40] Han et al describe data mining as sitting within a wider process of data cleaning, integration, selection and transformation, followed by mining activity which can then lead to pattern evaluation and knowledge presentation.[41] Their definition is: -

> Data mining is the process of discovering interesting patterns and knowledge from large amounts of data.[42]

Within the legal sphere, the process is typically referred to as Text and Data Mining (TDM) which is conceived as "a research technique to collect information from large amounts of digital data through automated software tools."[43]

Whilst exceptions to certain acts for non-commercial or research purposes existed in the Copyright, Designs and Patents Act 1988 the 2011 *Hargreaves Review* suggested that these were insufficient to cover text mining and noted that such activity was often prohibited by the terms of use for the database. The issues for the research community engaging in TDM were

---

[38] JISC (2012) *The Value and Benefit of Text Mining to UK Further and Higher Education*. Digital Infrastructure. Available at: http://bit.ly/jisc-textm Programme: Digital Infrastructure www.jisc.ac.uk/whatwedo/programmes/di_directions.aspx.

[39] Han, J., Kamber, M., & Pei, J., *Datamining: concept and techniques*, 3rd edn. (Morgan Kaufmann, Waltham, USA, 2011), p2.

[40] *Ibid* p6.

[41] *Ibid* p7.

[42] *Ibid* p8.

[43] Geiger, C., Frosio, G., and Bulayenko, O., "Text and data mining in the proposed copyright reform: making the EU ready for an age of big data? Legal analysis and policy recommendations (Legislative Comment)" (2018) *International Review of Intellectual Property & Competition Law*, Vol. 49(7), pp814-844.

perceived to be "political, social and legal" [44] and are part of a broader discussion about open access.[45] Professor Hargreaves recommended that the UK introduced a specific UK exception to allow use of analytics for non-commercial use as well as promoting a TDM exception for commercial use at EU level.[46] Whilst the proposed copyright exception for TDM was welcomed in principle, critics identified the very limited nature of the exception as problematic and called for wider scope to allow full utilisation of the potential of TDM outside the research community.[47] The UK implemented a specific copyright exception for TDM in 2014[48] and in 2016, the European Commission published a Proposal for the Directive on copyright in the Single Digital Market which aimed to address the 'legal uncertainty' around TDM.[49]

[44] For more see JISC (2012) *supra* note 11, Brook, M., Murray-Rust, P. and Oppenheim, C., "The Social, Political and Legal Aspects of Text and Data Mining (TDM)" (2014) *D-Lib Magazine*, 20(11/12), doi:10.1045/november2014-brook accessed at https://openaccess.city.ac.uk/id/eprint/4784/1/D-Lib%20Magazine.pdf p2 and Hargreaves, I., Guibault, L., Handke, C., Valcke, P., & Martens, B. (2014). *Standardisation in the area of innovation and technological development, notably in the field of Text and Data Mining: report from the expert group*. (Studies and reports). Publications Office of the European Union. https://doi.org/10.2777/71122.

[45] See Dietrich, N., Guibault, L., Margoni, T., Siewicz, K., Spindler, G., & Wiebe, A., "Safe to Be Open: Study on the Protection of Research Data and Recommendations for Access and Usage" (2013) Göttingen, Germany: Universitätsverlag Göttingen, available at: http://eprints.gla.ac.uk/129335 and Doldirina, C., et al. "Legal Approaches for Open Access to Research Data." *LawArXiv*, 1 Apr. 2018. Web. https://doi.org/10.31228/osf.io/n7gfa.

[46] Hargreaves, I, *Digital Opportunity: A Review of Intellectual Property and Growth*, Intellectual Property Office, May 2011, p48.

[47] European Copyright Society (2017) *General opinion on the EU copyright reform package*, 24 January 2017 avilable at https://europeancopyrightsocietydotorg.files.wordpress.com/2015/12/ecs-opinion-on-eu-copyright-reform-def.pdf at p5. See also Margoni, T., & Dore, G.,"Why we need a text and data mining exception (but it is not enough)", 1 April 2016. 10th edition of the Language Resources and Evaluation Conference (LREC 2016), Portorož (Slovenia), 23-28 May 2016 available at https://zenodo.org/record/248048#.WXdf2oiGNEY Geiger, C., Frosio, G., & Bulayenko, O., "Opinion of the CEIPI on the European Commission's Proposal to Reform Copyright Limitations and Exceptions in the European Union" Centre for International Intellectual Property Studies Research Paper No. 2017-09 available at http://infojustice.org/wp-content/uploads/2017/10/Opinion-of-the-CEIPI-on-the-European-Commissions-Proposal-to-Reform-Copyright-Limitations-and-Exceptions-in-the-European-Union.pdf , Hilty RM & Richter H (2017) Text and data mining. In: Hilty, R. M., Moscon, V., (eds) *Modernisation of the EU copyright rules. Position Statement of the Max Planck Institute for Innovation and Competition*, Max Planck Institute for Innovation and Competition Research Paper No. 17-12, available at https://pure.mpg.de/rest/items/item_2470998_12/component/file_2479390/content , Geiger, C., Frosio, G., & Bulayenko, O., "Text and data mining in the proposed copyright reform: making the EU ready for an age of big data? Legal analysis and policy recommendations (Legislative Comment)" (2018) *IIC*, Vol. 49(7), pp814-844, Rosati, E., "An EU text and data mining exception for the few: would it make sense?" *(2018) Journal of Intellectual Property, Law and Practice,* Vol. 13(6), pp429-430, Raue, B., "Free flow of data? The friction between the Commission's European data economy initiative and the proposed Directive on Copyright in the Digital Single Market (Editorial)," (2018) *IIC*, Vol. 49(4), pp379-383.

[48] Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014/1372 reg.3(2) inserting section 29A into the Copyright Designs and Patents Act 1988.

[49] European Commission, Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market, 2016/0280(COD).

However, as this paper will show, such an exception only applies to one aspect of the legal protection of data (copyright) and in a limited way so it is by no means clear that text and data mining is now permissible in the UK without incurring legal liability towards a rightsholder.

2. What types of data are protected by law and what exceptions are there?

2.1. What is meant by data?

> We therefore must not refer to data as a uniform entity, as this may lead to misunderstandings, oversimplifications and less effective policy.[50]

The term 'data' has a broad meaning. Within the legal sphere, the terms 'data' and 'information' appear to be used almost interchangeably,[51] but without rigorous analysis of what may or may not be data – for example, are naturally occurring phenomena capable of being 'data'?[52]

The discipline of information science devotes more time to such analysis, without arriving at conclusions that assist in analysing the law applicable to text and data mining. For example, Ballsun-Stanton perceives 'data', 'information' and 'knowledge' as distinct, ands suggests three 'philosophies' of data science: "data as bits" "data as hard numbers," and "data as recorded observations."[53] 'Information' on the other hand, is seen as data which is well formed and meaningful,[54] in contrast to 'raw data'[55] which may be abstract or basis.

For the purposes of this Working Paper, the definition of 'data' will be taken to encapsulate both 'raw' and 'informational' data, as both may be of interest for the purposes of text and data mining. However, as part 2 will show, protection afforded to data may depend on the level of information in that data. For example, the definition of 'personal data' requires it to be information capable of identifying a living individual. In their approach to 'disentangling different types of data'

---

[50] OECD, *Data in the Digital Age*, OECD Publishing, Paris, 2019 https://www.oecd.org/going-digital/data-in-the-digital-age.pdf p2.
[51] See for example, the definition of 'personal data' as 'information' in GDPR. Section 1.3.
[52] For more see Bygrave, L.A., "Information Concepts in Law: Generic Dreams and Definitional Daylight" (2015) *Oxford Journal of Legal Studies*, Vol. 35(1), pp91-120 and Bygrave, L. A., "The Data Difficulty in Database Protection" (June 20, 2012). *University of Oslo Faculty of Law Research Paper No. 2012-18*, Available at SSRN: https://ssrn.com/abstract=2088018.
[53] Ballsun-Stanton, B., "Asking about data: Experimental philosophy of Information Technology", (2010) *Computer Science: 5th International Conference on Computer Sciences and Convergence Information Technology* accessed at https://www.researchgate.net/publication/224218081_Asking_about_data_Experimental_philosophy_of_Information_Technology.
[54] See Floridi, L., "Philosophical Conceptions of Information" in Sommaruga G. (eds) *Formal Theories of Information. Lecture Notes in Computer Science* (Springer, Berlin, Heidelberg, 2009) and Floridi, L., "Is Semantic Information Meaningful Data?" (2005) *Philosophy and Phenomenological Research,* Vol. 70(2).
[55] Should such a concept be capable of recognition – see Gitelman, L., ed. *Is Raw Data an Oxymoron*, (MIT Press, 2013).

the OECD show proximate assessments of the personal content of data, with Business to Business GVC data having the lowest level of personal content and Citizen to Citizen communications the highest level of content.[56]

In the same way that copyright law can require the work to be 'fixed in some material form'[57] which is generally understood to mean that it is 'perceived, reproduced, or otherwise communicated for a period of more than transitory duration'[58], data is unlikely to be useful for mining if it is not fixed and therefore this paper limits its discussion to fixed data.

The rest of this section 2 will consider the protections that arise in the following types of data: data that constitutes copyright protected 'works', data protected by the *sui generis* database right (SGDR), data that identifies a living individual ('personal data') and finally data that is protected because of contractual obligations, confidentiality or technological protection measures (TPMs).

2.2. Copyright in 'data'

2.2.1. UK Legal regime for copyright.

The basis for copyright protection in the UK is statutory, with the Copyright, Designs and Patents Act 1988 (as amended)[59] (CDPA) applicable across the UK, copyright being a matter reserved to the UK government. Where UK law has been harmonised with EU Directives on copyright[60], this

[56] OECD, *Data in the Digital Age*, OECD Publishing, Paris, 2019 https://www.oecd.org/going-digital/data-in-the-digital-age.pdf p2.
[57] Article 2(2), Berne Convention
[58] The (US) Copyright Act of 1976, 17 U.S.C. § 101. In the UK, the requirement is simply specified as 'in writing or otherwise' – s3(2) CDPA.
[59] Copyright, Designs and Patents Act 1988, c48.
[60] See Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society ("InfoSoc Directive"), Directive 2006/115/EC of the European Parliament and of the Council of 12 December 2006 on rental right and lending right and on certain rights related to copyright in the field of intellectual property ("Rental and Lending Directive"), Directive 2001/84/EC of the European Parliament and of the Council of 27 September 2001 on the resale right for the benefit of the author of an original work of art ("Resale Right Directive"), Council Directive 93/83/EEC of 27 September 1993 on the coordination of certain rules concerning copyright and rights related to copyright applicable to satellite broadcasting and cable retransmission ("Satellite and Cable Directive") Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs ("Software Directive"), Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights ("IPRE Directive"), Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases ("Database Directive"), Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights ("Term Directive"), Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on certain permitted uses of orphan works ("Orphan Works Directive"), Directive 2014/26/EU of the European Parliament and of the Council of 26 February 2014 on collective management of copyright and related rights and multi-territorial licensing of rights in musical works for online use in the internal market ("CRM Directive"), Directive (EU) 2017/1564 of the European Parliament and of the

has largely been by way of regulations amending CDPA[61], although the Marrakech Treaty Regulation[62] and the Portability Regulation[63] both had direct effect in UK law. No major immediate changes are proposed to UK copyright law after the UK leaves the European Union by virtue of the Intellectual Property (Copyright and Related Rights) (Amended) (EU Exit) Regulations 2019. Copyright law is national, with international recognition achieved through the Berne Convention.[64]

2.2.2. What does copyright protect?

The UK copyright regime seeks to protect two categories of information/data.

The first is 'original works', which are 'fixed'[65] as a literary, artistic, musical or dramatic work (LMDA).[66] A literary work can include software, but cannot arise in one word, no matter how original as "a literary work is intended to afford either information and instruction, or pleasure, in the form of literary enjoyment."[67] The LMDA works are subject to an 'originality standard'. The criteria for this originality standard has varied in copyright history[68] and across jurisdictions (copyright being a national right that has been subject to harmonisation within the European Union) [69] and it is not the purpose of this paper to analyse this debate.[70] For the purposes of

---

Council of 13 September 2017 on certain permitted uses of certain works and other subject matter protected by copyright and related rights for the benefit of persons who are blind, visually impaired or otherwise print-disabled and amending Directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society ("Marrakech Treaty Directive"), Regulation (EU) 2017/1128 of the European Parliament and of the Council of 14 June 2017 on cross-border portability of online content services in the internal market ("Portability Regulation").

[61] See for example, Copyright and Related Rights Regulations 2003/2498, Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014/1372.

[62] Regulation (EU) 2017/1563 of the European Parliament and of the Council of 13 September 2017 on the cross-border exchange between the Union and third countries of accessible format copies of certain works and other subject matter protected by copyright and related rights for the benefit of persons who are blind, visually impaired or otherwise print-disabled ("Marrakech Treaty Regulation").

[63] Regulation (EU) 2017/1128 of the European Parliament and of the Council of 14 June 2017 on cross-border portability of online content services in the internal market ("Portability Regulation").

[64] Berne Convention for the Protection of Literary and Artistic Works, 1887.

[65] The requirement for recording of the work is specified in s3(2) CDPA and is based on the Berne Convention Article 2(1).

[66] ss1(1)(a) and 3, CDPA.

[67] See *Exxon Corporation v Exxon Insurance* [1982] Ch 119(Ch) quoting *Hollinrake v Truswell* [1894] 3 Ch 420

[68] *University of London Press v University Tutorial Press* [1916] 2 Ch 601 at 608. See also Waisman, A., "Revisiting Originality" (2009) *European Intellectual Property Review* p370.

[69] See Margoni, T., "The Harmonisation of EU Copyright Law: The Originality Standard" (June 29, 2016) published in Perry, M. (ed), *Global Governance of Intellectual Property in the 21st Century* (Springer International Publishing: Switzerland, 2016) pp. 85-105. Also available at SSRN: https://ssrn.com/abstract=2802327. For a US perspective see Ginsburg, J., "No Sweat Copyright and Other Protection of Works of Information and *Feist v Rural Telephone"*, (1992) *Columbia Law Review*, Vol. 92(2), pp338-388.

[70] Although within the literature the standard was of vital importance in relation to the protection of information works prior to the Database Directive in 1996 where the question of whether 'labour' mattered was highly relevant. For a more detailed examination of the originality standard and its harmonisation see

understanding the applicability of the originality standard in relation to TDM it is enough to note that the originality standard is generally accepted to be low, connected to original authorial input rather than novelty.[71] The UK originality standard that sought to protect the author's 'skill, judgment and labour' has, since 1991, been subject to European harmonisation into an author's own 'intellectual creation',[72] and post-Brexit there may be scope for the UK to independently revisit the originality standard again.

Considering TDM, reproducing (whether on or off screen) one data field will trigger copyright issues in relation to the specific data field where that target data field might qualify as an 'original' 'literary work', which will require fact based analysis in each situation – a logistic burden on the TDM community that is not easily resolved. The nature of TDM also usually requires a copy of the whole dataset to be made, meaning that it is necessary to consider more than the copyright position on the target data.[73]

Aside from possible protection for individual literary works sitting in data fields, throughout the copyright literature, difficulties over the qualification of *compilations* of factual information or data for protection have been identified, as explained by Ginsburg:

> Modern copyright comfortably embraces works manifesting a personal authorial presence. Protection depends on whether the work manifests authorial personality, not whether that personality demonstrates either taste or talent. On the other hand, modern copyright encounters far more difficulty accommodating works at once high in commercial value by low in personal authorship. The paradigm for this kind of work and its attendant problems is a compilation of factual information[74]

In the UK in 1944 the House of Lords had considered whether a diary containing tables and information might qualify for copyright protection,[75] with the Lord Chancellor concluding not as "there was no feature of them which could be pointed out as novel or specially meritorious or ingenious from the point of view of the judgment or skill of the compiler; it was not suggested

---

Margoni, T., "The Harmonisation of EU Copyright Law: The Originality Standard" (June 29, 2016) *supra* note 69.

[71] Waisman, A., "Revisiting Originality" (2009) *European Intellectual Property Review* p370.

[72] This was introduced in the Software Directive Article 1(3) in 1991. Similar terms are used in the Database Directive Article 3(1), the InfoSoc Directive Recitals 4, 9 and the Term Directive Article 6 for photographs. The CJEU confirmed this as the applicable standard in relation to copyright in *Football Dataco Ltd v Yahoo! UK Ltd and others* (C-604/10) [2012] 3 WLUK 1. See Margoni, T., "The Harmonisation of EU Copyright Law: The Originality Standard" (2016) *supra* note 69, at para 2.1.2.

[73] See Flynn, S., Geiger, C., Quintais, J., Margoni, T., Sag, M., Guibault, L. & Carroll, M.W., "Implementing User Rights for Research in the Field of Artificial Intelligence: A Call for International Action" (April 20, 2020) *European Intellectual Property Review*, Vol.7, and also *American University, WCL Research Paper No. 2020-12*, Available at SSRN: https://ssrn.com/abstract=3578819.

[74] Ginsburg, J., "Creation and Commercial Value: Copyright Protection of Works of Information", (1990), *Columbia Law Review*, Vol. 90 (7), pp1865-1938, p1866. For those interested in the rationale for protection of different types of work, this paper provides a useful overview of the theories of copyright and the protection of works of low originality in US law.

[75] *Cramp v Smythson* [1944] AC 329.

that there was any element of originality or skill in the order in which the table are arranged."[76] The originality standard in UK law prior to 1996 which required 'skill, judgment and labour' from the author therefore meant that recipes were perceived to be a list of ingredients and a set of instructions which did not trigger protection.[77] However, knitting guides,[78] catalogues[79] and exam papers[80] were held to be protected as literary works or 'compilations'. The question of protection for compilations of factual works was then directly addressed by the reforms in the 1990s that led to specific copyright protection for databases in addition to a new right to protect the data therein (see section 2.3 of this paper). Since 1997 UK law has therefore recognised a specific copyright arising in a database [81] which may itself may be an original work if 'by reason of the selection or arrangement of the contents of the database the database constitutes the author's own intellectual creation.'[82] The stipulation regarding authorial intellectual creation (harmonised across the European Union) requires a qualitative assessment which is distinct from the skill, judgment and labour test that had been applied in the UK. Indeed, in 2012 the CJEU confirmed that skill, judgment and labour may not be sufficient to meet the new author's own intellectual creation test[83] meaning that some of the compilations previously protected by copyright (knitting guides, catalogues and exam papers) may no longer qualify for protection once the test for the author's own intellectual creation is applied. Discomfort with the loss of the investment/labour test was expressed from the inception of the database directive.[84] However, the copyright protection for the selection and arrangement of databases will be of relevance to those engaging in TDM, who will have to consider whether the selection and arrangement of the target data represents the author's own intellectual creation. Specific examples of the application of this copyright are discussed in part 3 of this paper.

[76] Lord Chancellor, Viscount Simon in *Cramp v Smythson [1944] AC 329*.
[77] For discussion of categories of creativity that are not protected by copyright see e.g. Bonadio, E., & Lucchi, N., "How Far Can Copyright Be Stretched? Framing the Debate on Whether New and Different Forms of Creativity can be Protected," (2019) *Intellectual Property Quarterly* Vol 2, 115.
[78] *Bridget Foley Ltd v Ellott* [1982] RPC 433.
[79] *Harpers v Barry Henry & Co* (1892) 20 R 133.
[80] *University of London Press v University Tutorial Press* [1916] 2 Ch 601.
[81] Specified as 'a collection of independent works, data or other materials which (a) are arranged in a systematic or methodical way, and (b) are individually accessible by electronic or other means' - s3A(1) CDPA.
[82] s3A(2) CDPA.
[83] *Football Dataco Ltd v Yahoo! UK Ltd and others* (C-604/10) [2012] 3 WLUK 1.
[84] See for example Chalton, S., "The Amended Database Directive Proposal: A Commentary and Synopsis", (1994) *European Intellectual Property Review,* Vol.3, 94 at p278, who argued that the UK should consider a new right which could be applied generally to literary, dramatic, musical and artistic productions which lack human creativity but which demonstrate sufficiently substantial investment by their makers.

The second category of copyright protected works are the 'related rights' which protect sound recordings,[85] films,[86] broadcasts[87] and the typographical arrangement of published editions.[88] These types of works do not have to be 'original' to qualify for protection. However, for the purposes of TDM, these types of works are likely to have limited relevance.

In all situations, copyright in the UK arises automatically and does not require registration.

2.2.3. Effect of copyright

Copyright has the effect of granting the author[89] exclusive control of the work, and specifically the exclusive rights to control:

- Copying of the work;
- Issuing copies of the work to the public;
- Renting or lending the work to the public;
- Performing the work in public;
- Communicating the work to the public; and
- Making an adaptation of the work.[90]

Conducting any of the above activities without the authorisation of the copyright holder, except where permitted by specific exceptions (see Section 1.2.4 below) constitutes infringement of copyright,[91] which is actionable as a matter of both civil and criminal law in the UK.

The rights granted under s16 CDPA are often referred to as the 'economic' rights associated with copyright. UK law also recognises 'moral rights' for authors of copyright works. These are the right to be identified as the author (or director of a film)[92] and the right to object to derogatory treatment of the work.[93]

[85] s1(1)(b) & s5A CDPA.
[86] s1(1)(b) & s5B CDPA.
[87] s1(1)(b) & s6 CDPA.
[88] s1(1)(c) and s8 CDPA.
[89] According to s9 CDPA, the author is 'the person who creates [the work]' and in respect of a sound recording, the producer, and of a film, the producer and principal director. In respect of original works and films, if they are made by an employee 'in the course of his employment' they are owned by the employer per s11 CDPA.
[90] s16(1) CDPA.
[91] ss17-26 CDPA.
[92] s77 CDPA. This right is not absolute and must be asserted (s78). There are certain exceptions listed in s79 CDPA.
[93] s80 CDPA. Exceptions and qualifications are contained in ss81-82 CDPA.

Collectively referred to as 'copyright', the exclusive rights granted to authors and other copyright holders have relatively long protected periods, shown in Table 2.1 below.

Table 2.1: Length of copyright protection for types of work

Type of work	*Length of copyright protection*
Literary, artistic, musical or dramatic	Lifetime of author plus 70 years.[94]
Sound recording	50 years from the end of the year the recording is made or 70 years from publication/making available[95]
Film	Lifetime of the principal director, author of screenplay, author of dialogue or composer of music, plus 70 years (after the last to die)[96]
Broadcast	50 years from the end of the year the broadcast is made[97]
Typographical arrangement of published editions	25 years from the end of the calendar year in which the edition was first published[98]

In relation to TDM, copyright is a relevant consideration because most (if not all)[99] TDM involves some act of reproduction, whether intransient or permanent, and the exclusive rights granted under section 16 include the right to control reproduction.[100] The scope of copyright protection could include 'pre-processing to standardise materials into machine-readable formats.'[101]

2.2.4. Permitted acts with copyright works

UK copyright law grants exclusive rights and there are limited circumstances in which copyright works can be copied, shared or adapted during the period of protection without the authorisation of the rightsholder. UK law does not have the concept of 'fair use' seen in the United States of America, and any use of a copyright right work is only permitted if it meets

[94] s12 CDPA.
[95] s13A CDPA.
[96] s13B CPDA.
[97] s14 CDPA.
[98] s15 CDPA.
[99] See Geiger, C., Frosio, G., & Bulayenko, O., "Text and data mining in the proposed copyright reform: making the EU ready for an age of big data? Legal analysis and policy recommendations (Legislative Comment)," (2018) *IIC* Vol.49(7), pp814-844, at p818: "*Only TDM tools involving minimal copying of a few words or crawling through data and processing each item separately could be operated without running into potential liability for copyright infringement.*" C.f. Flynn, S., Geiger, C., Quintais, J., Margoni, T., Sag, M., Guibault, L. & Carroll, M.W., 'Implementing User Rights for Research in the Field of Artificial Intelligence: A Call for International Action' (April 20, 2020) *European Intellectual Property Review* 2020, Issue 7, American University, WCL Research Paper No. 2020-12, Available at SSRN: https://ssrn.com/abstract=3578819 at p4: "*Temporary reproductions are made any time a researcher makes a query of a database*.".
[100] S16 CDPA.
[101] Geiger, C., Frosio, G., & Bulayenko, O., "Text and data mining in the proposed copyright reform: making the EU ready for an age of big data? Legal analysis and policy recommendations (Legislative Comment)" (2018) *IIC* Vol.49(7), pp814-844, P818.

a defined list of exceptions listed in ss28-76 CDPA (which harmonise some of the exceptions permitted in Article 5 of the InfoSoc Directive).

In summary, the permitted exceptions in UK law apply to:

- The making of temporary copies[102] as part of technological processes;
- Fair dealing for the purposes of research for non-commercial purposes (with sufficient acknowledgement);[103]
- Copying for the purpose of text and data analysis for non-commercial research;[104]
- Fair dealing for the purpose of criticism, review, news reporting or quotation from the work;[105]
- Fair dealing for the purpose of caricature, parody or pastiche;[106]
- Incidental inclusion;[107]
- Copies for personal use by disabled persons;[108]
- Making accessible copies available through authorised bodies;[109]
- Certain teaching uses for non-commercial purposes;[110]
- Certain copying by librarians and archivists;[111]
- Certain uses 'for the purposes of parliamentary or judicial proceedings' and other public administration purposes;[112]

Additionally, CDPA provides for more specialist exceptions for computer programs, design documents/models, typesetting, anonymous works, orphan works, certain spoken works and public readings, journal abstracts, folksong recordings, artistic works on public display (e.g. sculptures and buildings), advertising art for sale, building reconstruction, and certain broadcasting activity.[113] These will not be considered in depth in this paper.

[102] s28A CDPA.
[103] s29 CDPA.
[104] s29A CDPA.
[105] s30 CDPA.
[106] s30A CDPA.
[107] s31 CDPA.
[108] s31A CDPA.
[109] s31B CDPA.
[110] ss32-36 CDPA.
[111] ss37-43 CDPA.
[112] ss45-50 CDPA.
[113] ss50-76 CDPA.

The introduction of the exception for text and data mining under s29A CDPA was introduced into UK law in 2014.[114] This provides that the making of a copy of a copyright protected work does not infringe copyright law where

> the copy is made in order that a person who has lawful access to the work may carry out a computational analysis of anything recorded in the work for the sole purpose of research for a non-commercial purpose, and
>
> the copy is accompanied by a sufficient acknowledgement (unless this would be impossible for reasons of practicality or otherwise)[115]

The exception specifically prohibits transfer of the copy to another person, onwards sale,[116] or use for any other purpose.[117] As discussed in section 1, this exception is extremely limited in scope and many academics called for the scope of the TDM copyright exception to be reconsidered.[118]

2.2.5. Interpretation of the copyright exceptions in the courts

The case law considering the exceptions set out in sections 28A-50 CDPA is somewhat mixed in quantity. Research using the Thomson Reuters Westlaw Database shows the total number of cases cited[119] for each exception below.

Table 2.2: Number of cases cited for exceptions in ss28-50 CDPA

Section (CDPA)	*Exception for*	*Total cases citing*
s28A	Temporary copies	13[120]
s29	Research & Private Study	9[121]
s29A	Text & Data Mining	0
s30	Criticism, review, quotation and news reporting	34[122]
s30A	Caricature, Parody & Pastiche	0
s31	Incidental Inclusion	6[123]
s31A	Disabled persons: copies of works for personal use	0
s31B	Accessible copies by authorised bodies	0
s31BA	Intermediate copies by authorised bodies	0

[114] Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014/1372 reg.3(2).
[115] s29A(1) CDPA.
[116] s29A(3)-(4) CDPA.
[117] s29A(2) CDPA.
[118] *Supra* notes 36 & 47.
[119] As at 14 December 2020.
[120] Key cases citing this exception were ITV Broadcasting Ltd v TV Catch Up Ltd [2011] EWHC 2977, Newspaper Licensing Agency Ltd v Meltwater Holding BV [2011] EWCA 890, Football Association Premier League Ltd v QC Leisure [2008] EWHC 1411, Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd [2013] UKSC 18.
[121] The key case citing this exception is HM Stationary Office v Green Amps Ltd [2007] EWHC 2755.
[122] See Appendix 1 for the full list.
[123] The key cases citing this exception were Football Association Premier League Ltd v Panini UK Ltd [2003] 4 All ER 1290, Football Association Premier League Ltd v Panini UK Ltd [2002] EWHC 2779(Ch) IPC Magazines Ltd v MGN Ltd [1998] 2 WLUK211.

s32	Illustration for instruction	4[124]
s33	Anthologies for educational use	1[125]
s34	Performance in the course of activities of educational establishment	2[126]
s35	Recording by educational establishments of broadcasts	1[127]
s36	Copyright and use of extracts by educational establishments	3[128]
s36A	Lending of copies by educational establishments	0
s40A	Lending of copies by libraries or archives	0
s40B	Libraries/educational establishments making work available through dedicated terminals	0
s41	Single copies by libraries for other libraries	0
s42	Replacement copies by librarians	0
s42A	Single copies of published works by librarians	0
s43	Single copies of unpublished works by librarians	0
s44	Copy of work to be made for export	1[129]
s44A	Legal deposit libraries	0
s44B	Permitted uses of orphan works	0
s45	Parliamentary and judicial proceedings	9[130]
s46	Royal commissions and statutory enquiries	1[131]
s47	Materials open to public inspection or official register	1[132]
s48	Material communicated to Crown in course of public business	4[133]
s49	Public record keeping	1[134]
s50	Acts done under statutory authority	1[135]

* Note repeal of s28B (private copies?) after R (on app of British Academy of Songwriters, Composers and Authors) v Secretary of State for Business, Innovation and Skills [2015] EWHC 1723

As illustrated in this table, some exceptions have not yet been subject to judicial interpretation. The most debated exception is the exception in section 30 for quotation. Many cases consider more than one exception.

[124] Taylor v Macguire [2013] EWHC 3804 (IPEC), Lucasfilm Ltd v Ainsworth [2011] UKSC 39, Universities UK Ltd v Copyright Licensing Agency Ltd [2001] 12 WLUK 37, Phonographic Performance Ltd v South Tyneside MBC [2001] 1 WLR 400.
[125] Newspaper Licensing Agency Ltd v Marks & Spencer Plc [2001] Ch 257.
[126] Football Association Premier League Ltd v QC Leisure [2012] EWCA Civ 1708 and Football Association Premier League Ltd v QC Leisure [2012] EWHC 108 (Ch).
[127] Lucasfilm Ltd v Ainsworth [2011] UKSC 39.
[128] Dyson v Vax Ltd [2010] EWHC 1923, Universities UK Ltd v Copyright Licensing Agency Ltd [2001] 12 WLUK 373, Newspaper Licensing Agency Ltd v Marks & Spencer Plc [2001] Ch 257.
[129] R V Registered Designs Appeal Tribunal Ex P. Ford Motor Co Ltd [1994] RPC 545.
[130] Key cases citing this exception were BBC Petitioners [2012] HCJ 10, Vitof Ltd v Altoft [2006] EWHC 1678, A v B (Copyright: Diary Pages) [2007] 7 WLUK 934.
[131] National Guild of Removers and Storers Ltd v Milner (t/a Intransit Removals and Storage) [2014] EWHC 670 (IPEC).
[132] National Guild of Removers and Storers Ltd v Milner (t/a Intransit Removals and Storage) [2014] EWHC 670 (IPEC).
[133] ITV Broadcasting Ltd v TV Catchup Ltd [2015] EWCA Civ 204, Lucasfilm Ltd v Ainsworth [2011] UKSC 39, Media CAT Ltd v Adams [2011] EWPCC 10, Navitaire Inc v EasyJet Airline Co Ltd (No.3) [2004] EWHC 1725 (Ch).
[134] University Court of the University of St Andrews v Student Gowns Ltd [2019] CSOH 86.
[135] Kabushi Kaisha Sony Computer Entertainment Inc v Owen (t/a Neo Technologies) [2002] EWHC 45.

At the date of writing this paper, there is no case law on the new exception for text and data mining.

2.2.6. Data commons

In relation to copyright protection it is of course possible that data which might qualify for protection otherwise is available for access and use without liability because the copyright owner has chosen to license on an open basis. The OECD propose 'data commons' as an alternative ownership mechanism along with data trusts.[136] Creative Commons[137] is the most well-known model of open licensing, but other open licensing models exist. For example, in the UK, the Open Government Licence for public sector information[138] and the UK Scholarly Communications licence and Model Policy[139] provide mechanisms for open copyright licensing in those specific contexts, although conditions such as non-commercial use and attribution are often imposed that may present issues to a non-publishing user engaging in TDM.

2.3. The Sui Generis Database Right (SGDR)

Separately from copyright, a specific right may arise in databases, which was created much more recently. This is referred to as the Sui Generis Database Right (SGDR).

2.3.1. *UK Legal regime for SGDR*.

The basis for the SGDR is also statutory, the UK having implemented the 1996 Directive[140] through the Copyright and Rights in Databases Regulations 1997. It is not recognised by any international convention.[141] As the name implies, the SGDR was created as a specific creature of legislation and is applicable as well as any rights arising under copyright or other legal doctrines:[142] "a form

---

[136] OECD, *Data in the Digital Age*, OECD Publishing, Paris, 2019 https://www.oecd.org/going-digital/data-in-the-digital-age.pdf p4.
[137] Creativecommons.org
[138] Available at http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/ The Open Government Licence (OGL) is mandated (with certain exceptions) as the default licence for Crown bodies and recommended for other public sector bodies. It doesn't specifically say that it may not be used by non public bodies, but this is implied. It is compatible with the Creative Commons Attribution License 4.0 and the Open Data Commons Attribution License.
[139] https://ukscl.ac.uk/.
[140] Directive 96/9/EC of 11 March 1996 on the legal protection of databases.
[141] International protection was considered in 1996 but was never taken forward - WIPO Proposal for the Substantive Provisions of the Treaty on Intellectual Property in respect of Databases to be considered by the Diplomatic Conference 30 August 1996 available at https://www.wipo.int/edocs/mdocs/diplconf/en/crnr_dc/crnr_dc_6.pdf . See also Powell, M., "The European Union's Database Directive: An International Antidote to the Side Effects of Feist", (1997) *Fordham International Law Journal* Vol. 20, p1215, Sanks, T., "Database Protection National and International Attempts to Provide Legal Protection for Databases," (1998) *Florida.State University Law Review* Vol. 25, p991.
[142] See e.g. Devici, H.A., "Databases: Is Sui Generis a Stronger Bet than Copyright", (2004) *International Journal of Law & Information Technology*, Vol 12(2), pp178-208. See also Chalton, S., "The Amended

of industrial property which goes further than the present copyright law".[143] The goal of the SGDR has been summarised as "to rescue database producers from the threat of market-destructive appropriations by free-riding competitors who contributed nothing to the costs of collecting or distributing the relevant data."[144]

The SGDR was designed to fill a gap in national legislation – as databases were felt to be 'not sufficiently protected in all Member States' and that even where protected such regimes 'had different attributes.'[145] The SGDR was aimed at preventing 'the unauthorised extraction and/or reutilisation of the contents of a database…whereas the making of databases requires the investment of considerable human technical and financial resources while such databases can be copied or accessed at a fraction of the cost needed to design them independently.'[146]

The introduction of the SGDR was not universally welcomed and the provisions subject to rigorous criticism[147] although welcomed by some.[148] According to Hansucker, the EU had taken "a bold, intellectually honest step forward by adopting a sui generis regime" rather than awarding property-based rights.[149] According it Reichman and Samuelson, it was lacking in "any legal or economic analysis of what investment protection law should consist, or how it should differ from regimes of exclusive property rights"[150] Other mechanisms such as unfair competition law,

---

Database Directive Proposal: A Commentary and Synopsis", (1994) *European Intellectual Property Review* Vol.3, p94.

[143] Cornish, W., "European Community Directive on Database Protection", (1996) *Columbia-VLA Journal of Law and the Arts*, Vol. 21, p1.

[144] Reichman, J.H., & Samuelson, P., "Intellectual Property Rights in Data" (1997) *Vanderbilt Law Review* Vol. 50, 51 at p54.

[145] Recital 1, Directive 96/9/EC of 11 March 1996 on the legal protection of databases.

[146] Recitals 6-7, Directive 96/9/EC of 11 March 1996 on the legal protection of databases.

[147] See e.g. Hughes, J., & Weightman, E., "E.C. Database Protection: Fine Tuning the Commission's Proposal," (1992) *European Intellectual Property Review*, Vol.14, p147, von Simson, C., "Feist or Famine American Database Copyright as an Economic Model for the European Union" (1994) *Brooklyn Journal of International Law,* Vol.20, p729, Bloch, F., "Aspects of the protection of electronic databases and data banks in Europe, in view of the draft Directive. Towards a synthesis of European laws?" (1994*) Illinois Business Law Journal,* Vol. 4, pp457-467, Rosler, D. B., "The European Union's Proposed Directive for the Legal Protection of Databases: A New Threat to the Free Flow of Information", (1995) *High Technology Law Journal*, Vol. 10, p105. See also *Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases, Final Report: A study prepared for the European Commission DG Communications Networks, Content & Technology by JIIP and Technopolis* (European Union, 2018), Reichman, J.H., & Samuelson, P., "Intellectual Property Rights in Data" (1997) *Vanderbilt Law Review* Vol. 50, p51.

[148] Cerina, P., "The Originality Requirement in the Protection of Databases in Europe and the United States," (1993) *IIC* p579, Eisenschitz, T., "The EC Draft Directive on the Legal Protection of Databases – an information scientist's reaction", (1993*) Journal of Information Sciences*, Vol 19(1), pp77-80.

[149] Hansuker, G.M., "The European Database Directive: Regional Stepping Stone to an International Model?" (1997) *Fordham Intellectual Property, Media and Entertainment Law Journal*, Vol. 7, p697.

[150] Reichman, J.H., & Samuelson, P., "Intellectual Property Rights in Data" (1997) *Vanderbilt Law Review* Vol. 50, p51 at p164.

contract and technological protection measures in additional to intellectual property law offer protections that can address the economic goals of the SGDR.[151]

The wording of the Database Directive was suggested to be 'imprecise' as the European "Commission [had] given up the idea of achieving a synthesis of different national approaches."[152] Writing in 1994, Bloch predicted that "editors of databases will suffer the legal consequences of this uncertain situation as will users of the databases,"[153] in particular predicting that the question of substantial or unsubstantial use would "be the subject of innumerable conflicts in future."[154] As a Directive, it was also subject to national implementation.[155]

2.3.2. What does SGDR protect?

For the purposes of the SGDR, a database is defined as

> a collection of independent works, data, or other materials which
>
> (a) are arranged in a systematic or methodical way, and
>
> (b)are individually accessible by electronic or other means.[156]

The terms 'independent works', 'data' and 'other materials' are not defined.[157]

The requirement for the database to be the authors own intellectual creation[158] does not apply to the SGDR; therefore, a database which qualifies for copyright protection may qualify for SGDR, but not all databases protected by the SGDR will be protected by copyright. However, the protections are different, copyright protects the selection or arrangement of the database

---

[151] See Derclaye, E, *The Legal Protection of Databases: A Comparative Analysis* (Edward Elgar Publishing, 2008).

[152] Bloch, F., "Aspects of the protection of electronic databases and data banks in Europe, in view of the draft Directive. Towards a synthesis of European laws?" (1994*) Illinois Business Law Journal,* Vol. 4, pp457-467 at p467.

[153] *Ibid*.

[154] *ibid* p465.

[155] For a comparison of implementation see Gaster, J,. "The EC sui generis right revisited after two years: a review of the practice of database protection in the 15 EU Member States", (2000) *Communications Law, Vol.* 5(3), pp87-98.

[156] Article 2, Directive 96/9/EC of 11 March 1996 on the legal protection of databases and Regulation 6, Copyright and Rights in Databases Regulations 1997.

[157] For further discussion of the drafting of some terms see Derclaye, E., "What is a database? A critical analysis of the definition of a database in the European Database Directive and suggestions for an international definition", (2002) *Journal of World Intellectual Property* Vol.5, p981 and Derclaye, E., "Do sections 3 and 3A of the CDPA violate the Database Directive? A closer look at the definition of a database in the U.K. and its compatibility with European law", (2002) *European Intellectual Property Review,* Vol.24(10), pp466-474.

[158] s3A, Copyright, Designs and Patents Act 1988 (as amended).

from copying, whereas SGDR protects the data from substantial or repeated extractions and reutilisation. [159]

Databases do not have to be electronic to qualify for protection under SGDR.[160]

The wording of the legislation regarding the SGDR has generated case law and controversy across the EU – by 2001 Hugenholtz had identified 25 cases concerning the SGDR.[161] The limitation of the right to 'obtaining, verifying or presenting' the contents of the database has led to a series of CJEU decisions starting in 2005[162] that the act of *creating* the data is not protected by SGDR although later verification or presentation work (subsequent to the creation process) might give a creator the benefit of SGDR.[163]

The question of whether there is a minimum threshold for 'substantial investment' has also not been clarified – the UK legislation specifies:

> "substantial" in relation to any investment, extraction or re-utilisation, means substantial in terms of quantity, quality or a combination of both.

The Directive suggests a minimum threshold in that a 'compilation of several recordings of musical performances on a CD does not come within the scope' of SGDR as 'it does not represent a substantial enough investment' but this leaves questions unanswered. The Directive suggests that 'such investment may consist of the implementation of financial resources and/or the expending of time, effort and energy'[164] and that it is a question of 'the investment of considerable human, technical and financial resources'[165] The Advocate General states that the term 'substantial' must be construed in relative terms, 'first in relation to costs and their redemption and secondly in relation to the scale, nature and contents of the database and

---

[159] Regulation 13, Copyright and Rights in Databases Regulations 1997.
[160] Recital 14, Directive 96/9/EC of 11 March 1996 on the legal protection of databases.
[161] For discussion of non-UK law see Hugenholtz, P., "The New Database Right: Early Case Law from Europe" *Fordham University School of Law Ninth Annual Conference on International IP Law & Policy* New York, 19-20 April 2001 available at https://www.ivir.nl/publications/intellectual-property/database-right/, Hugenholtz, P.B., "Program Schedules, Event Data and Telephone Subscriber Listings under the Database Directive The 'Spin-Off' Doctrine in the Netherlands and elsewhere in Europe", Paper presented at *Fordham University School of Law Eleventh Annual Conference on International IP Law & Policy* New York, 14-25 April 2003, available at https://www.ivir.nl/publicaties/download/spinofffordham.pdf and Derclaye, E., "Recent French decisions on database protection: towards a more consistent and compliant approach with the Court of Justice's case law?" (2012) *European Journal of Law and Technology,* Vol 3(2).
[162] Derclaye, E., "The Court of Justice interprets the database sui generis right for the first time" (2005) *European Law Review* Vol. 30, p420.
[163] Case C-203/02 *British Horseracing Board v William Hill Organisation Ltd* [2004] ECR I-10415 (ECJ), Case C-338/02 *Fixtures Marketing Ltd v Svenska Spel AB* [2004] ECR I-10497, Case C-444/02 *Fixtures Marketing Ltd v OPAP* [2004] ECR I-10549, Case C-46/02 *Fixtures Marketing Ltd v Oy Veikkaus Ab* [2004] ECR I-10365.
[164] Recital 40, Directive 96/9/EC of 11 March 1996 on the legal protection of databases.
[165] Recital 7, Directive 96/9/EC of 11 March 1996 on the legal protection of databases.

the sector to which is belongs'[166] and that 'this is not only investment which have a high value in absolute terms that are protected.'[167] Derclaye reviewed the approaches of different national courts in the EU in 2005 and commented that generally:

> the courts of the Member States seem to favour a low level [of investment] and it is rare that a database does not qualify because the investment is insubstantial. The courts also generally do not go into elaborate reasoning as to whether the investment threshold should be high or low. The issue is therefore not clear-cut in the case law.[168]

The other effect of the case law discussions on the investment criteria is that 'spin off databases' (those created as an additional or ancillary output to other activity) probably do not qualify for protection,[169] although the Advocate General has been clear that the purpose of a database does not determine the unequivocal solution to this question.[170]

2.3.3. Effect of SGDR

The SGDR is a *sui generis* property right, but there are three key differences between the effect of SGDR and the effect of copyright.

Firstly, the protection under SGDR is aimed at the control of the extraction and reutilisation of the data in the database.[171] Unauthorised extraction means 'the permanent and temporary transfer of those contents to another medium'[172] Unauthorised reutilisation means 'making these contents available to the public by any means'.[173] Whilst similar, the effect is not as wide-ranging as the controls given to copyright owners over the licensing and use of their works.

The extent of the right extends to 'all or substantial part' of the database, and 'the repeated and systemic extraction or reutilisation of insubstantial parts of the contents of a database may

---

[166] Opinion of Advocate General Stix-Hackl delivered on 8 June 2004 regarding Case C-338/02 *Fixtures Marketing Ltd v Svenska Spel AB*, para 38.
[167] Opinion of Advocate General Stix-Hackl delivered on 8 June 2004 regarding Case C-338/02 *Fixtures Marketing Ltd v Svenska Spel AB*, para 39.
[168] Derclaye, E., "Database sui generis right: what is a substantial investment? A tentative definition", (2005) *International Review of Intellectual Property and Competition Law*, Vol. 36(1), pp2-30 at p21.
[169] See Hugenholtz, P.B. "Program Schedules, Event Data and Telephone Subscriber Listings under the Database Directive The 'Spin-Off' Doctrine in the Netherlands and elsewhere in Europe", Paper presented at *Fordham University School of Law Eleventh Annual Conference on International IP Law & Policy* New York, 14-25 April 2003, available at https://www.ivir.nl/publicaties/download/spinofffordham.pdf
Derclaye, E., "Databases sui generis right: should we adopt the spin off theory?", (2004) *European Intellectual Property Review*, Vol.26(9), pp402-413, Davison, M.J., & Hugenholtz, P.B., "Football fixtures, horseraces and spin-offs: the ECJ domesticates the database right", [2005] *European Intellectual Property Review*, pp113-118.
[170] Opinion of Advocate General Stix-Hackl delivered on 8 June 2004 regarding Case C-338/02 *Fixtures Marketing Ltd v Svenska Spel AB*, para 41-44.
[171] Regulations 13 &16, Copyright and Rights in Databases Regulations 1997.
[172] Regulation 12(1) Copyright and Rights in Databases Regulations 1997.
[173] Regulation 12(1) Copyright and Rights in Databases Regulations 1997.

amount to the extraction or re-utilisation of a substantial part'.[174] Again, the question of what constitutes a 'substantial part' has been considered by the courts.[175]

Secondly, the protection under SGDR is granted for 15 years only,[176] from the completion of the making of the database, or the making available to the public, whichever is later,[177] which is substantially more limited than the period of time granted to authors of copyright works. However, there is provision for a new term of protection in the event of 'substantial new investment',[178] which might include 'substantial verification of the contents'.[179]

Finally, breach of SGDR is enforceable through the civil courts, whereas copyright carries potential criminal penalties as well as civil enforcement.[180]

Forensic Telecommunications Services Ltd v Chief Constable of West Yorkshire[181] is a case which illustrates the complexity of applying both copyright and SGDR protection to databases as well as the interaction of these rights with the law of confidence. The Claimant had created a confidential list of permanent absolute memory addresses from forensic analysis of mobile phones, which was licensed to the UK security services but not the police force. The court held this list was not the company's own intellectual creation by virtue of the selection and arrangement of its contents and that it was not protected by copyright because the lists would look the same whether copied from the Claimant or obtained from the mobile phones directly. It was not a true compilation as it was not planned and there was no overall design. However, a police officer who had copied the list from a licensee of the Claimant was found to be infringing the SGDR as the court decided that the Claimant had made ' a substantial investment in obtaining and verifying the data' and the second defendant had extracted and re-utilised a substantial part of the contents of that database both quantitatively and qualitatively. Additionally, an equitable duty of confidence arose in the data.

2.3.4. Permitted acts with SGDR protected databases

As with copyright, there are exceptions to the protection granted by SGDR. The UK has not implemented all the exceptions listed in the EU Directive. For example private use of non-

[174] Regulation 16, Copyright and Rights in Databases Regulations 1997.
[175] *Supra* note 159.
[176] Regulation 17(1), Copyright and Rights in Databases Regulations 1997.
[177] Regulation 17(2), Copyright and Rights in Databases Regulations 1997.
[178] Regulation 17(3), Copyright and Rights in Databases Regulations 1997.
[179] Recital 55, Directive 96/9/EC of 11 March 1996 on the legal protection of databases.
[180] ss107-110 Copyright, Designs and Patents Act 1988 (as amended).
[181] *Forensic Telecommunications Services Ltd v Chief Constable of West Yorkshire* [2011] WLUK 243.

electronic databases,[182] or extraction and reutilisation for the purposes of public security or an administrative or judicial procedure[183] have not been incorporated into UK law.[184]

Firstly, it is clear in the legislation that the SGDR does not go so far as to allow control over the use of 'insubstantial' parts of the database where the database has been made available to the public 'in any manner' and contractual terms to that effect are void.[185] Merely 'consulting' a database has been decreed not to be extraction or reutilisation.[186]

Secondly, there are three named exceptions. These are shown in the table below.

Table 2.3: Statutory exceptions to SGDR

When Exception applies	*Right to Extraction*	*Right to Reutilisation*	*Conditions*
'the purpose of illustration for teaching or research and not for any commercial purpose'[187]	Yes	No	• Already available to the public • Fair dealing • Lawful user • Attribution of source
'copying [of a work from the internet] by a deposit library or a person acting on its behalf'[188] Deposit Libraries	Yes	No	• Connection with UK • Compliance with conditions set out in the Legal Deposit Libraries Act 2003
'making of an accessible copy of a work [by a disabled person or authorised body] for the benefit of a Marrakesh beneficiary'[189]	Yes	No	• Personal use of a disabled person • Authorised bodies on a not-for-profit basis
'anything done for the purposes of parliamentary or judicial proceedings or for the purposes of reporting such proceedings'[190]	Yes	Yes	
Royal Commission, Statutory Inquiry or reports from the same[191]	Yes	No	
Database is open to public inspection or is a statutory register[192]	Yes	No	• Factual information • With authority of the 'appropriate person'
Database is open to public inspection[193]	Yes	Yes	• With authority of the 'appropriate person'

[182] Article 9(a) Directive 96/9/EC of 11 March 1996 on the legal protection of databases.
[183] Article 9(c) Directive 96/9/EC of 11 March 1996 on the legal protection of databases.
[184] Regulation 20, Copyright and Rights in Databases Regulations 1997.
[185] Regulation 19(2), Copyright and Rights in Databases Regulations 1997.
[186] Case C-203/02 *British Horseracing Board v William Hill Organisation Ltd* [2004] ECR I-10415 (ECJ)
[187] Regulation 20, Copyright and Rights in Databases Regulations 1997.
[188] Regulation 20A, Copyright and Rights in Databases Regulations 1997 inserted by the Legal Deposit Libraries Act 2003.
[189] Regulation 19, The Copyright and Related Rights (Marrakesh Treaty etc.)(Amendment) Regulations 2018 amending the Copyright and Rights in Databases Regulations 1997.
[190] Schedule 1, para 1, Copyright and Rights in Databases Regulations 1997.
[191] Schedule 1, para 2, Copyright and Rights in Databases Regulations 1997.
[192] Schedule 1, para 3(1), Copyright and Rights in Databases Regulations 1997.
[193] Schedule 1, para 3(2), Copyright and Rights in Databases Regulations 1997.

			• purpose of enabling the contents to be inspected at a more convenient time or place
Database is open to public inspection and the contents 'contain information about matters of general scientific, technical, commercial or economic interest'	Yes	Yes	• With authority of the 'appropriate person' • Purpose of disseminating that information
'contents of a database have in the course of public business been communicated to the Crown for any purpose, by or with the licence of the owner of the database right and a document or other material thing recording or embodying the contents of the database is owned by or in the custody or control of the Crown.'[194]	Yes	Yes	• Applies to Crown only • No prior publication[195]
Database is comprised in public records open to public inspection[196]	Yes	Yes	• Public records • Authority of officer under Public Records legislation
Act is specifically authorised by Act of Parliament[197]	Yes	Yes	• Act of Parliament

The very limited number of exceptions to the SGDR stands in direct contrast to the longer list of exceptions to copyright law. Notably, whilst data in a database might qualify for both copyright and SGDR protection, any exceptions for text and data mining will only apply to the copyright protection, and there is no exception to the SGDR protection available specifically in relation to TDM. In the Technical Review of the exceptions the UK government observed that some respondents had suggested a need for a TDM exception for the SGDR and stated: "*the Government's view is that this existing exception will permit the extraction of whole works if required for text and data mining through the provision for "fair dealing with a substantial part*"."[198] This existing exception (Regulation 20) only applies to use "for the purpose of illustration for teaching or research and not for any commercial purpose" to data already available to the public to lawful users.[199]

[194] Schedule 1, para 4, Copyright and Rights in Databases Regulations 1997.
[195] Schedule 1, para 4(3), Copyright and Rights in Databases Regulations 1997.
[196] Schedule 1, para 5, Copyright and Rights in Databases Regulations 1997.
[197] Schedule 1, para 6, Copyright and Rights in Databases Regulations 1997.
[198] Intellectual Property Office, *Technical Review of Draft Legislation on Copyright Exceptions: Government Response* (2014) available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/308732/response-copyright-techreview.pdf page 13.
[199] Regulation 20, Copyright and Rights in Databases Regulations 1997.

2.3.5. Interpretation of the SGDR exceptions in the courts

As shown in the table below, the exceptions for legal deposit libraries and Marrakesh beneficiaries have been subject to no judicial interpretation, and only two cases have considered the exception under regulation 20 for illustration for teaching and research. As some cases considered more than one exception there are in fact only four cases identified to be relevant to the exceptions to the SGDR. A synthetic table providing further detail about these cases is included in Appendix 1.

Table 2.4: Case law mapped onto the exceptions in Regulation 20 CRDR

Regulation (CRDR)	*Exception for*	*Total cases citing*
Reg 20	Illustration for teaching or research	2[200]
Reg 20A	Deposit libraries	0
Reg 20B	Marrakesh beneficiaries	0
Schedule 1	Exceptions to the SGDR for public administration	
Schedule 1 para 1	Parliamentary and judicial proceedings	0
Schedule 1 para 2	Royal Commissions and statutory enquiries	1[201]
Schedule 1 para 3	Materials open to public inspection or on official register	1[202]
Schedule 1 para 4	Material communicated to the Crown in the course of public business	2[203]
Schedule 1 para 5	Public records	1[204]
Schedule 1 para 6	Acts done under statutory authority	2[205]

However, the cases cited regarding the regulation 20 exception do not deal with interpretation of the exception. In both cases, the judge merely noted that the requirements of the exception under regulation 20 were not met.[206] In *R (on the application of the Office of Communications) v Information Commissioner*[207] (an unsuccessful appeal against a decision requiring OFCOM to disclose information relating to the location and other details of mobile phone base stations protected by SGDR) the extraction and re-utilisation of the database in question was ordered by the court as, whilst intellectual property rights (in this instance SGDR) were a valid reason for

[200] *Forensic Telecommunications Services Ltd v Chief Constable of West Yorkshire* [2011] WLUK 243 and *R (on the application of the Office of Communications) v Information Commissioner* [2009] EWCA Civ 90.
[201] *Sietech Hearing Ltd v Borland* [2003] 2 WLUK 568.
[202] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch).
[203] *Sietech Hearing Ltd v Borland* [2003] 2 WLUK 568 and *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch).
[204] *Sietech Hearing Ltd v Borland* [2003] 2 WLUK 568.
[205] Sietech Hearing Ltd v Borland [2003] 2 WLUK 568 and 77M Ltd v Ordnance Survey Ltd [2019] EWHC 3007 (Ch).
[206] Forensic Telecommunications Services Ltd v Chief Constable of West Yorkshire [2011] WLUK 243, para 120 and R (on the application of the Office of Communications) v Information Commissioner [2009] EWCA Civ 90, para 48.
[207] R (on the application of the Office of Communications) v Information Commissioner [2009] EWCA Civ 90.

non-disclosure under the Environmental Information Regulations 2004,[208] the regulations further allowed for a balancing act of such rights against the public interest. Lord Justice Richards found

> The legislative scheme involves a weighing of pros and cons, with a presumption in favour of disclosure and in the context of a strong legislative policy of promoting access to, and dissemination of, information. Where use of information in breach of intellectual property rights has beneficial as well as adverse consequences, the proposition that only the adverse consequences can be taken into account seems to me to run wholly counter to that scheme.[209]

Whilst *Sietech* gives no consideration to the exceptions, in the most recent case concerning the SGDR in the UK *77M Ltd v Ordnance Survey Ltd*[210] Justice Birss considered the exceptions under Schedule 1 para 3 in relation to materials open to public inspection or on an official register where 77M Ltd had re-utilised and extracted various databases from the UK Land Registries using both manual and 'scraping' techniques to create a database for sale. He referred to these exceptions as the authorised extraction defence (paragraph 3(1)), the time and place shifting defence (paragraph 3(2)) and the general information dissemination defence (paragraph 3(3)).[211]

In this case, 77m had acquired the data from the UK Land Registries but the database right that would be infringed came from other databases – specifically providers of centroid data and address data. Justice Birss concluded that the intention of the exception was 'to provide a defence irrespective of whether the acts were consented to by the database right holder.'[212] The applicability of the Berne three step test was also considered, with the judge concluding that the Berne Convention did not purport to apply to SGDR.[213]

In *77M Ltd v Ordnance Survey Ltd*, whilst the initial downloading (extraction) was held to be authorised by the Land Registries, it was held that '*none of these defences provide that once the contents have been acquired by an authorised initial act of extraction, the user is free to do anything at all with them irrespective of the authorisation of the appropriate person*.'[214] In this case, data had been both scraped and manually downloaded: it was held that 'the relevant act of extraction was the putting of that data to commercial use'[215] in some instances authorised by

[208] Regulation 12(5)(c), Environmental Information Regulations 2004.
[209] R (on the application of the Office of Communications) v Information Commissioner [2009] EWCA Civ 90 Lord Justice Richards at Para 56.
[210] 77M Ltd v Ordnance Survey Ltd [2019] EWHC 3007 (Ch).
[211] 77M Ltd v Ordnance Survey Ltd [2019] EWHC 3007 (Ch) at para 298.
[212] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch) at para 311.
[213] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch) at para 314.
[214] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch) at para 320.
[215] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch) at para 321.

the relevant terms. For one database it was held that 'No re-utilisation was involved because the FAP addresses were not made available to the public.'[216] Justice Birss confirmed that:

> The defence allows the public to take at face value the terms of an authorisation granted by a public body over the use of data that public body has made available on a public database it is responsible for.[217]

Whilst detailed commentary on the exceptions in paragraph 3(2) and (3) was not required, however on the requirements for purpose under those exceptions, Justice Birss observed

> I am not convinced that the fact that a user wanted to operate for a profit must necessarily rule out the idea that they were doing acts for the purpose of disseminating information about matters of general scientific, technical, commercial or economic interest.[218]

Ultimately, this case found in favour of the rightsholders, with 77m able to rely on the exception under paragraph 3(1) of Schedule 1 only in relation to its *manual* downloading of certain data, and found to be in breach of the SGDR through its scraping activities.[219]

2.4. Data protection

2.4.1. UK Legal Regime for personal data.

The legal basis for UK law on data protection is found in the Data Protection Act 2018 which implements and expands on the General Data Protection Regulation 2016 (GDPR).[220] There are additional related protections specifically around marketing in Privacy and Electronic Communications Regulations 2003. As all European legislation has been implemented in the UK, Brexit will not affect data protection law, and an 'adequacy decision' is expected[221] from the EU confirming the UK's eligibility to continue to control and process the personal data of EU citizens post-Brexit.

2.4.2. What does data protection law protect?

Data protection law is aimed primarily at regulating the capture and use ('controlling' and 'processing') of 'personal data'. This is defined as:

> any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an

[216] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch) at para 321.
[217] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch) at para 321.
[218] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch) at para 324.
[219] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch) at para 342.
[220] Regulation 2016/679.
[221] https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en

> online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;[222]

The definition excludes applicability to dead persons[223] and requires the personal data to be capable of identifying a person, meaning that it could apply to small amounts of data as well as data capable of cross-referencing identification. Pseudonymised personal data is captured by the legislation, but anonymised data is not.[224] For the purpose of specifying the requirements as to protection of personal data, additional categories of 'special category' data are specified,[225] as well as criminal conviction data. These categories replace the category of 'sensitive personal data' afforded by the Data Protection Act 1998 and are of particular relevance where TDM is considered in relation to health data.[226]

Special category data are: -

> personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation[227]

The material scope of data protection law is set out in Article 2 GDPR. Individuals carrying on the processing of personal data 'in the course of a purely personal or household activity' are not covered,[228] neither are processing 'by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security.'[229] The 'course of an activity which falls outside the scope of Union law' or which are in matters of common foreign and security policy.[230] Those seeking information on conducting text and data mining in relation to competent authorities carrying out criminal or security enforcement or other matters of foreign and security policy should consult the appropriate specialist legislation as this is out with the scope of this paper.

---

[222] Article 4(1) GDPR.
[223] Recital 27, GDPR.
[224] Recital 26, GDPR.
[225] Article 9, GDPR.
[226] See Comande, G., & Schneider, G., "Regulatory challenges of data mining practices: the case of the never-ending lifecycles of "health data"" (2018) *European Journal of Health Law* Vol.25(3), pp284-307 and Terry, N., "Protecting Patient Privacy in the Age of Big Data", (2012) *UMKC Law Review* Vol. 81, p385 for a U.S. perspective.
[227] Article 9(1) GDPR.
[228] Article 2 (1)(c) GDPR.
[229] Article 2 (1)(d) GDPR.
[230] i.e. as set out in Chapter 2 of Title V of the Treaty on European Union and the Treaty on the Functioning of the European Union 2012/C 326/01.

2.4.3. Effect of data protection law.

The goals of data protection law are two-fold: to grant enforceable rights to individual 'data subjects' (to preserve the fundamental rights of citizens[231]) and to impose restrictions on data controllers and processors in relation to most data processing activities.[232]

It is beyond the scope of this paper to give a detailed summary of the extensive obligations imposed on data controllers and processors by UK data protection legislation. There are six principles on which data protection law is based: -

1. Lawfulness, fairness and transparency – personal data can only be processed where a legal basis exists (as specified in Article 6 GDPR) and in a fair and transparent manner;[233]
2. Purpose limitation – data must be 'collected for specified, explicit and legitimate purposes and not further processed in a manner incompatible with those purposes';[234]
3. Data Minimisation – data must be 'adequate, relevant and limited to what is necessary for the purpose';[235]
4. Accuracy – data must be kept up to date and 'erased or rectified without delay';[236]
5. Storage Limitation – data must not be kept for longer than necessary;[237]
6. Integrity and Confidentiality – data must be 'processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures'[238]

In respect of the question of data or text mining, the second principle listed above presents a clear barrier to accessing and using personal data from another source, particularly when coupled with the obligations in Article 13 GDPR to inform a data subject of the purpose of processing. It is also likely that the last principle above will prevent personal data being accessible for text and data mining purposes without contractual agreements being put in place to preserve the integrity and confidentiality of the personal data.

[231] Recital 1, GDPR.
[232] Recital 10, GDPR.
[233] Article 5(1)(a) GDPR.
[234] Article 5(1)(b) GDPR.
[235] Article 5(1)(c) GDPR.
[236] Article 5(1)(d) GDPR.
[237] Article 5(1)(e) GDPR.
[238] Article 5(1)(f) GDPR.

The starting point with special category data is that processing is prohibited completely,[239] unless the conditions listed in data protection legislation are met.[240]

2.4.4. Permitted acts with personal data?

The starting basis for determining what acts are permitted with 'personal data' are the legal bases set out in Article 6 GDPR and supplemented by the conditions for processing of special category data in Article 9 GDPR. However, specific exemptions to the general rules are also set out in legislation,[241] as well as exclusion from the scope of the law as set out in 1.4.2. These exemptions are very specific and the ability to rely on them often depends on the purpose of processing. The following categories of use may qualify for exemptions if all criteria are met:

- Crime & Immigration
- Taxation
- Law and Public Protection
- Regulation, Parliament and the Judiciary
- Journalism
- Research & Archiving
- Health
- Social Work, Education and Child Abuse
- Corporate finance, management and negotiations
- Confidential references
- Exam marking

The applicability of the exemptions is complex and depends both on satisfying preconditions (such as data protection law being likely to prejudice the purpose of processing) as well as having purposes of processing that meet the criteria. Meeting the exemption criteria will usually only provide an exemption from certain provisions in data protection legislation rather than a blanket exemption from considering the legislation at all. This is illustrated in the table below:

Table 1.5: Breakdown of Exemption for Crime and Taxation (General)

Exemption	Pre-condition	Specified Purposes	Permitted Exemptions
Data Protection Act 2018 Schedule 2, Part 1 (2) 'Crime and taxation general'	Application of the provisions would prejudice the specified purposes	• Prevention or detection of crime • Apprehension or prosecution of offenders	• Article 5 – general principles – where applicable • Article 13(1) – (3) & 14 (1)-(4) – requirement to provide certain information to data subject

[239] Article 9(1) GDPR.
[240] See Article 9 GDPR.
[241] See Schedules 2-4 Data Protection Act 2018.

		• Assessment or collection of a tax or duty	• Article 15(1)-(3) – confirmation of processing and safeguards for third country transfers • Article 16 – data subject's right to rectification • Article 17(1)-(2) – data subject's right to erasure • Article 18(1) – data subject's right to restrict processing • Article 19 – notification obligation regarding rectification or erasure of personal data • Article 20(1)-(2) – data subject's right to data portability • Article 21 (1) – objections to processing • Article 34 (1) & (4) GDPR – communication of personal data breach to data subject

As demonstrated, merely operating in the field of crime or taxation will not entitle a user to carry out the mining of personal data without complying with data protection law and therefore any mining of personal data carries a high burden of familiarisation with the applicable provisions and exemptions. It is beyond the scope of this paper to provide guidance on each individual exemption: the mining of personal data should be undertaken only after the applicability of data protection law and the exemptions have been explored on a case by case basis.

The most applicable exemption to text or data mining activity *per se*, is likely to be Schedule 2 DPA 2018 paras 26-28. These are shown in the table below. This shows that the extent to which the activity is exempt from the data protection provisions varies, with the scope of the exemption for 'special purposes' under paragraph 26 granting exemption from many more provisions of GDPR than the exemptions under paragraphs 27-28.

For those engaging in text and data mining, the scope for those engaged in 'academic' TDM appears slightly wider than those engaging in TDM for scientific or historical research purposes' or 'statistical purposes. However, the ethical standards imposed by academic institutions on their researchers are likely to require voluntary compliance with data protection law as a requirement of research,[242] regardless of the availability of the exemption.

[242] See e.g. JISC (2014) *Data protection and research data* available at https://www.jisc.ac.uk/full-guide/data-protection-and-research-data and JISC (2019) *Data Protection Regulation* available at https://rdmtoolkit.jisc.ac.uk/manage-store-and-preserve/data-protection-regulation/.

Table 1.6: Breakdown of data protection exemptions under paras 26-28 of Schedule 1 to the Data Protection Act 2018

Exemption	Pre-condition(s)	Specified Purposes	Permitted Exemptions	Case law citing
Data Protection Act 2018 Schedule 2, Part 1 para 26	'The controller reasonably believes that the application of those provisions would be incompatible with the special purposes'	'Special purposes': journalism, academic purposes, artistic purposes, literary purposes - Carried out with a view to publication by a person of journalistic, academic, artistic or literary material - The controller reasonably believes that the publication of the material would be in the public interest	- Article 5 (1)-(e) – general principles – where applicable - Article 6 – lawfulness - Article 7 (conditions for consent) - Article 8 (1) & (2) child's consent - Article 9 (special category data) - Article 10 (criminal conviction data) - Article 11(2) – processes not requiring identification - Article 13(1) – (3) & 14 (1)-(4) – requirement to provide certain information to data subject - Article 15(1)-(3) – confirmation of processing and safeguards for third country transfers - Article 16 – data subject's right to rectification - Article 17(1)-(2) – data subject's right to erasure - Article 18(1) – data subject's right to restrict processing - Article 19 – notification obligation regarding rectification or erasure of personal data - Article 20(1)-(2) – data subject's right to data portability - Article 21 (1) – objections to processing - Article 34 (1) & (4) GDPR – communication of personal data breach to data subject - Article 36 (obligation to consult ICO) - Article 44 (general principles of transfers to 3rd countries) - Articles 60-62 (co-operation) - Articles 63-67 (consistency)	0
Data Protection Act 2018 Schedule 2, Part 1 para 27	Application of the provisions would prejudice the specified purposes	'scientific or historical research purposes' or 'statistical purposes;	- Article 15(1)-(3) – confirmation of processing and safeguards for third country transfers - Article 16 – data subject's right to rectification - Article 18(1) – data subject's right to restrict processing - Article 21 (1) – objections to processing	0
Data Protection Act 2018 Schedule 2, Part 1 para 28	Application of those provisions would prevent or seriously impair the achievement of those purposes Personal data must be processed per Article 89 GDPR	'archiving purposes in the public interest'	- Article 15(1)-(3) – confirmation of processing and safeguards for third country transfers - Article 16 – data subject's right to rectification - Article 18(1) – data subject's right to restrict processing - Article 19 – notification obligation regarding rectification or erasure of personal data - Article 20(1) – data subject's right to data portability - Article 21 (1) – objections to processing	0

As shown above, there have been no cases considering the new exemptions under the Data Protection Act 2018.

However, it is worth noting that the 2018 exemptions are very similar to those in s32 of the (now repealed) Data Protection Act 1998 and some cases have considered the interpretation of these exceptions.[243]

2.5. Other protections for data?

In addition to copyright, SGDR and data protection there are other mechanisms for the protection and lawful use of data that need to be considered before engaging in TDM.

2.5.1. Contract

Whilst the protections listed so far in this Section 2 arise automatically and are not recorded, the principle of freedom of contract in the UK allows parties to determine the creation, use and exploitation of information in the same way as other contractual deliverables. Contract and copyright have always been in tension, and contract law can be used to alter the balance of supply and demand that copyright seeks to preserve. [244]

In 2015 case of *Ryanair Ltd v PR Aviation BV*[245] the terms of use on the database owner's website were upheld as a means to protect against data scraping and this has been described as 'back door' protection for otherwise un-protectable information.[246] Whilst there was criticism that property based protections such as SGDR were not necessary[247] or that contract can solve

---

[243] *NT1 v Google* [2018] EWHC 799 (right to be forgotten*), Stunt v Associated Newspapers Ltd* [2017] EWHC 695, *Murray v Express Newspapers Plc* [2007] EWHC 1908 (child's privacy in photograph) and *Campbell v Mirror Group Newspapers Ltd* [2002] EWCA Civ 1373 (drug details published) and s33 was considered by *Christian Institute v Lord Advocate* [2016] UKSC 51 (Scottish legislation on data sharing on children), *Dawson-Damer v Taylor Wessing LLP* [2015] EWHC 2366 (Ch) and *Department of Health v Information Commissioner* [2011] EWHC 1430 (anonymised statistics on late termination not personal data) and *Common Services Agency v Scottish Information Commissioner* [2008] UKHL 47.

[244] Kretschmer, M., Derclaye, E., Favale, M., & Watt, R., *The Relationship between Copyright and Contract Law: A Review commissioned by the UK Strategic Advisory Board for Intellectual Property Policy* (SABIP), 2010, Available at SSRN: https://ssrn.com/abstract=2624945. See also Geiger, C., "The future of copyright in Europe: striking a fair balance between protection and access to information", (2010) *Intellectual Property Quarterly* Vol. 1, pp1-14

[245] *Ryanair Ltd v PR Aviation BV* (C-30/14) EU:C:2015:10; [2015] 2 All E.R. (Comm) 455; [2015] 1 WLUK 181 (ECJ (2nd Chamber)). See also Synodinou, T., "Databases and screen scraping: lawful user's rights and contractual restrictions do not fly together (Case Comment)", (2016) *European Intellectual Property Review* Vol. 38(5), pp312-315

[246] Borghi, M., & Karapapa, S., "Contractual Restrictions on the lawful use of information: sole-source databases protected by the back door?", (2015) *European Intellectual Property Review,* Vol. 37(8), pp505-514. See also Myska, M., & Harasta, J., "Less is More? Protecting Databases in the EU After Ryanair", (2016), *Masaryk University Journal of Law and Technology*, Vol.10(2), p170

[247] Ginsburg, J., "Creation and Commercial Value: Copyright Protection of Works of Information", (1990), *Columbia Law Review*, Vol. 90(7), pp1865-1938, Sheils, P.T., & Penchina, R., "What's all the Fuss About Feist – The Sky Is Not Falling on the Intellectual Property Rights of Online Database Proprietors", (1992) *University of Dayton Law Review*, Vol.17, p563 p585

the legal uncertainty around the protection of data,[248] Weibe argues that the time has not yet come to consider contract law as the exclusive way to protect industrial purposes.[249] However, it is one of a suite of protections that may be available to data producers or owners and one that users should consider carefully.[250]

Contract law is a creature of the common law, and therefore there are limited legal provisions that govern how information or databases might be protected by the law of contract. Aside from consumer protection law, the law of contract has been developed through precedent. Many of the exceptions listed above specify that any attempt to contract out of them is void[251] but often parties deal in databases that are available through the internet through end user licencing terms that specify use. The end user licensing terms were a key consideration in *77M Ltd v Ordnance Survey Ltd.*[252]

For end user consumers the protections of unfair contract doctrines and legislation[253] might potentially offer some comfort but for the majority of those seeking to engage in TDM, consumer protection will not be invoked.

2.5.2. Confidentiality and trade secrets

UK law offers protection for confidential information through the common law, a protection that sits uneasily between concepts of property and contractual rights.[254] Following the enactment of the Human Rights Act 1998 protection from the 'misuse of private information' has also arisen through rights to privacy.[255] The principles of the protection are set out in *Coco v AN Clark Engineers Ltd.*[256] These require:

1. Information to have the necessary quality of confidence;
2. The creation of an obligation of confidence; and
3. Unauthorised use to the detriment of the rightsholder.

---

[248] Koscik, M., & Myska, M., "Database authorship and ownership of sui generis database rights in data-driven research", (2017) *International Review of Law, Computers & Technology*, Vol.31(1), p43.
[249] Wiebe, A., "Protection of industrial data – a new property right for the digital economy?"(2017) *Journal of Intellectual Property Law & Practice,* Vol.12(1), pp62–71.
[250] Spence, W.C., & Pedersen, G.L., "Scraping Electronic Data from websites", 2009 *Copyright World* Vol.193, pp18-21.
[251] See for example, Regulation 20 Copyright and Rights in Databases Regulations 1997/3032. For discussion see Raue, B., "Free flow of data? The friction between the Commission's European data economy initiative and the proposed Directive on Copyright in the Digital Single Market (Editorial)" (2018) *IIC*, Vol. 49(4), pp379-383.
[252] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch).
[253] Unfair Contract Terms Act 1977.
[254] Gurry, F., *Breach of Confidence* (Oxford: Clarendon Press, 1984), pp46-56.
[255] Human Rights Act 1998. *Campbell v Mirror Group Newspapers Ltd* [2004] UKHL 22, paras 11 & 14.
[256] *Coco v AN Clark Engineers Ltd* [1968] FSR 415.

The protection of confidentiality arising in relation to databases has been considered by the UK courts. For example, applying the *Coco* three step test to a scenario where three different companies were utilising horse racing data for betting purposes,[257] Justice Zarcoli found:

1. That confidentiality in race day data arose 'because there is a substantial commercial value in the information;'
2. That a reasonable person would have known that racecourse operators valued the data and therefore that any permitted data collection was not unlimited in scope; and
3. That a party authorised to collect some data had breached its obligation of confidence by collecting additional data and releasing it all to a third party, who then created a database for commercial purposes.

In *Racing Partnership Ltd*, the data had not been copied or scraped from the party authorised to create the 'official' database but collected independently of that database. However, the outcome was detrimental to Racing Partnership Ltd who were then able to successfully act against Sports Information Services Ltd who had continued to sell their database after their exclusive rights of access to the racecourses ended. Miller et al suggest that this case means that "breach of confidence" [can be added] to the list of potential claims that right holders may be able to bring against those who collect and/or receive unofficial data streams from sporting events."[258] Sports databases will be considered in detail in Section 3 of this paper.

The protection of confidential information is not absolute. A defence exists where publication of the information is in the 'public interest', with a court required to carry out a 'balancing act' which weighs the public interest in maintaining protection for confidential information against 'a countervailing public interest favouring disclosure'.[259]

2.5.3. Protection by design

Aside from legal protection, it is also possible that data can be protected by non-legal methods. In 1999, Lessig published his proposition that in cyberspace 'code is law'[260] where he proposed that there were four major regulators of activity online (Law, Norms, Market and Architecture) and that technology would play an ever-greater role in regulating copying online.

257 Racing Partnership Ltd v Done Brothers (Cash Betting) Ltd [2019] EWHC 1156 (Ch); [2019] 3 W.L.R. 779; [2019] 5 WLUK 112 (Ch D).
258 Muller, A., " "Scouting" for unofficial data: SIS pays the price - adding 'breach of confidence' to the list of potential claims by rightsholders" (2019) Computer and Telecommunications Law Review Vol. 25(8), pp187-191.
259 Attorney General v Guardian Newspapers Ltd (No 2) [1988] 3 WLR 776 at 807.
260 Lessig, L., Code and other laws of cyberspace (Basic Books, 1999) now Code 2.0.

'Technological protection measures' (TPMs) are mechanisms by which software controls the ability of users to use or even access copyright protected material. Additionally, rightsholders might impose security and integrity measures to protect the integrity of the data, even if the data is available for consultation. In 2001, the Directive required member states to provide adequate legal protection against the circumvention of effective TPMs.[261]

By 2016, the European Commission took the position that rightsholders "should be allowed to apply measures where there is risk that the security and integrity of the system or databases where the works or other subject-matter are hosted would be jeopardised" but specifies that these measures should "not exceed what is necessary to pursue the objective of ensuring the security and integrity of the system."[262]

The concept of protection by design also appears in the 2016 General Data Protection Regulation, where data controllers and processors are obliged to adopt a data protection by design and default process.[263]

2.5.4. Other lawful use

As illustrated by In *R (on the application of the Office of Communications) v Information Commissioner*[264](discussed in Section 2.3.4) other legislation can open data to public access and lawful use. The Freedom of Information (FOI) legislation in the UK[265] provides a basis for users to request access to information held by public authorities or those providing services for them. Further, Ducato and Strowel argue that the 'transparency' principle in EU consumer law provides a mechanism to permit TDM.[266]

2.6. Summary of the legal protection available to 'data'

Section 2 has explained the types of information or data that are protected by various legal regimes in the UK. These are summarised in the Table below. Before engaging in TDM, the key question to ask is whether the 'data' in question constitutes any of the protected types of data: copyright 'works', a qualifying 'database', 'personal data' or any other type of information that benefits from protection. If so, the next question is whether there are any lawful exceptions that will allow the reproduction and use of the data without the rightsholders permission. Failing that,

[261] Article 6, Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society.
[262] Recital 12, Proposal for a Directive on copyright in the digital single market COM (2016) 593.
[263] Recital 78, Article 25, GDPR.
[264] *R (on the application of the Office of Communications) v Information Commissioner* [2009] EWCA Civ 90.
[265] Freedom of Information Act 2000, Freedom of Information (Scotland) Act 2002.
[266] Ducato, R., & Strowel, S., "Limitations to text and data mining and consumer empowerment: making the case for a right to "machine legibility"" (2019) *IIC* Vol.50(6), pp649-684.

the rightsholder will need to grant a licence, which carries cost and resource implications that may deter TDM activity.

Table 2.6: Summary of legal protection for different types of data

Type of protection	*Type of data protected*	*Length of protection*	*What is restricted?*	*Are there exceptions?*
Copyright	Original works and related works	up to 70 years after death of author	Copyright, distribution, making derivatives of the 'work'	Yes
SGDR	Databases	15 years	Substantial or repeated extraction from qualifying 'database'	Yes
Data protection	Personal information about a living individual	Lifetime of data subject	Storage and processing of 'personal data'	Yes
Contract	Anything specified in contract	As specified in contract	As specified in contract	Not unless permitted in contract

Section 3 will go on the explore how the legal protection and exceptions to the protections outlined in this Section have been interpreted in the UK courts covering specific factual scenarios such as sports fixtures lists, airline pricing, newspaper clippings and collections of works such as maps.

3. What is permitted and prohibited in relation to data use under the current legal regime?

In this section, the applicability of legal protection and exceptions will be considered in a range of factual scenarios where the courts have interpreted the law.

3.1. Sports Databases

The betting industry is dependent on databases and a few cases look at what is permitted and prohibited by UK law in relation to these databases.

From 2001 a series of cases considered the applicability of the SGDR to sports databases. Fixtures Marketing Ltd put considerable investment in time, effort and planning into creating football fixtures lists in the UK that satisfied a number of criteria including ensuring that no team played more than two consecutive away matches, proximate teams did not play at home on the

same date, avoidance of major national events etc.[267] The British Horseracing Board Ltd made a substantial investment in creating a database of race day results. These databases had value for companies engaging in offering fixed odds and pools betting services. However, considering the applicability of SGDR the European Court of Justice held that the SGDR did not arise in the databases because SGDR was not designed to protect situations where the investment was in creation of the database – such as the dates, times and pairings for the football fixtures. In relation to the horse racing lists, the subsequent investment in the 'obtaining, verification and presentation' of the database could not be determined independently from the creation of the data.[268] Whilst these decisions were somewhat surprising and felt to create a lack of clarity for database owners,[269] they were welcomed as providing a better balance of the rights of users and database owners.[270]

However, regardless of the protection of SGDR, the court held in both *Fixtures Marketing Ltd* and *British Horseracing Board Ltd* that infringement of the SGDR would not arise as it was unlikely that repeated and systematic extraction or re-utilisation of substantial parts had occurred – the intrinsic value of parts of data in a database did not make them more substantial as substantiality was determined based on volume.[271]

Having failed to establish SGDR, a new round of case law followed from 2009, considering the applicability of copyright law to sports databases. Whilst Football Dataco Ltd was initially successful in establishing copyright in its UK football fixtures lists based on the "sweat of the brow" judgment and skill exercised by the database creator,[272] the CJEU found against Football Dataco Ltd, holding that the skill and effort in the "selection and arrangement" criteria refer to the structure of the database and not to the creation of the data. The CJEU confirmed that the criterion was "not satisfied when the setting up of the database is dictated by technical considerations, rules or constraints which leave no room for creative freedom". In summary, any

[267] Fixtures Marketing Limited v Ab Svenska Spel [2001] E.C.D.R. 29 and Fixtures Marketing Ltd v Svenska Spel AB (C-338/02) [2004] 11 WLUK 203.
[268] British Horseracing Board Ltd v William Hill Organisation Ltd (C-203/02) and British Horseracing Board Ltd v William Hill Organisation Ltd [2001] EWCA Civ 1268, Fixtures Marketing Ltd v Oy Veikkaus AB (C-46/02) [2004] 11 WLUK 205, Fixtures Marketing Ltd v Organismos Prognostikon Agonon Podosfairou (OPAP) (C-444/02) [2004] 11 WLUK 204.
[269] Davison, M.J., & Hugenholtz, P.B., 'Football fixtures, horseraces and spin-offs: the ECJ domesticates the database right', (2005) European Intellectual Property Review, pp113-118.
[270] E. Derclaye "The Court of Justice interprets the database sui generis right for the first time" (2005) European Law Review Vol.30, p420.
[271] This is also the position taken in *Fixtures Marketing Limited v Ab Svenska Spel* where SGDR was said to be a narrow protection only applicable to the unauthorised copying of "all or large parts of a product" or to "thinly disguised plagiarism" - *Fixtures Marketing Limited v Ab Svenska Spel* [2001] E.C.D.R. 29.
[272] *Football Dataco Ltd v Brittens Pools Ltd* [2009] EWHC 3294 (Ch), *Football Dataco Ltd v Smoot Enterprises Ltd* [2011] EWHC 973(Ch). See also *Forensic Telecommunications Services Ltd v Chief Constable of West Yorkshire* [2011] WLUK 243 (discussed above).

subsisting English copyright rule around originality or creativity was overridden by the harmonisation of copyright in databases.[273]

Racing Partnership Ltd v Done Brothers (Cash Betting) Ltd[274] concerned the collection of 'live' data from attendance at racecourses pursuant to exclusive arrangements with course owners, and the resale of that data and algorithmic 'betting shows' information generated from the collected data to off-course bookmakers. The betting shows data was held not to be copyright protected (the process of generating the output on the betting shows was 'pure routine work' which did not involve sufficient originality) and in any case there was no copying – the shows were consulted by the defendant to ensure its own shows were closer. There was also held not to be any conduct amounting to extraction or re-utilisation of any part of the database, as each betting show was a new entry. In any case, following *British Horseracing Board Ltd,* there would be no reconstitution of the database as a whole or substantial part. Breach of contract did not arise. The sole means of protection of the data was in the confidentiality that arose in it, the court finding that 'a reasonable person in the Tote's position would have known of the steps taken by the claimants to preserve confidentiality in the data and to grant an exclusive right to exploit it for fixed-odds betting purposes.'

Viewing the racing data cases from a TDM perspective, what can these cases offer by way of guidance as to the legitimate scope of TDM? Following *British Horseracing Board*, if the investment in the database is substantially in the creation of the data, then SGDR protection may not subsist. Yet, for the ordinary person or business conducting TDM, how can substantial investment in the database be established? The very making of enquiries into the database owner's investment in a database is likely to incite assertion of the SGDR and the transparency of the database owner's finances may make it difficult to make independent enquiries. Secondly, again following *British Horseracing Board* if the volume of data is very extensive, extracting a small volume of that data may not infringe the SGDR, regardless of the intrinsic value of that data. This is more useful for those engaging in TDM within one large database, but less useful if TDM is implemented against a range of databases, where the data sought may constitute a small volume of one database but a large component of another. Finally, *Racing Partnership Ltd* offers only caution for those seeking to engage in TDM when considering competing databases which have been created independently – might the data be subject to any obligations of confidentiality?

[273] C-604/10 *Football Dataco Ltd v Yahoo! UK Ltd and others,* [2012] 3 WLUK 1.
[274] *Racing Partnership Ltd v Done Brothers (Cash Betting) Ltd* [2019] 3 W.L.R. 779.

Further, the sports database cases constitute data which was not capable of identifying living individuals, and therefore data protection laws were not applicable.

3.2. Price Databases

However, whilst the case law on sports databases suggests that the protection of databases may, in fact, be low, the case of *Ryanair Ltd v PR Aviation BV* (C-30/14)[275] concerns a database where protection lay in the law of contract. This case considered the exceptions in Article 6 of the Database Directive, the rights and obligations of lawful users in Article 8 of the Database Directive and the prohibition on contracting out of these exceptions and rights[276] and concluded that where the database in question was not protected by the Database Directive, then the owner of that database was entitled to place contractual limits on the use of that database, and prohibit scraping of the data.

In the case, PR Aviation had been operating a price comparison website where customers could search the flight data on low cost airlines, including Ryanair, having obtained such data through 'scraping'. Ryanair's terms of use for the website prohibited unauthorised screen scraping and this contractual limitation was upheld by the CJEU.[277] Concerns regarding competition law have been raised, as price comparison websites are seen as performing an important anti-competitive function.[278] The data in question was not personal data, so the applicability of data protection law was not considered.

3.3. Location Databases

A more complex set of scenarios arise in relation to databases that contain location information. The 2007 case of *HM Stationary Office v Green Amps Ltd*[279] concerned Ordnance Survey mapping data, which was subject to Crown copyright. This case held use of the data to infringe the copyright in the data, even where the research was for non-commercial purposes. A year later, *Common Services Agency v Scottish Information Commissioner*[280] considered the use of postcode data, and the applicability of the exceptions in the Data Protection Act 1998 for research and other public interest activities. These cases were followed by *R (on the application*

[275] *Ryanair Ltd v PR Aviation BV* (C-30/14) EU:C:2015:10; [2015] 2 All E.R. (Comm) 455 (ECJ (2nd Chamber))
[276] Article 15, Database Directive.
[277] *Ryanair Ltd v PR Aviation BV* (C-30/14) EU:C:2015:10; [2015] 2 All E.R. (Comm) 455 (ECJ (2nd Chamber)).
[278] Borghi, M., & Karapapa, S., "Contractual Restrictions on the lawful use of information: sole-source databases protected by the back door?" (2015) *European Intellectual Property Review.,* Vol. 37(8), pp505-514.
[279] *HM Stationary Office v Green Amps Ltd* [2007] EWHC 2755.
[280] *Common Services Agency v Scottish Information Commissioner* [2008] UKHL 47.

of the Office of Communications) v Information Commissioner[281] looked at mobile phone base station databases, this time under the Freedom of Information legislation exceptions (which permit consideration of intellectual property protection as discussed in Part 2 above). Finally, as discussed in Part 2 of this paper, *77M Ltd v Ordnance Survey Ltd*[282] gives extensive consideration to the scraping of data from the Land Registry and other location based sources and ultimately, whilst some of the manual data extraction was permissible, the automated 'scraping' was not.

3.4. Databases containing copyright protected 'works'

Finally, there is a small set of case law (mainly from outside the UK) which consider the protection of databases which are collections of more extensive or complex data, potentially consisting of individual works protected by copyright law, in addition to any rights in the databases. However, these cases consider the protection from the point of view of the owner of the compilation rather than of the original works.

The transfer of poems from a university poetry database to a CD was considered in the German case of *Directmedia*.[283] The German court found the professors skill in compiling the database met the criteria for a substantial investment in the obtaining, verification or presentation of the contents. The defendant had exercised independent assessment in selection of the poems for the DS. On referral to the CJEU, the court considered what constituted 'extraction' for the purposes of the Database Directive[284] and confirmed that it had to be understood as referring to any unauthorised act of appropriation, not just to copying and that the exercise of independent, manual judgment in the extraction process did not stop the transfer, so the question for the German court was whether the extraction had been substantial.

The Bulgarian case of *Apis-Hristovich EOOD v Lakorda AD*[285] considered legal databases where the claimant claimed the defendant had extracted material from its database to create a rival one with many similar features, which the defendant claimed was a result of its own efforts. As with *Directmedia* the CJEU confirmed that the concept of 'extraction' was to be interpreted broadly and that temporary or permanent transfers could be considered – so even if the defendant had transformed the data, the court had to consider whether there had been an extraction.

---

[281] *R (on the application of the Office of Communications) v Information Commissioner* [2009] EWCA Civ 90.
[282] *77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch).
[283] *Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg* (C-304/07) [2008] E.C.R. I-7565; [2009] 1 C.M.L.R. 7.
[284] Article 7(2)(a).
[285] *Apis-Hristovich EOOD v Lakorda AD* (C-545/07) [2009] E.C.R. I-1627; [2009] 3 C.M.L.R. 3.

Focusing on copyright law, the UK cases of *Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd*[286] and *Newspaper Licensing Agency Ltd v Meltwater Holding BV*,[287] considered the copying activity taking place when an end user viewed a collection of newspaper clippings (protected by copyright). These cases established the need for an end user (including anyone engaging in temporary reproduction for TDM) to have an end user licence to avoid copyright infringement but the CJEU confirmed that the exception for temporary, transient or incidental copyright would apply to on-screen and cached copies made by the end user when viewing the website.

Given the derivative nature of TDM, it is unlikely that these cases give comfort to anyone engaging in TDM in respect of data that constitutes copyright protected work. Even where copying is temporary, transient or incidental and so copyright protection is not relevant, the broad interpretation of 'extraction' may bring the TDM within the scope of the SGDR (if the database is protected). The transformative nature of TDM is less problematic under US law where two cases concerning literary collections considered the creation of full scale databases of copyright protected works not to constitute infringement, *Authors Guild Inc v HathiTrust*[288] (which considered the GoogleBooks project to digitise entire libraries as databases) and *Perfect 10*[289] established that Google (a search engine that essentially creates a database of websites and images) did not infringe copyright in images by creating a database reproducing the same as thumbnails. The US courts view such activity as 'quintessentially transformative'[290] for the purposes of copyright protection whereas the CJEU considering this type of activity in *Innoweb v Wegener (C-202/12)*[291] held that it was capable of being classed as 're-utilisation' of a protected database.

3.5. Discussion

The scenarios considered in the case law outlined above show the complexity of the legal position for anyone engaging in TDM. Even where permissible under copyright law (as with the new, limited TDM exception), the wide definition of 'extraction', the fact dependent nature of the volume of data extracted, the possibility of contract protection where the database is not

[286] *Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd* [2013] UKSC 18 & *Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd* (C-360/13) [2014] WLR 1025.
[287] *Newspaper Licensing Agency Ltd v Meltwater Holding BV* [2011] EWCA Civ 890.
[288] *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014).
[289] *Perfect 10 v. Amazon.com*. 508 F.3d 1146 (CA9 2007).
[290] *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87 (2d Cir. 2014).
[291] *Innoweb v Wegener* (C-202/12) [2014] Bus. L.R. 308.

SGDR protected, and the 'reasonable person' test for confidentiality all present hurdles for the average user.

Using location information, or any data that could be protected by data protection law, creates more hurdles than purely commercial data scenarios where the data owner is often less successful in relying on intellectual property rights. Typically anonymisation may be both ethical and legal solution for those seeking to use personal data databases however again this illustrates the different standards applicable to those engaging in research or non-commercial purposes verses other purposes as those in academia are often subject to stringent ethical guidelines that require anonymisation as a matter of course.[292] *Department of Health v Information Commissioner* [2011] EWHC 1430 illustrates the use of anonymised abortion statistics which were lawfully published as not containing any personal data but illustrates the importance of true anonymisation.

For anyone engaging in TDM, the use case may be case specific, and the current case law provides limited examples to form a view on whether activity might be lawful. Mindful of the new narratives around a 'data producers right',[293] which may strengthen protection for database owners, the use case for a broader exception for TDM may need to be made to provide clarity to those engaging in TDM.[294]

4. Conclusion

Ascertaining the protection for data in order to carry out text and data mining is no mean feat. This paper has summarised the legal protection available under copyright law, the sui generis database law, data protection law and contract and other creatures akin to contract like confidentiality. As demonstrated, these protections had different rationales and histories, and have different levels of international harmonisation and recognition. Collectively they constitute a patchwork of protection that could be daunting for anyone seeking to engage in TDM.

Specific provision to allow for TDM has been limited to copyright law, and limited in that context to research for a non-commercial purpose. No equivalent exception is available against the sui

[292] *Department of Health v Information Commissioner* [2011] EWHC 1430 interpreting s33 DPA 1998 now Schedule 2 DPA 2018 and *Common Services Agency v Scottish Information Commissioner* [2008] UKHL 47
[293] *Supra* note 22.
[294] See e.g. Margoni, T., & Kretschmer, M., (2018) "The text and data mining exception in the proposal for a Directive on Copyright in the Digital Single Market: why it is not what EU copyright law needs" Blog published at https://www.create.ac.uk/blog/2018/04/25/why-tdm-exception-copyright-directive-digital-single-market-not-what-eu-copyright-needs/ and Rosati, E., "An EU text and data mining exception for the few: would it make sense?" (2018) *Journal of Intellectual Property Law & Practice*, Vol.13(6), pp429-430.

generis database right or in data protection law. TDM has scope and potential beyond both research and the non-commercial sphere, but this is not recognised by the current exception, and therefore those engaging in TDM for purposes not protected by the exception run the risk of infringement of copyright, where this protects the information or database in question, in addition to infringing any SGDR rights or falling foul of contractual provisions or data protection laws.

The multiple layers of legal protection are difficult to navigate, and the case law often highly contextual (for example, the confidentiality obligation in relation to race day data may not translate into other contexts). In addition, the general paucity of case law provides few contexts for users to translate to the wide variety of contexts that big data may be relevant to. As of yet no case law has considered the new copyright exception for TDM.

Appendix 1: Review of the database case law (UK/EU)

Year	*Court*	*Case Citation*	*Type of Data*	*Method of extraction/utilisation*	*Issue for court*	*Legislation cited*	*Summary of decision by court*
2000	Sweden - Tingsratt	*Fixtures Marketing Limited v Ab Svenska Spel* [2000] 4 WLUK 334	Database of football league fixtures	The defendant used the information about the league fixtures to create pools games	Whether there had been a copying of the data to infringe copyright law or an 'extraction' from the database to infringe the SGDR	Article 7(5) of Directive 96/9 & (Swedish) Copyright Act 1993	Swedish 'catalogue' protection was held to apply but, on the facts, there had been no reproduction in the same compilation. See below for appeal and final decision.
2001	England – Court of Appeal (civil division)	*British Horseracing Board Ltd v William Hill Organisation Ltd* [2001] EWCA Civ 1268	Database of horse racing fixtures, including registered horses, jockeys etc	The defendant published information on its website that were alleged to be feeds from the database that were obtained from third party licensees without the right to sublicense.	Whether (a) The database was protected by SGDR; and (b) there had been an 'extraction and reutilisation' of a 'substantial part' of the database to infringe the SGDR.	Article 7(1) of Directive 96/9	The original court decision found that the database was protected by SGDR and that there was unauthorised extraction (which did not have to mean permanent removal) and reutilisation of a substantial part of the database but the appeal court referred this to the CJEU (see below).
2001	Sweden - Hovratt	*Fixtures Marketing Limited v Ab Svenska Spel* [2001] 5 WLUK 84	Database of football league fixtures	The defendant used the information about the league fixtures to create pools games	Whether there had been an 'extraction' from the database to infringe the SGDR	Article 7(1) of Directive 96/9 & (Swedish) Copyright Act 1993	The original decision was upheld as it was not shown that there had been an 'extraction' from the database, but interpretation of Article 7(1) was

							referred to the CJEU (see below).
2003	Scotland – Court of Session (Outer House)	*Sietech Hearing Ltd v Borland* [2003] 2 WLUK 568	Property, documents and computer records – database of customers, CNP cars, audiogram cards	The defendants had copied the database onto floppy disks during employment and then when they started their new rival business had approached the customers in the database	Whether there had been Infringement of copyright; and (a) an 'extraction and reutilisation' of a 'substantial part' of the database(s) to infringe the SGDR; (b) Breach of confidentiality regarding the trade secrets in the data	CDPA 1988, s17(2) CRDR 1997 reg 13(1), 13(2), 15, 16(1)	Copying to floppy disk was copyright infringement. Removal of the CNP cars and audiogram cards from the office was extraction but there was no re-utilisation. Downloading to floppy disks constituted extraction but there was insufficient evidence of re-utilisation. There was also a finding of a duty of confidentiality for trade secrets
2004	CJEU	*Fixtures Marketing Ltd v Oy Veikkaus AB* (C-46/02) [2004] 11 WLUK 205	Database of football league fixtures	The defendant used the information about the league fixtures to create pools games	Whether there had been an 'extraction' from the database to infringe the SGDR	Article 7(1) of Directive 96/9	The court found that the database was not protected by SGDR as the investment had been in creating the data rather than obtaining or verifying the contents.
2004	CJEU	*Fixtures Marketing Ltd v Organismos Prognostikon Agonon Podosfairou (OPAP)* (C-444/02) [2004] 11 WLUK 204	Database of football league fixtures	The defendant used the information about the league fixtures to create pools games	Whether there had been an 'extraction' from the database to infringe the SGDR	Article 7(1) of Directive 96/9	The court found that the database was not protected by SGDR as the investment had been in creating the data rather than obtaining or verifying the contents.

2004	CJEU	*Fixtures Marketing Ltd v Svenska Spel AB* (C-338/02)	Database of football league fixtures	The defendant used the information about the league fixtures to create pools games	Whether there had been an 'extraction' from the database to infringe the SGDR	Article 7(1) of Directive 96/9	The court found that the database was not protected by SGDR as the investment had been in creating the data rather than obtaining or verifying the contents.
2004	CJEU	*British Horseracing Board Ltd v William Hill Organisation Ltd* (C-203/02) [2005] RPC 13	Database of horse racing fixtures, including registered horses, jockeys etc	The defendant published information on its website that were alleged to be feeds from the database that were obtained from third party licensees without the right to sublicense.	The CJEU was asked to explain the scope of Article 7(1) of the Database Directive	Article 7(1) of Directive 96/9	'substantial' was to be measured by volume of data in database; 'extraction/re-utilisation' were any unauthorised act of appropriation and public distribution
2005	England – Court of Appeal (Civil Division)	*British Horseracing Board Ltd v William Hill Organisation Ltd* [2005] EWCA Civ 863	Database of horse racing fixtures, including registered horses, jockeys etc	The defendant published information on its website that were alleged to be feeds from the database that were obtained from third party licensees without the right to sublicense.	Whether (a) The database was protected by SGDR; and (b) there had been an 'extraction and reutilisation' of a 'substantial part' of the database to infringe the SGDR.	Article 7(1) of Directive 96/9	Original decision was overturned. The court found that the database was not protected by SGDR as the investment had been in creating the data rather than obtaining or verifying the contents.
2007	England – Chancery Division	*HM Stationary Office v Green Amps Ltd* [2007] EWHC 2755	Ordnance Survey mapping data	Data was used to create a mapping tool for research for commercial purposes	Whether the defendants use fell within the exception in s29 CDPA (research)	CDPA 1988, s29 Re-use of Public Sector Information Regulations 2005, reg 15	The court found that as the defendant wanted to use the data for a commercial purpose and was not fair dealing so s29 exception did not apply and copyright was infringed.

2008	CJEU	*Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg (C-304/07)* [2009] 1 C.M.L.R. 7	Anthology of German poetry	Inclusion of certain poems from the database in a CD of 1,000 poems	Whether the transfer of material to another database with a manual intermediated process of selection constituted an 'extraction' within Article 7(1) of the Database Directive	Article 7(1) of Directive 96/9	The court found that there could be a transfer of a substantial part of a protected database despite having a manual selection of material.
2008	Scotland – House of Lords	*Common Services Agency v Scottish Information Commissioner* [2008] UKHL 47	Data of details by census wards of all incidents of leukaemia for both sexes	The data was requested under FOI and refused on the basis that it would breach data protection laws.	Whether the data requested was 'personal data' and whether, if so, it would be a breach of data protection law to release it	DPA 1998, s1(1) FOI(S)A 2002 s38(1)(b)	The court found that the information was capable of being personal data and would need to be anonymised.
2009	CJEU	*Apis-Hristovich EOOD v Lakorda AD (C-545/07)* [2009] 3 WLUK 131	Electronic databases containing legal materials	The defendant reportedly had a very similar database to one that they would have had access to when an employee of the claimant	Whether there was capable of being an 'extraction' from a database protected by SGDR	Article 7(1) of Directive 96/9	The court found that extraction took place when the materials were stored in a new medium – the objectives pursued were not relevant, the fact of some material being publicly relevant was not relevant.
2009	England – Court of Appeal (Civil Division)	*R (on the application of the Office of Communications) v Information Commissioner* [2009] EWCA Civ 90	Database of information about mobile phone base stations	The data was requested under FOI and refused on the basis that it would breach environmental information laws, including an exception for SGDR	Whether where there was more than one exception under the Environmental Information Regulations 2004 (including intellectual property rights) they had to be considered together in the public interest balancing exercise or if the intellectual property exception stood outside this	Environmental Information Regulations 2004, reg 12(5)	Intellectual property rights were subject to the public interest balancing test and it was properly open to the court to find that SGDR did not prevent disclosure of the database in the public interest

2009	England – Chancery Division	*Football Dataco Ltd v Brittens Pools Ltd, Football Dataco Ltd v Stan James (Abingdon) Ltd, Football Dataco Ltd v Yahoo! UK Ltd* [2009] EWHC 3294 (Ch)	Football fixtures lists and live data from football matches (see below)	Use of the data for betting purposes allegedly based on independent compilation of the same data from the live matches and presented to customers by hyperlink	The court was asked to consider whether it was lawful to have a trial of the preliminary issue (legal protection of the fixtures lists) that pertained to three cases on the same issue	-	It was lawful for the court to consider protection of the fixtures lists first – the judge had discretion to order this and was not seriously wrong in doing so.
2010	England – Chancery Division	Football Dataco Ltd v Sportradar GmbH [2010] EWHC 2911 (Ch)	Live football data collected at matches about goals, goal-scorers, yellow cards, penalties, substitutions etc collected by analysts paid to attend matches and put into database	Use of the data for betting purposes allegedly based on independent compilation of the same data from the live matches and presented to customers by hyperlink	Whether (assuming SGDR applied) extraction and re-utilisation had to take place in the UK for the English court to be able to hear the case.	Reg 16 CRDR 1997 Art 7(2) Directive 96/6 and Art 3 Directive 2001/29	The end users viewing in the UK would be 'extracting' data. However, transmission to the public was only re-utilisation in the place where the transmission took place. Referral to CJEU (see below)
2010	England – Chancery Division	*Football Dataco Ltd v Brittens Pools Ltd* [2010] EWHC 841 (Ch)	Football fixtures lists created by sending out questionnaires, sequencing and pairing teams etc	Use of the data for betting purposes	Whether the data and database were protected by copyright or SGDR	CDPA 1988 s3, s3A CRDR 1997 reg 12-13	The effort in obtaining, verifying or presenting the data was not sufficient to attract SGDR protection. The intellectual creation in the selection and arrangement of the database did attract copyright protection. The data in the database was not otherwise protected by copyright.

2011	England – Court of Appeal (Civil Division)	*Football Dataco Ltd v Sportradar GmbH* [2011] EWHC Civ 330	Live football data about goals, goal-scorers, yellow cards, penalties, substitutions etc collected by analysts paid to attend matches and put into database	Use of the data for betting purposes allegedly based on independent compilation of the same data from the live matches and presented to customers by hyperlink	Whether copyright or SGDR protected the database if there was no creative skill in compilation and, if it did, whether there had been infringement	Reg 16 CRDR 1997 Art 7(2) Directive 96/6 and Art 3 Directive 2001/29	The court referred the questions on interpretation to the CJEU.
2011	England – Chancery Division	*Forensic Telecommunications Services Ltd v Chief Constable of West Yorkshire* [2011] WLUK 243	Database of permanent memory absolute (PM Abs) addresses for different types of mobile phone	The defendant had compiled his own list of PM Abs and used some of the database in doing so received from a third party with a licence, this was then shared online where others added to the compilation. 32/33 PM Abs addresses were in both databases.	Whether the list was a 'database' or a table/compilation, if a database whether the database was protected by copyright or SGDR. Whether there was a breach of confidence.	CDPA 1988 s3	The court held that there was no skill, judgment or labour in writing down the PM Abs addresses. However, the addresses were systematically arranged and individually accessible so was a database rather than a compilation. It was not protected by copyright, but a substantial investment had been made in obtaining and verifying the data so SGDR protected the database. There had been extraction and re-utilisation of a substantial part and therefore infringement of SGDR.

							There was also a breach of confidence in sharing the data online.
2012	CJEU	*Football Dataco v Yahoo UK (C-604/10)* EU:C:2012:115; [2012] 2 C.M.L.R. 24	Football fixtures lists created by sending out questionnaires, sequencing and pairing teams etc	Use of the data for betting purposes	Whether the database was protected by copyright	Art 3(1) Directive 96/6	The court held that the concepts of selection and arrangement did not extend to creating the data. The intellectual creation criterion was not satisfied if the database was dictated by technological constraints.
2012	England – Chancery Division	*Football Dataco Ltd v Sportradar GmbH, Stan James (Abingdon) Ltd* [2012] EWHC 1185 (Ch)	Live football data about goals, goal-scorers, yellow cards, penalties, substitutions etc collected by analysts paid to attend matches and put into database	Use of the data for betting purposes allegedly based on independent compilation of the same data from the live matches	Whether SGDR protected the database and if it did whether there had been 'extraction'	Reg 16 CRDR 1997 Art 7(2) Directive 96/6	The court held that data collected/recorded at live events was 'obtained' rather than 'created', recording existing facts was not creating new information, so SGDR could apply as there had been substantial investment in this process. There was a likelihood that there had been an extraction of a substantial part of the protected database, infringing the SGDR. Appealed (see below)
2013	CJEU	*Football Dataco v Sportradar (C-173/11)*	Live football data about goals,	Use of the data for betting purposes	Whether (assuming SGDR applied) extraction and re-	Art 7(2) Directive 96/6	Re-utilisation could take place in the UK

		EU:C:2012:642; [2013] 1 C.M.L.R. 29.	goal-scorers, yellow cards, penalties, substitutions etc collected by analysts paid to attend matches and put into database		utilisation had to take place in the UK for the English court to be able to hear the case.	and Art 3 Directive 2001/29	where there was evidence that the website owner intended to target the public there
2013	Court of Appeal (Civil Division)	*Football Dataco Ltd v Sportradar GmbH, Stan James (Abingdon) Ltd* [2013] EWCA Civ 27	Live football data about goals, goal-scorers, yellow cards, penalties, substitutions etc collected by analysts paid to attend matches and put into database	Use of the data for betting purposes allegedly based on independent compilation of the same data from the live matches	Whether SGDR protected the database (since the analysts provided literary commentary rather than the data and so the database owner was creating the data for the database)	Reg 16 CRDR 1997 Art 7(2) Directive 96/6	The appeal was rejected and the original finding that SGDR applied and that there had been infringement was upheld.
2013	UK Supreme Court	*Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd* [2013] UKSC 18	Monitoring reports containing clippings and headlines sourced through software trawling websites looking for search terms	Receipt of the monitoring report by emails allowed the user to click on hyperlinks to access the clippings	Whether the end user needed a licence to receive the services as the clippings/headlines could be copyright protected works and reproduction would occur when opening on screen or in the cache	Art 5(1) Directive 2001/29	The court held that Article 5(1) extended in principle to temporary copies made for the purpose of browsing by an unlicensed end-user (subject to CJEU referral)
2014	CJEU	*Innoweb v Wegener (C-202/12)* EU:C:2013:850; [2014] Bus. L.R. 308	An online collection of car advertisements and list of second-hand cars	The defendant operated a meta search engine for car advertisements that, amongst other things, allowed the end users to search the	Whether the activities were making available to the public a whole or substantial part of the database protected by SGDR.	Art 7(1) Directive 96/9	The court held that re-utilisation had a broad meaning and referred to any unauthorised act of distribution to the public of a whole or substantial part of the

				claimants' database without accessing it from the claimants' website			database. In this instance the activity was re-utilisation and therefore an infringement of SGDR.
2015	CJEU	*Ryanair Ltd v PR Aviation BV* (C-30/14) EU:C:2015:10; [2015] 2 All E.R. (Comm) 455 (CJEU (2nd Chamber))	Website database of flights available, times, prices etc	Operation of an airfare comparison and booking site using the data 'scraped' from the claimant's website	Whether the claimants' database was protected by copyright and/or SGDR and if not, whether the owner was free to set their own contractual limitations on use.	Arts 1(2), 6(1), 8, 15 Directive 96/9	The court held that if the database was not protected by the Database Directive then the provisions barring owners from contracting out of certain uses did not apply and the owner of the database was free to set their own contractual terms for use. In this instance even if copyright applied, there had been no unauthorised reproduction as the site use was the same as normal users. There was insufficient investment in the database to give rise to SGDR.
2019	England - Chancery Division	*Racing Partnership Ltd v Done Brothers (Cash Betting) Ltd* [2019] EWHC 1156 (Ch); [2020] Ch. 289; [2019] 5 WLUK 112 (Ch D)	Horseracing data collected live at racecourses under exclusivity agreements with the racecourse owners	Betting shows information closely related to the prices in the horseracing database	Whether the betting shows infringed any rights held by the claimant.	Reg 16 CRDR 1997	The court held: (a) there was no copyright in the date as there was insufficient skill, judgment or labour in arriving at the information.

							(b) Each betting show was separate from the database entries. No extraction or re-utilisation was involved and even if there was it was insubstantial. No SGDR was infringed. (c) There was no contractual restriction on feeding of live data from the racecourse – the analysts had a statutory right to be there (d) A reasonable person would have realised that there was an obligation of confidentiality and in this instance, it was breached. (e) There was no requisite knowledge of unlawfulness by the defendants to establish conspiracy.

							(f) This case was appealed (see below).
2019	England – Chancery Division	*77M Ltd v Ordnance Survey Ltd* [2019] EWHC 3007 (Ch)	Multiple (18) datasets. The most relevant discussed in the case were UK Land Registries databases and council databases, Post Office address files, various maps and lists of sporting venues and churches	Two competing database shad been created containing geo-spatial co-ordinates of all addresses in the UK which had been created by combining datasets from the UK Land Registries and open access data from a council.	Whether (a) the acts had complied with the terms of the licences; (b) if they had not, whether there was breach of SGDR through substantial extraction and/or reutilisation.	Reg 20 and Schedule 1 para 4 CRDR 1997	In this instance the acts of data scraping breached the terms of most of the licences (multiple licences considered by the court). The court held that transferring data to another medium could constitute extraction that breached SGDR. None of the exceptions applied.
2020	England – Court of Appeal (Civil Division)	*Racing Partnership Ltd v Sports Information Services Ltd* [2020] EWCA Civ 1300	Horseracing data collected live at racecourses under exclusivity agreements with the racecourse owners	Betting shows information closely related to the prices in the horseracing database	Whether (a) there was an unlawful means conspiracy between the defendants and (b) whether the defendant was liable for misuse of confidential information.	-	The court held (a) there was an element of trespass by the analysts collecting live data in respect of those purposes (b) in relation to an obligation of confidentiality the true condition was inaccessibility rather than confidentiality. The appeals were successful.

Appendix 2: Chronology of key events and literature on the protection of data, databases and TDM

1981	Denicola, R.C., "Copyright in Collections of Facts: A Theory for the Protection of Nonfiction Literary Works" (1981) *Columbia Law Review* Vol. 18, p516	Proposed that information should be protected by copyright and that use of databases should be subject to the fair use doctrine.
1982	Gorman, R.A., "Fact or Fancy – The Implications for Copyright – The Twelfth Annual Donald C. Brace Memorial Lecture" (1982) *Journal of the Copyright Society of the USA,* Vol. 29, p560	Summarised the case law to date in US on maps, directories etc
1982	WIPO, Copyright: Monthly Review of the World Intellectual Property Organization (WIPO) 2918, no.9 accessed at https://www.wipo.int/edocs/pubdocs/en/copyright/120/wipo_pub_120_1982_09.pdf	Recommended copyright law used to address issues with data flow
1986	Francione, G.L., "Facing the Nation: The Standards for Copyright, Infringement, and Fair Use of Factual Works" (1986) *University of Pennsylvania Law Review,* Vol. 134, p519	Looks at the US case *Harper & Row, Publishers v Nation Enterprises (The Nation)* and the 'totality' approach taken by courts to factual works as the case limited the fair use doctrine as a cure all and recommended rejecting totality approach.
1987	Sunders, E.M., "Copyright Protection for Compilations of Fact: Does the Originality Standard Allow Protection on the Basis of Industrious Collection" (1987) *Notre Dame Law Review*, Vol.62, p763	Considered the US doctrine of 'industrious collection' and argued that US copyright law seek intellectual rather than mechanical input in compilations to uphold the constitutional aims of copyright law
1987	Hicks, J.B., "Copyright and Computer Databases: Is Traditional Compilation Law Adequate" (1987) *Texas Law Review* Vol.65, p993	Argues that the current copyright regime around the selection, arrangement and 'effort' in creating a database would allow copyright to adequately protected computer databases
1988	Copyright Designs and Patents Act 1988	UK Legislation on copyright
1988	*Green Paper on copyright and the challenge of technology Copyright Issues requiring immediate action* dated 7 June 1988 Catalogue number CB-CO-88-267-EN-C ISSN 0254-1475.	An initial consideration of the key legal issues around data bases and the possibility of a similar right to that for photographers for database operators.
1989	Patterson, L.R., & Joyce, C., "Monopolizing the Law: The Scope of Copyright Protection for Law Reports and Statutory Compilations", (1989) *UCLA Law Review*, Vol. 36, p719-64	A review of the US law on protection of law reports and the cases concerning the electronic legal databases.
1990	Ginsburg, J., "Creation and Commercial Value: Copyright Protection of Works of Information", (1990), *Columbia Law Review,* Vol. 90 (7), pp1865-1938	A good summary of the academic position in the US to 1990 and proposal for collective licensing of works of 'low authorship'
1990	Metaxas, G, "Protection of databases: quietly steering in the wrong direction?", (1990) *European Intellectual Property Review*, Vol.12(7), pp227-234	Argues that sui generis protection should be considered by the European Commission
1991	*Feist Publications, Inc. v. Rural Telephone Service Co* 111 S. Ct. 1282 (1991)	Limitation of the scope of US copyright law to information works

1991	Miller, P.H., "Life after Feist: Facts, the First Amendment, and the Copyright Status of Automated Databases", (1991) *Fordham Law Review,* Vol. 60, p507	Reviews the Feist decision and proposes 2 strategies to protect information – the commerce clause or amendment to unfair competition rule (ref Ginsburg, 1990)
1992	*Proposal for a Council Directive on the Legal Protection of Databases* (EC Commission, Brussels, January 1992)	Draft EU legislation for copyright and SGDR protection for databases
1992	Ginsburg, J., "No Sweat Copyright and Other Protection of Works of Information and Feist v Rural Telephone", (1992) *Columbia Law Review*, Vol. 92(2), pp338-388	Considered the case of Feist and argues that the US courts adopted 'high authorship' and therefore are creating issues regarding protection of database.
1992	Sheils, P.T., & Penchina, R., "What's all the Fuss About Feist – The Sky Is Not Falling on the Intellectual Property Rights of Online Database Proprietors", (1992) *University of Dayton Law Review*, Vol.17, p563	Argues that the Feist decision is not as bad as some argue.
1992	Proposal for a Council Directive on the Legal Protection of Databases (EC Commission, Brussels, January 1992).	Harmonisation of the originality standard The sui generis right
1992	Hughes, J., & Weightman, E., "E.C. Database Protection: Fine Tuning the Commission's Proposal, (1992) *European Intellectual Property Review*, Vol.14, p147	Discussion of the issues with the drafting of the draft Directive
1993	Cerina, P., "The Originality Requirement in the Protection of Databases in Europe and the United States", (1993) *International Review of Intellectual Property and Competition Law*, p579	Discusses the role of originality for database protection
1993	Narayanan, A.S., "Standards of Protection for Databases in the Europeanm Community and the United States: Feist and the Myth of Creative Originality", (1993) *George Washington Journal of International Law & Economics*, Vol.27, p457	Argues that Feist should be reversed and US should consider sui generis protection as set out in EC proposal
1993	Eisenschitz, T., The EC Draft Directive on the Legal Protection of Databases – an information scientist's reaction' (1993) *Journal of Information Sciences*, Vol 19(1), pp77-80	Argues that the Directive is better than the current UK legal position from an information scientist perspective
1994	Chalton, S., "The Amended Database Directive Proposal: A Commentary and Synopsis", (1994) *European Intellectual Property Review,* Vol.3, p94	Review of the text and provisions of the amended Database Directive
1994	von Simson, C., "Feist or Famine American Database Copyright as an Economic Model for the European Union", (1994) *Brooklyn Journal of International Law,* Vol.20, p729	Considers the EU database protection and argues it is anticompetitive or inefficient
1995	Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data	Harmonisation of (personal) data protection law in Europe
1995	Rosler, D. B., "The European Union's Proposed Directive for the Legal Protection of Databases: A New Threat to the Free Flow of Information" (1995) *High Technology Law Journal*, Vol. 10, p105	Argues that the Directive will create transaction costs for databases in Europe and is poorly drafted
1995	Kaye, L. M., "The Proposed E.U. Directive for the Legal Protection of Databases: A Cornerstone of the Information Society?" (1995) *European Intellectual Property Review,* Vol. 12, p583	Consideration of the proposed Database Directive and its provisions and need
1995	Gaster, J., "The EU Council of Ministers' common position concerning the legal protection of databases: a first comment", (1995) *Entertainment Law Review*, Vol. 6(7), p258-262	Considers the amendments to the initial proposals to protect databases in the EU
1996	WIPO Proposal for the Substantive Provisions of the Treaty on Intellectual Property in respect of Databases to be considered by the Diplomatic Conference 30 August 1996 available at https://www.wipo.int/edocs/mdocs/diplconf/en/crnr_dc/crnr_dc_6.pdf	Proposal for international database protection – not implemented.
1996	Beutler, S., "The Protection of Multimedia Products through the European Community's Directive on the Legal Protection of Databases", (1996) *Entertainment Law Review*, Vol.7(8), p31	Considers whether the new SGDR will offer protection to multimedia

1996	Cornish, W., "European Community Directive on Database Protection", (1996) *Columbia-VLA Journal of Law and the Arts*, Vol. 21, p1	Summarises the protection under the Database Directive and its international implications
1996	WIPO Copyright Treaty	Did not mention SGDR
1996	Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases	EU Legislation on database protection
1997	Copyright and Rights in Databases Regulations 1997 (SI 1997/3032)	UK legislation on database protection
1997	Reichman, J.H., & Samuelson, P., "Intellectual Property Rights in Data" (1997) *Vanderbilt Law Review,* Vol. 50, p51	Extensive analysis of the protection of data, criticism of the Database Directive, and two alternative models for protection of data – unfair competition laws and modified liability provisions.
1997	Gaster, J., "The New EU Directive Concerning The Legal Protection Of Databases" (1997) *Fordham International Law Journal* Vol. 20, p1129	Describes the evolution of the Directive and its provisions.
1997	Powell, M., "The European Union's Database Directive: An International Antidote to the Side Effects of Feist", (1997) *Fordham International Law Journal,* Vol.20, p1215	Looks at the Directive provisions and considers their suitability for international expansion – concluding that there is limited evidence and the impact of the SGDR should be assessed first
1997	Hansuker, G.M., "The European Database Directive: Regional Stepping Stone to an International Model?", (1997) *Fordham Intellectual Property, Media and Entertainment Law Journal*, Vol.7, p697	Commended the Database Directive and discussed the economic incentives for database production.
1998	Data Protection Act 1998	UK law on personal data protection
1998	Chalton, S., "The Copyright and Rights in Databases Regulations 1997: some outstanding issues on implementation of the Database Directive" (1998) *European Intellectual Property Review,* Vol. 20(5), p178-182	This article considers only five aspects of the Database Regulations' transposition of the Database Directive into U.K. law, namely: --the consequences of creating a sub-class of databases as protected literary works distinct from other tables and compilations; --the separate and new criterion of originality applied only to databases, and the effect of this provision on the United Kingdom's concept of copyright in computer-generated works; --the application of the new lawful user rights to databases; --the potential conflict between the rights of copyright authors and database makers; and --the underlying problems of protection of database content.
1998	Lai, S., "Database protection in the United Kingdom: the new deal and its effects on software protection" (1998), *European Intellectual Property Review*, Vol.20(1), pp32-35	Critical of the wording of the Database Directive and its effect on protection of software.

1998	Wolken, J.C., "Just the Facts, Ma'am – A Case for Uniform Federal Regulation of Information Databases in the New Information Age", (1998 *) Syracuse Law Review*, Vol. 48, p1263	Proposes a system of database regulation in the US that would include registration and fair use provisions
1998	Austin, W.L. "A Thoughtful and Practical Analysis of Database Protection under Copyright Law, and a Critique of *Sui Generis* Protection" (1997) *Journal of Technology Law & Policy,* Vol.3, p35	Critical review of the US proposals around database directive and the protection against freeriding
1998	Sanks, T., "Database Protection National and International Attempts to Provide Legal Protection for Databases," (1998) *Florida State University Law Review,* Vol.25, p991	Reviewed the EU Directive and the US proposals and was critical of the role of the SGDR and the balance between users and owners
1999	Davison, M.J., "Proposed U.S. database legislation: a comparison with the U.K. Database Regulations" (1999), *European Intellectual Property Review*, Vol 21(6), pp279-284	Discussion of the proposed The Collections of Information Antipiracy Act introduced into the House of Representatives in US in 1999 and comparing the draft with the database regulations in the UK – particularly around exceptions and lawful use.
1999	Adams, J.N., "The Reporting Exception: Does it Still Exist?" (1999) *European Intellectual Property Review,* Vol. 21(8), pp282-285	Expressed concern that the database regulations had no equivalent of the fair dealing for news reporting exception
1999	National Research Council. 1999. *A Question of Balance: Private Rights and the Public Interest in Scientific and Technical Databases*. Washington, DC: The National Academies Press. https://doi.org/10.17226/9692	US research position – critical of the Database Directive and sceptical as to whether statutory protection is needed.
1999	Reichman, J.H., & Uhlir, P.H., "Database Protection at the Crossroads: Recent Development and Their Impact on Science and Technology" (1999) *Berkeley Technology Law Journal* Vol.14, p793	Argues that the EU Directive and proposed US provisions are too strong and could compromise research and innovation
2000	Hugenholz, P., "Implementing the Database Directive" (2000) in Kabel, J.C.K., & Mom, G.J.H., (eds.), *Intellectual Property and Information Law – Essays in Honour of Herman Cohen Jehoram*, (The Hague/Londen/Boston,) p. 183.available at https://www.ivir.nl/publications/intellectual-property/database-right/	Summarises the key provisions and looks at how the database directive was implemented in different EU member states.
2000	Gaster, J,. "The EC sui generis right revisited after two years: a review of the practice of database protection in the 15 EU Member States" (2000) *Communications Law*, Vol. 5(3), pp87-98	Comparison of the data protection enacted in EU Member States
2000	Leistner, M., "The legal protection of telephone directories relating to the new database maker's right",(2000) *International Review of Intellectual Property and Competition Law,* Vol. 31(7/8*),* 950-967	Considers the German case of *Bundesgerichtshof (Tele-Info-CD)*(I ZR 199/96
2001	Hugenholz, P., "The New Database Right: Early Case Law from Europe" *Fordham University School of Law Ninth Annual Conference on International IP Law & Policy* New York, 19-20 April 2001 available at https://www.ivir.nl/publications/intellectual-property/database-right/	Summary of case law on the Database Directive from Belgium, France, Germany, the Netherlands, UK and concludes that parts of the directive are still unclear
2001	Maurer, S., Hugenholtz, P., Onsrud, H., "Europe's Database Experiment" (2001), *Science,* Vol. 294, pp789-790	Suggests US congress should take a "long, hard look" at the drawbacks of the Database Directive due to he high cost, the erosion of the public domain, overprotected synthetic data of doubtful

		worth and new barriers to data aggregation
2001	Porter, H., & Roy, R., "All Bets Are Off! (Case Comment)" (2001) *Entertainment Law Review,* Vol 12(8), pp243-248	Case comment on *British Horseracing Board Ltd*
2001	Bovenberg, J., "Should genomics companies set up databases in Europe? The EU Database Protection Directive Revisited" (2001) *European Intellectual Property Review* p364	Considers the applicability of the SGDR to human genome banks and suggested a Database Right pool to avoid overlapping rights becoming prohibitive to research and development.
2001	Westkamp, G., "Balancing database sui generis right protection with European monopoly control under Article 82 E.C" (2001) *European Competition Law Review,* Vol. 22(1), pp13-24	Considers the balance between SGDR and competition law in relation to information law monopolies following *Radio Telefis Eireann v Commission of the European Communities* (1995)
2002	Karnell, G., "The European sui generis protection of data bases: Nordic and U.K. law approaching the court of the European Communities – some comparative reflections", (2002*) Journal of the Copyright Society of the USA,* p983-1002	Comparative consideration of *Fixtures Marketing v Svenska Spel* and *British Horseracing Board*
2002	Derclaye, E., "Do sections 3 and 3A of the CDPA violate the Database Directive? A closer look at the definition of a database in the U.K. and its compatibility with European law" (2002) *European Intellectual Property Review,* Vol.24(10), pp466-474	Analyses sections 3 and 3A CDPA and the protection available to databases concluding there is the possibility of overlap and 'protection shopping'
2002	Derclaye, E., "What is a database? A critical analysis of the definition of a database in the European Database Directive and suggestions for an international definition", (2002) *Journal of World Intellectual Property* Vol. 5, 981	Review of the definition of 'database' in the Directive
2002	Leistner, M., "Legal Protection for the Database Maker – Initial Experience from a German Point of View", (2003) *International Review of Intellectual Property and Competition Law,* Vol.33, p439	Looked at German case law where protection of telephone directory was upheld based on the SGDR
2002	Grosheide, W., "Database protection, the European way", (2002) *Washington University Journal Law & Policy,* Vol.8, pp39-74	Summarises the scope of the protection of the SGDR
2003	Lipton, J., "Databases as intellectual property: new legal approaches", (2003) *European Intellectual Property Review*, Vol.25(3), pp139-145.	Discussion of the issues with the US and UK approaches being too closely modelled on copyright law
2003	Westkamp, G., "Protecting Databases Under US and European Law – Methodical Approaches to the Protection of Investments Between Unfair Competition and Intellectual Property Concepts", (2003) *International Review of Intellectual Property and Competition Law* Vol.34, pp772-778	Compares the approaches to database protection in the UK and UK and suggests that linking investment and infringement would resolve some of the issues
2003	Hugenholtz, P.B., "Program Schedules, Event Data and Telephone Subscriber Listings under the Database Directive The 'Spin-Off' Doctrine in the Netherlands and elsewhere in Europe", Paper presented at *Fordham University School of Law Eleventh Annual Conference on International IP Law & Policy* New York, 14-25 April 2003, available at https://www.ivir.nl/publicaties/download/spinofffordham.pdf	Discusses interpretation of 'investment' requirement under the Database Directive and the value of the 'spin off doctrine'
2004	National Centre for Data Mining established	-
2004	Derclaye, E., "Databases sui generis right: should we adopt the spin off theory?", (2004) *European Intellectual Property Review*, Vol.26(9), pp402-413	A review of the European Case law on protection of the 'investment' in databases and explanation and analysis of the 'spin off doctrine'

2004	Davies, G., "Database rights and wrongs" (2004) *Scots Law Times* Vol.24, pp147-151	Commentary on the Advocat General's Opinion on the referral from *British Horseracing Board Ltd*
2004	Hugenholtz, P., "Abuse of Database Right: Sole-source information banks under the EU Database Directive" Paper presented at Conference 'Antitrust, Patent and Copyright', École des Mines/UC Berkeley, Paris, January 15-16, 2004 available at https://www.ivir.nl/publicaties/download/abuseofdatabaseright.pdf	Looks at the potential for abuse of the SGDR where a database is the only source for information and remedies against this
2005	Derclaye, E., "Database sui generis right: what is a substantial investment? A tentative definition" (2005) *International Review of Intellectual Property and Competition Law*, Vol. 36(1), pp2-30	Reviews the literature relating to the concept of substantial investment – shows decisions rarely specify what the level of substantiality of an investment is.
2005	Derclaye, E., "The Court of Justice interprets the database sui generis right for the first time" (2005) *European Law Review* Vol.30, p420	Review of the CJEU decision in the *Fixtures* and *British Horseracing* questions and concludes some questions on interpretation answered and some remain.
2005	Davison, M.J., & Hugenholtz, P.B., "Football fixtures, horseraces and spin-offs: the CJEU domesticates the database right", [2005] *European Intellectual Property Review*, pp113-118	Analysis of the CJEU rulings on *British Horseracing Board* and *Fixtures Marketing* and the options for database makers who create their own data.
2005	Leistner, M., "European Union: Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, Arts.7 and 10(3) – "British Horseracing (Case Comment)" (2005) *International Review of Intellectual Property and Competition Law,* Vol.36(5), pp581-595	Case comment on *British Horseracing Board*
2005	European Commission, First evaluation of Directive 96/9/EC on the legal protection of databases, 12 December 2005	-
2006	Kur, A., Hilty, R., Geiger, C., & Leistner, M., "First Evaluation of Directive 96/9/EC on the Legal Protection of Databases – Comment by the Max Planck Institute for Intellectual Property, Competition and Tax Law, Munich", (2006) *International Review of Intellectual Property and Competition Law* Vol.37, pp551-558	Review of the European Commissions evaluation and commentary on their proposals
2007	Derclaye, E., "Intellectual property rights on information and market power – comparing European and American protection of databases" (2007) *International Review of Intellectual Property and Competition Law,* Vol. 38(3), pp275-298	Critical review of US and EU law, and call for a focus on evidence based decision making by European Commission
2007	Allgrove, B., & Ganley, P., "Search engines, data aggregators and UK copyright law: a proposal", (2007) *European Intellectual Property Review* Vol.29(6), pp227-237	Considers the need for a fair dealing exception in UK copyright law to deal with issues around search engines
2007	OECD, *Directorate for Science, Technology and Industry Committee for Information, Computer and Communications Policy Working Party on Information Economy: Participative Web: User-Created Content*, (2007) available at https://www.oecd.org/sti/38393115.pdf	-
2007	OECD, *Giving Knowledge for Free: The Emergence of Open Educational Resources* (2007) available at http://www.oecd.org/education/ceri/38654317.pdf	-
2008	Derclaye, E., *The Legal Protection of Databases: A Comparative Analysis*, (Edward Elgar Publishing, 2008)	Comparative work on database protection internationally
2008	Derclaye, E., "Of maps, Crown copyright, research and the environment (Case Comment)" (2008) *European Intellectual Property Review,* Vol. 30(4), pp162-164	Case comment on *HM Stationery Office v Green Amps Ltd*
2008	Massey, R., "Referee! Illicit device and copyright issues in football broadcasting referred to the European Court of Justice (Case Comment)", (2008) *Entertainment Law Review* Vol.19(8), pp174-177	Case comment on *Football Association Premier League Ltd v QC Leisure*
2008	European Commission, *Green Paper: Copyright in the Knowledge Economy*, COM(2008) 466 final	-

2009	OECD, *Directorate for Science, Technology and Industry Committee for Information, Computer and Communications Policy Counterfeiting and Piracy: Phase II: Piracy of Digital Content* (2009) available at http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/IND(2008)24/REV1&docLanguage=En	-
2009	Larusson, H., "Uncertainty in the scope of copyright: the case of illegal file-sharing in the UK" (2009) *European Intellectual Property Review* Vol.31(3), pp124-134	Reviews copyright law in the context of file sharing and the acts of making available and copying.
2009	Spence, W.C., & Pedersen, G.L., "Scraping Electronic Data from websites", (2009) *Copyright World*, Vol.193, pp18-21	Describes UK and US protections – the SGDT, breach of contract, trespass to personal property and copyright – against web scraping.
2010	Sawdy, C., "High Court decision revisits protection of databases in the United Kingdom – Football Dataco Ltd v Brittens Pools Ltd (Case Comment)", (2010), *Entertainment Law Review,* Vol. 21(6), pp221-224	Commentary on *Football Dataco Ltd v Brittens Pools Ltd*
2010	British Library, *Driving UK Research: Is copyright a help or a hindrance – a perspective from the research community*, 25 July 2010	Series of essays by UK academics on the interplay of copyright with research
2010	Kretschmer, M., Derclaye, E., Favale, M., & Watt, R., *The Relationship between Copyright and Contract Law: A Review commissioned by the UK Strategic Advisory Board for Intellectual Property Policy* (SABIP), 2010, Available at SSRN: https://ssrn.com/abstract=2624945	Consideration of the relationship between copyright and contract law and the role of exceptions in relation to supply and demand
2010	Geiger, C., "The future of copyright in Europe: striking a fair balance between protection and access to information" (2010) *Intellectual Property Quarterly,* Vol. 1, pp1-14	Considered the new technologies upsetting the copyright balance and the possibilities for copyright reform as well as contractual measures
2011	Hargreaves Review	-
2011	HM Government, "Consultation on Copyright" *Intellectual Property Office,* 14 December 2011 **https://www.gov.uk/government/consultations/copyright**	-
2011	James, S., "Based on Ryanair Limited v Billingfleuege.de GmbH", (2011), *E-Commerce Law Reports*, Vol.11(3), pp18-19	Considers the Ryanair Ltd v Billigfluege.de GmbH case and the possible causes for action in relation to owners of data that has been scraped
2011	James, S., "Screen scraping and web harvesting: the legal issues" (2011) E-*Commerce Law & Policy,* Vol.13(6), pp13-15	Explores the protection of databases and the causes of action for scraping
2011	Clark, S., "Just browsing? An analysis of the reasoning underlying the Court of Appeal's decision on temporary copies exemption in *Newspaper Licensing Agency Ltd v Meltwater Holding BV*" (2011) *European Intellectual Property Review* Vol.33(111), pp725-728	Considers the temporary copies exception in copyright law in light of Newspaper Licensing Agency v Meltwater – decision should not have been a surprise
2011	Schermer, B.W., "The limits of privacy in automated profiling and data mining", (2011) *Computer Law & Security Review*, Vol.27(1), pp45-52	Consideration of the risks of automated profiling and TDM and the inadequacy of current laws to provide adequate safeguards for privacy
2011	Guibault, L., "Creative Commons Licences: What to do with the Database Right" (2011) *Computers and Law*, Vol. 21(6), pp22-24	Description of the creative commons licences and treatment of 'works' which excludes the SGDR
2011	Manyika J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A., (2011) *Big data: the next frontier for innovation, competition, and productivity.* McKinsey Global Institute, San Francisco, USA	Considers the scope of big data as a global phenomenon and opportunity
2012	European Commission (2012) Communication on content in the digital single market, COM(2012) 789 final, 18 December 2012	-

2012	McDonald, D., & Kelly, U., JISC (2012) *The Value and Benefit of Text Mining to UK Further and Higher Education. Digital Infrastructure*. Available at: http://bit.ly/jisc-textm Programme: Digital Infrastructure www.jisc.ac.uk/whatwedo/programmes/di_directions.aspx pp21	Report on the potential value of TDM in the UK research industry
2012	Clark, J., *Text mining and scholarly publishing* (2012) Publishing Research Consortium available at https://www.stm-assoc.org/2012_01_01_PRC_Clark_Text_Mining_and_Scholarly_Publishing.pdf	Overview of the process of TDM and key terminology
2012	Derclaye, E., "Recent French decisions on database protection: towards a more consistent and compliant approach with the Court of Justice's case law?" (2012) *European Journal of Law and Technology,* Vol 3(2), p1	Summary of the French case law on SGDR
2012	Stanganelli, M., "Spreading the news online: a fine balance of copyright and freedom of expression in news aggregation" (2012) *European Intellectual Property Review,* Vol.34(11), pp745-753	Looks at the case law on news aggregators and the balance between media providers and aggregators
2012	Haynes, J., "Critically reconceptualising the United Kingdom's fair dealing exception to copyright infringement in light of the government's most recent proposals for reform and lessons learnt from civil law countries" (2012) *European Intellectual Property Review,* Vol.34(12), pp811-814	Critical examination of the UK 'fair dealing' exceptions and looks at alternative options
2012	Terry, N., "Protecting Patient Privacy in the Age of Big Date" (2012) *UMKC Law Review,* Vol. 81, p385	US specific review of the tension between exceptionalism for health privacy and the growth of electronic medical records
2012	Intellectual Property Office, *Impact Assessment BIS0312: Exception for copying of works for use by text and data analytics*, 13th December 2012 available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/308738/ia-exception-dataanalytics.pdf	Review of the potential impact of an exception for TDM
2012	OECD, *New Sources of Growth: Knowledge-based Capital Driving Investment and Productivity in the 21st Century: Interim Project Findings* available at https://www.oecd.org/sti/50498841.pdf	
2012	Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions *The Digital Agenda for Europe – Driving European Growth Digitally*	
2013	OECD, *New Sources of Growth: Knowledge-based Capital – Key Analyses and Policy Conclusions Synthesis Report* available at https://www.oecd.org/sti/inno/knowledge-based-capital-synthesis.pdf	
2013	Pila, J., "Copyright and internet browsing" (2013) *Law Quarterly Review,* Vol.128 (Apr), pp204-208	Looks at Newspaper Licensing Agency and other cases on internet browsing
2013	Dietrich, N., Guibault, L., Margoni, T., Siewicz, K., Spindler, G., & Wiebe, A., "Safe to Be Open: Study on the Protection of Research Data and Recommendations for Access and Usage" (2013) Göttingen, Germany: Universitätsverlag Göttingen, available at: http://eprints.gla.ac.uk/129335	Review of the legal protection of research data and exceptions and licensing models with a view to open access
2013	Smith, J., & Montagnon, R., "Databases hosted outside the UK can infringe rights in UK databases: Football Dataco v Sportradar (C-173/11) (Case Comment)," (2013) *European Intellectual Property Review,* Vol. 35(2), pp111-113	Commentary on Football Dataco v Sportradar
2014	OECD Survey on Open Government Data 2.0 available at http://www.oecd.org/gov/digital-government/2014-open-government-data-survey.pdf	
2014	OECD, Data-driven Innovation for Growth and Well-being: Interim Synthesis Report, October 2014 available at https://www.oecd.org/sti/inno/data-driven-innovation-interim-synthesis.pdf	
2014	Hargreaves, I., Guibault, L., Handke, C., Valcke, P., & Martens, B. (2014) *Standardisation in the area of innovation and technological development, notably in the field of Text and Data Mining: report from the expert group*. (Studies and reports). Publications Office of the European Union. https://doi.org/10.2777/71122	Review the position for TDM and expressed concern that the EU research industry was being held back
2014	Triaille, J.P., de Meeûs d'Argenteuil, J., & de Francquen, A., "Study on the legal framework of text and data mining", 2014, Publications Office of the European Union DOI:10.2780/1475	Review of the definitions and legal framework for TDM
2014	Stokes, S., & Reeves, S., "UK Supreme Court decides web-browing doesn't infringe copyright but nevertheless refers the matter to the CJEU: PRCA Ltd v The NLA (Case Comment)" (2013) *Entertainment Law Review*, Vol.24(5), pp174-176	Commentary on Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd

2014	Cartwright, E., "CJEU condemns the 'scraping' of databases: Innoweb v Wegener (Case Comment)", (2014) *Entertainment Law Review*, Vol.25(5), pp195-197	Commentary on Innoweb v Wegener
2014	Brook, M., Murray-Rust, P., and Oppenheim, C., (2014). "The Social, Political and Legal Aspects of Text and Data Mining (TDM)." D-Lib Magazine, 20(11/12), doi: 10.1045/november2014-brook accessed at https://openaccess.city.ac.uk/id/eprint/4784/1/D-Lib%20Magazine.pdf	Review of the non-technological barriers to text and data mining including legal uncertainty, publisher reluctance and lack of awareness
2014	Hirschey, J, Symbiotic Relationships: Pragmatic Acceptance of Data Scraping (April 1, 2014). *Berkeley Technology Law Journal*, Vol. 29, 2014, Available at SSRN: https://ssrn.com/abstract=2419167 or http://dx.doi.org/10.2139/ssrn.2419167	This paper discusses the current legal regimes surrounding data scraping online. Although these doctrines can be used to protect data, the paper highlights situations when businesses can benefit from working with, instead of against, scrapers.
2014	Hoeren, T., "Big Data and the Ownership in Data: Recent Developments in Europe", (2014) *European Intellectual Property Review* pp751-754	Considers civilian law concepts of property ownership and looks at the emerging rights in data in the UK and Germany
2014	Boulanger, J., Carbonnel, A., De Coninck R.,& Langus, G., "Assessing the economic impacts of adapting certain limitations and exceptions to copyright and related rights in the EU", May 2014, European Commission	Considers digital preservation, e-lending by public libraries, TDM and private use
2014	Truyens, M., & van Eecke, P., "Legal aspects of text mining", (2014) *Computer Law & Security Review,* Vol.30(2), pp153-170	Review of the legal protections of data – contract law may pose issues – the US regime provides more flexibility
2014	Rajaretnam, T., "Data Mining and Data Matching: Regulatory and Ethical Considerations relating to Privacy and Confidentiality in Medical Data", (2014) *Journal of International Commercial Law and Technology* Vol. 9(4), p294	Review of ethical balance between practitioner and data miner in Australia
2014	Intellectual Property Office, *Eight Great Technologies: Big Data* (2014) available at https://www.gov.uk/government/publications/eight-great-technologies-big-data	Report produced by informatics team on big data in patents
2014	Intellectual Property Office, *Technical Review of Draft Legislation on Copyright Exceptions: Government Response* (2014) available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/308732/response-copyright-techreview.pdf	Summarises key points made by respondents on the draft legislation
2014	Copyright and Rights in Performances (Quotation and Parody) Regulations 2014, the Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014, the Copyright and Rights in Performances (Disability) Regulations 2014, and the Copyright and Rights in Performances (Personal Copies for Private Use) Regulations 2014	New exceptions introduced into UK law
2015	Borghi, M., & Karapapa, S., "Contractual Restrictions on the lawful use of information: sole-source databases protected by the back door?" (2015) *European Intellectual Property Review,* Vol. 37(8), pp505-514	Looks at the Ryanair case and whether there are emerging issues with monopolies over information
2015	De Franceschi, A. & Lehmann, M. "Data as tradable commodity and new measures for their protection", (2015) *Italian Law Journal,* Vol. 1, pp51-72	Discusses the balance between the economic trade in data and the personal rights arising
2015	Stannard, E., "A copyright snapshot: the impact of new copyright legislation on information professionals" (2015), *Legal Information Management* Vol.15(4), pp233-239	Looks at the proposed changes to copyright law in 2014
2015	European Commission (2015) *Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A digital single market strategy for Europe*, COM(2015) 192 final, 6 May 2015	
2015	Handke, C., Guibault, L., Vallbé, J.J., "Is Europe falling behind in data mining? Copyright's impact on data mining in academic research" in Schmidt, B., & Dobreva, M., (eds) *New avenues for electronic publishing in the age of infinite collections and citizen science: scale, openness* and trust (IOS Press, Amsterdam, Netherlands, 2015)	Uses bibliometric data to show the increase of data mining research articles and shows that countries where specific

		consent is required have less publications
2015	OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being* OECD Publishing, Paris, 2015 http://dx.doi.org/10.1787/9789264229358-en	
2015	OECD, Enquiries into Intellectual Property's Economic Impact, OECD Publishing, 2015 available at https://www.oecd.org/sti/ieconomy/Chapter7-KBC2-IP.pdf	Review of the economic impact of intellectual property law
2016	The General Data Protection Regulation 2016	EU legislation on data protection.
2016	Proposal for a Directive on copyright in the digital single market COM(2016) 593	
2016	European Commission (2016) Commission staff working document. Impact assessment on the modernisation of EU copyright rules, 14 September 2016, SWD(2016) 301 final	
2016	OECD, The OECD Ministerial Declaration on the Digital Economy: Innovation, Growth and Social Prosperity (the Cancun Declaration), OECD, Paris, 2016 https://www.oecd.org/internet/Digital-Economy-MinisterialDeclaration-2016.pdf	
2016	Synodinou, T., "Databases and screen scraping: lawful user's rights and contractual restrictions do not fly together (Case Comment)" (2016) *European Intellectual Property Review,* Vol. 38(5), pp312-315	Case comment on Ryanair Ltd v PR Aviation BV
2016	Casserly, Z., "The more things change the more they stay the same? An analysis of the new fair dealing provisions in the UK copyright law" (2016) *Computer and Telecommunications Law Review,* Vol.22(5), pp114-119	Considers the various copyright exceptions introduced in the UK in 2014
2016	Zech, H., "A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data", (2016) *Journal of Intellectual Property Law & Practice,* Vol. 11, pp460-470	In light of the Digital Single Market Strategy for Europe, reviews the protection for data and considers the potential for a property like right
2016	Drexl, J., "Designing Competitive Markets for Industrial Data: Between Propertisation and Access" (2016) Max Planck Institute for Innovation and Competition Research Paper No. 16-13 available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2862975	Reviews the competition law framework for ownership and access to data
2016	Pek San Tay & Cheng Peng Sik, "Data mining and copyright: a bittersweet technology gift for copyright owners and the Malaysian public?" (2016) *Computer Law & Security Review*, Vol. 32(6), pp898-906	Malaysian focus on copyright, data mining and database protection
2016	Leung, P., "European Union: copyright legislation – EU leaked doc floats copyright changes for news, data mining", (2016) *World Intellectual Property Review*, Vol.30(9), pp9-10	Note of possible changes to EU law regarding news publishing and copyright
2016	Mowbray, A., Chung, P., & Greenleaf, G., "A free access, automated law citator with international scope: the LawCite project" (2016) *European Journal of Law & Technology*, Vol.7(3)	Specific example of LawCite from Australian Legal Information Institute which uses data mining techniques
2016	Hugenholtz, P.B., "Something Completely Different: Europe's Sui Generis Database Right", in: Frankel, S., & Gervais, D., (eds.), *The Internet and the Emerging Importance of New Forms of Intellectual Property* (Information Law Series, Vol. 37, Kluwer Law International 2016), pp205-222	Chapter on the SGDR
2016	Margoni, T., & Dore, G.,"Why we need a text and data mining exception (but it is not enough)", 1 April 2016. 10th edition of the Language Resources and Evaluation Conference (LREC 2016), Portorož (Slovenia), 23-28 May 2016 available at https://zenodo.org/record/248048#.WXdf2oiGNEY	Recommends a two-step process to TDM – creating an exception not limited to non-commercial purposes, and then European fair use copyright reform
2016	Margoni, T., "The Harmonisation of EU Copyright Law: The Originality Standard" (June 29, 2016) published in Perry, M. (ed), *Global Governance of Intellectual Property in the 21st Century* (Springer International Publishing: Switzerland, 2016) pp. 85-105. Also available at SSRN: https://ssrn.com/abstract=2802327	Review of the international landscape on originality standard and the attempts to harmonise this in the EU
2016	Cocoru, D., & Boehm, M., "An analytical review of text and data mining practices and approaches in Europe", 2016, Open Forum Europe available at https://www.creative-destruction.org/publication/ofe-tdm/	Review different options for TDM in Europe
2016	Myska, M., & Harasta, J., "Less is More? Protecting Databases in the EU After Ryanair" (2016), *Masaryk University Journal of Law and Technology*, Vol.10(2), pp170	Considered the protection of databases after the Ryanair cases and the role of contract law in giving 'sole source'

		databases legal protection – does it render the SGDR unnecessary?
2017	Hugenholtz, P.B., "Data Property: Unwelcome Guest in the House of IP" in J. Reda (Ed.), *Better Regulation for Copyright : Academics meet Policy Makers*: Wed 6 Sept 2017 15:00-18:30 : European Parliament, Room AP 1G3 : University of Southampton, MEP Julia Reda, The Greens EFA (pp. 65-77). The Greens EFA in the European Parliament. https://juliareda.eu/events/better-regulation-for-copyright/ available at https://www.ivir.nl/publicaties/download/Data_property_Muenster.pdf	Makes the case against introducing a data property right
2017	European Copyright Society (2017) General opinion on the EU copyright reform package, 24 January 2017 available at https://europeancopyrightsocietydotorg.files.wordpress.com/2015/12/ecs-opinion-on-eu-copyright-reform-def.pdf	Generally supportive of the package – concerns around ambition, TDM, platform liability, reprobel article, and fair compensation
2017	European Commission Communication concerning the building of a European data economy	
2017	European Commission, Open Science Monitoring – Impact Case Study – The Social Science Open Access Repository	
2017	Wiebe, A., "Protection of industrial data – a new property right for the digital economy?" (2017) *Journal of Intellectual Property Law & Practice,* Vol.12(1), pp62–71	Considers the ability to trade in data contractually without an exclusive legal protection and suggests that there is not yet a case for this
2017	Zech, H., "Building a European Data Economy" (2017), *International Review of Intellectual Property and Competition Law*, Vol.48(5), pp501-503	Discusses the EC Communication 2017 and the possibility of the data producers right
2017	Thorburn, R., Stalla-Bourdillon, S., & Rosati, R. 'iCLIC data mining and data sharing workshop: the present and future of data mining and data sharing in the EU,' (2017) *Computer Law & Security Review*, Vol.33(1), pp129-137	Report on workshop on Data Mining and Data Sharing in the EU Southampton, UK, 23 September 2016
2017	Geiger, C., Frosio, G., & Bulayenko, O., "Opinion of the CEIPI on the European Commission's Proposal to Reform Copyright Limitations and Exceptions in the European Union" Centre for International Intellectual Property Studies Research Paper No. 2017-09 available at http://infojustice.org/wp-content/uploads/2017/10/Opinion-of-the-CEIPI-on-the-European-Commissions-Proposal-to-Reform-Copyright-Limitations-and-Exceptions-in-the-European-Union.pdf	Supportive of the proposals but recommends expansion
2017	Hilty, R.M., & Richter, H., (2017) "Text and data mining" in Hilty R.M., Moscon, V., (eds) *Modernisation of the EU copyright rules. Position Statement of the Max Planck Institute for Innovation and Competition*, Max Planck Institute for Innovation and Competition Research Paper No. 17-12, available at https://pure.mpg.de/rest/items/item_2470998_12/component/file_2479390/content	Part of a more extensive response to the European Commissions proposals for copyright reform – in principle TDM exception welcomed but detailed consideration of the issues with the proposal and an alternative put forward
2017	Koscik, M., & Myska, M., "Database authorship and ownership of sui generis database rights in data-driven research" (2017) *International Review of Law, Computers & Technology*, Vol.31(1), p43	Considers the rules on jointly owned data and proposes contract law as a possible solution to legal uncertainty
2017	Höppner, T., Kretschmer, M., Xalabarder, R., "CREATe public lectures on the proposed EU right for press publishers", (2017) *European Intellectual Property Review* Vol. 39(10), pp.607-622	
2018	The Data Protection Act 2018	UK legislation. Repealed the Data Protection Act 1998 and introduced GDPR into UK law.
2018	European Commission, Evaluation of Directive 96/9/EC on the legal protection of databases, 25th April 2018	Suggested that the SGDR had no proven impact on the overall production of databases or the competitiveness of the industry and that there was an

		appropriate balance between makers and users
2018	European Commission. Commission Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information	-
2018	Barbero, M., Cocoru, D., Graux, H., Hillebrand, A., Linz, F., Osimo, D., Siede, A., Wauters, P., *Study on emerging issues of data ownership, interoperability, (re-)usability and access to data, and liability*, 25 April 2018. Available at https://ec.europa.eu/digital-single-market/en/news/study-emerging-issues-data-ownership-interoperability-re-usability-and-access-data-and	Considers the legal, technical and other types of barriers which currently prevent the full deployment of the European Data Economy and which limit Business to Business data sharing and re-use in Europe.
2018	Hoeren, T. & Bitter, P., "Data ownership is dead: long live data ownership" (2018) *European Intellectual Property Review* Vol. 40(6), pp347-348	Considers the relevance of the property right to individual data ownership, with focus on Germany
2018	Geiger, C., Frosio, G., & Bulayenko, O., "Text and data mining in the proposed copyright reform: making the EU ready for an age of big data? Legal analysis and policy recommendations (Legislative Comment)" (2018) *International Review of Intellectual Property and Competition Law,* Vol.49(7), pp814-844	Considers the proposed exception for text and data mining – welcomes proposal but recommends extending the rights
2018	Rosati, E., "An EU text and data mining exception for the few: would it make sense?" (2018) *Journal of Intellectual Property Law & Practice* Vol.13(6), pp429-430	Consideration of the new TDM exception and recommendation that it should not be restricted to research organisations
2018	Raue, B., "Free flow of data? The friction between the Commission's European data economy initiative and the proposed Directive on Copyright in the Digital Single Market (Editorial)" (2018) *International Review of Intellectual Property and Competition Law*, Vol.49(4), pp379-383	Consideration of the new TDM exception and recommends that commercial research is allowed, and that the provisions regarding contractual circumvention and deletion are reconsidered
2018	Comande, G., & Schneider, G., "Regulatory challenges of data mining practices: the case of the never-ending lifecycles of "health data"' (2018) *European Journal of Health Law*, Vol.25(3), pp284-307	Considers the interaction of GDPR with data mining and suggests that data mining challenges privacy – further consideration is required
2018	Margoni, T., "Artificial Intelligence, Machine Learning and EU copyright law: Who Owns AI?" (10 November 2018) CREATe Working Paper 2018/12 available at https://www.create.ac.uk/artificial-intelligence-machine-learning-and-eu-copyright-law-who-owns-ai/	Considers EU copyright in relation to 'data capturing' and the question of 'ownership' of AI and concludes the EU is falling behind.
2018	Margoni, T., & Kretschmer, M., (2018) *The text and data mining exception in the proposal for a Directive on Copyright in the Digital Single Market: why it is not what EU copyright law needs*. https://www.create.ac.uk/blog/2018/04/25/why-tdm-exception-copyright-directive-digital-single-market-not-what-eu-copyright-needs/	Review of article 3 on TDM – what is needs is a broad and flexible EU wide exception that covers future technological developments
2018	Doldirina, C., et al. "Legal Approaches for Open Access to Research Data." LawArXiv, 1 Apr. 2018. Web. https://doi.org/10.31228/osf.io/n7gfa	Reviews and assesses legal mechanisms to facilitate the open sharing and use of data in support of achieving public-interest societal objectives.
2018	Taylor, L., "The Ethics of Big Data as a Public Good: Which Public? Whose Good?." LawArXiv, 19 June 2018. Web. https://doi.org/10.1098/rsta.2016.0126	Looks at the 'responsible data' movement and the argument for 'data as a public good'

2018	Veale, M., & Edwards, L., "Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling" (forthcoming)(2018) *Computer Law and Security Review*, Vol. 34(2), p398	-
2018	Working Party on Security and Privacy in the Digital Economy OECD Expert Workshop on Enhanced Access to Data: Reconciling Risks and Benefits of Data Re-use Summary of the discussions 15-16 May 2018	-
2019	Campbell, F., 'Data scraping – what are the privacy implications" (2019) *Privacy & Data Protection*, Vol.20(1), pp3-5	Considers the practical implications under GDPR for organisations seeking to engage in data scraping activities
2019	Carroll, M., "Copyright and the Progress of Science: Why Text and Data Mining is Lawful" (2019) *University of California Davis Law Review,* Vol.53, p893	Considers the US position on TDM – arguing that the fair use doctrine permits researchers to conduct TDM.
2019	Ducato, R., & Strowel, A., "Limitations to text and data mining and consumer empowerment: making the case for a right to "machine legibility"" (2019) *International Review of Intellectual Property and Competition Law,* Vol.50(6), pp649-684	Looks at smart disclosure systems and data mining
2019	Muller, A., Hoy, R., & Fitzpatrick, N., "'Scouting" for unofficial data: SIS pays the price – adding 'breach of confidence' to the list of potential claims by rightsholders" (2019) *Computer and Telecommunications Law Review* Vol. 25(8), pp187-191	Commentary on Racing Partnership Ltd v Done Brothers
2019	Stepanov, I., "Introducing a property right over data in the EU: the data producer's right – an evaluation" (2019) *International Review of Law, Computers and Technology* Vol.34(1), pp65 – 86.	Discusses the nature and justifications for the data producers right
2019	Sag, M., "The New Legal Landscape for Text Mining and Machine Learning" (2019) *Journal of the Copyright Society of the USA*, Vol.66, p3	A US view of TDM – demonstrating that fair use should apply to copying for non-expressive purposes and considers the effect of Authors Guild v HathiTrust and Authors Guild v Google.
2019	Vaclav, J., & Malgieri, G., "Commerce in Data and the Dynamically Limited Alienability Rule." *LawArXiv*, 27 Oct. 2019. Web. https://doi.org/10.31228/osf.io/7ztys	Looks at ownership of personal data and the ownership of data in the Internet of Things
2019	Intellectual Property Office, *Music 2025: the Music Data Dilemma* (2019) available at https://www.gov.uk/government/publications/music-2025-the-music-data-dilemma	Interviews with music business representatives on the use of data and databases in the industry
2019	OECD, "Using digital technologies to improve the design and enforcement of public policies", OECD Digital Economy Papers, No. 274, OECD Publishing, Paris, 2019 https://doi.org/10.1787/99b9ba70-en	-
2019	OECD, "Vectors of digital transformation", OECD Digital Economy Papers, 2019, No. 273, OECD Publishing, Paris, https://doi.org/10.1787/5ade2bba-en	-
2019	OECD, *Going Digital: Shaping Policies, Improving Lives*, OECD Publishing, Paris, 2019 https://doi.org/10.1787/9789264312012-en	-
2019	OECD, "Trade and Cross-border data flows", OECD Trade Policy Papers, No. 220, OECD Publishing, Paris, 2019 http://dx.doi.org/10.1787/b2023a47-en	-
2019	OECD, Measuring the Digital Transformation: A Roadmap for the Future, OECD Publishing, Paris, 2019 https://doi.org/10.1787/9789264311992-en	-
2019	OECD, *Data in the Digital Age*, OECD Publishing, Paris, 2019 https://www.oecd.org/going-digital/data-in-the-digital-age.pdf	-
2019	OECD, *Enhanced access to and sharing of data: Reconciling risks and benefits of data re-use across societies*, OECD Publishing, Paris, 2019 available at https://www.oecd.org/publications/enhancing-access-to-and-sharing-of-data-276aaca8-en.htm	-
2020	Drexl, J., "Connected Devices – An Unfair Competition Law Approach to Data Access Rights of Users" (Max Planck Institute for Innovation & Competition Research Paper, No. 20-22), (2020) Available at SSRN: https://ssrn.com/abstract=3746163	Considers the problem of data owners refusing access to data from a competition law perspective
2020	Kop, M., "Machine Learning & EU Data Sharing Practices" available at https://ec.europa.eu/futurium/en/system/files/ged/mauritz_kop_machine_learning_eu_data_sharing_practices_stanford_law_school.pdf	

2020	Horton, A., "Finding the plot for database rights: Ordnance Survey v 77M: database rights in and infringement of a geospatial address dataset (Case Comment)" (2020) *European Intellectual Property Review*, Vol.42(7), pp454-457	Commentary on 77M Ltd v Ordnance Survey Ltd
2020	Ducato, R., "Data protection, scientific research and the role of information" (2020) *Computer Law & Security Review*, Vol. 37, pp105-412	Considers scientific research in light of GDPR – detailed discussion of the relevant provisions of the GDPR for research
2020	Tombal, T., "Economic dependence and data access" (2020) *International Review of Intellectual Property and Competition Law* Vol.51(1), 70-98	Looks at two unreported cases in US and whether there is abuse of market position if access to data is refused to weaker competitors
2020	Aplin., T., & Bently, L., *Global Mandatory Fair Use: The Nature and Scope of the Right to Quote Copyright Works* (Cambridge: Cambridge University Press, 2020)	Makes the case that the quotation exception in Article 10 of the Berne Convention constitutes a global, mandatory, fair use provision replacing the three step test.
2020	Flynn, S., Geiger, C., Quintais, J., Margoni, T., Sag, M., Guibault, L. & Carroll, M.W., "Joint Comment to WIPO on Copyright and Artificial Intelligence" (17 February 2020) available at https://infojustice.org/archives/42009	Commentary on proposed questions from WIPO in request to the impact of artificial intelligence on IP
2020	Flynn, S., Geiger, C., Quintais, J., Margoni, T., Sag, M., Guibault, L. & Carroll, M.W., "Implementing User Rights for Research in the Field of Artificial Intelligence: A Call for International Action" (April 20, 2020) *European Intellectual Property Review* 2020, Issue 7, *American University, WCL Research Paper No. 2020-12*, Available at SSRN: https://ssrn.com/abstract=3578819	Summary of the key points of the submission to WIPO.

Bibliography

I. Legislation

United Kingdom

Copyright Designs and Patents Act 1988, c48

Copyright and Rights in Databases Regulations 1997

Copyright and Related Rights Regulations 2003/2498

Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014/1372

Copyright and Rights in Performances (Quotation and Parody) Regulations 2014 Copyright and Rights in Performances (Disability) Regulations 2014

Copyright and Rights in Performances (Personal Copies for Private Use) Regulations 2014

The Copyright and Related Rights (Marrakesh Treaty etc.) (Amendment) Regulations 2018

Data Protection Act 1998

Data Protection Act 2018

Environmental Information Regulations 2004

Freedom of Information Act 2000

Freedom of Information (Scotland) Act 2002

Human Rights Act 1998

International

Berne Convention for the Protection of Literary and Artistic Works, 1887

European Union

Council Directive 93/83/EEC of 27 September 1993 on the coordination of certain rules concerning copyright and rights related to copyright applicable to satellite broadcasting and cable retransmission Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases

Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases

Directive 2001/84/EC of the European Parliament and of the Council of 27 September 2001 on the resale right for the benefit of the author of an original work of art,

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

Directive 2006/115/EC of the European Parliament and of the Council of 12 December 2006 on rental right and lending right and on certain rights related to copyright in the field of intellectual property

Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights

Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights

Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs

Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on certain permitted uses of orphan works

Directive 2014/26/EU of the European Parliament and of the Council of 26 February 2014 on collective management of copyright and related rights and multi-territorial licensing of rights in musical works for online use in the internal market

Directive (EU) 2017/1564 of the European Parliament and of the Council of 13 September 2017 on certain permitted uses of certain works and other subject matter protected by copyright and related rights for the benefit of persons who are blind, visually impaired or otherwise print-disabled

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (GDPR)

Regulation (EU) 2017/1563 of the European Parliament and of the Council of 13 September 2017 on the cross-border exchange between the Union and third countries of accessible format copies of certain works and other subject matter protected by copyright and related rights for the benefit of persons who are blind, visually impaired or otherwise print-disabled

Regulation (EU) 2017/1128 of the European Parliament and of the Council of 14 June 2017 on cross-border portability of online content services in the internal market

Other

The (US) Copyright Act of 1976, 17 U.S.C. § 101

II. Case Law

A v B (Copyright: Diary Pages) [2007] 7 WLUK 934

Apis-Hristovich EOOD v Lakorda AD (C-545/07) [2009] E.C.R. I-1627; [2009] 3 C.M.L.R. 3

Attorney General v Guardian Newspapers Ltd (No 2) [1988] 3 WLR 776

Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014)

BBC Petitioners [2012] HCJ 10

Bridget Foley Ltd v Ellott [1982] RPC 433

British Horseracing Board Ltd v William Hill Organisation Ltd [2001] EWCA Civ 1268

*British Horseracing Board v William Hill Organisation Ltd (*Case C-203/02) [2004] ECR I-10415 (CJEU)

Campbell v Mirror Group Newspapers Ltd [2002] EWCA Civ 1373

Campbell v Mirror Group Newspapers Ltd [2004] UKHL 22

Christian Institute v Lord Advocate [2016] UKSC 51

Coco v AN Clark Engineers Ltd [1968] FSR 415

Common Services Agency v Scottish Information Commissioner [2008] UKHL 47 DPA

Cramp v Smythson [1944] AC 329

Dawson-Damer v Taylor Wessing LLP [2015] EWHC 2366 (Ch)

Department of Health v Information Commissioner [2011] EWHC 1430

Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg (C-304/07) [2008] E.C.R. I-7565; [2009] 1 C.M.L.R. 7

Dyson v Vax Ltd [2010] EWHC 1923

Exxon Corporation v Exxon Insurance [1982] Ch 119(Ch)

Fixtures Marketing Ltd v Oy Veikkaus AB (C-46/02) [2004] 11 WLUK 205

Fixtures Marketing Ltd v Organismos Prognostikon Agonon Podosfairou (OPAP) (C-444/02) [2004] 11 WLUK 204

*Fixtures Marketing Ltd v OPAP (*Case C-444/02) [2004] ECR I-10549,

Fixtures Marketing Limited v Ab Svenska Spel [2001] E.C.D.R. 29

Fixtures Marketing Ltd v Svenska Spel AB (C-338/02) [2004] 11 WLUK 203

Football Association Premier League Ltd v QC Leisure [2008] EWHC 1411

Football Association Premier League Ltd v QC Leisure [2012] EWCA Civ 1708

Football Association Premier League Ltd v QC Leisure [2012] EWHC 108 (Ch)

Football Association Premier League Ltd v Panini UK Ltd [2002] EWHC 2779(Ch)

Football Association Premier League Ltd v Panini UK Ltd [2003] 4 All ER 1290,

Forensic Telecommunications Services Ltd v Chief Constable of West Yorkshire [2011] WLUK 243

Football Dataco Ltd v Brittens Pools Ltd [2009] EWHC 3294 (Ch),

Football Dataco Ltd v Smoot Enterprises Ltd [2011] EWHC 973(Ch).

*Football Dataco Ltd v Yahoo! UK Ltd and others (*C-604/10) [2012] 3 WLUK 1

Forensic Telecommunications Services Ltd v Chief Constable of West Yorkshire [2011]

Harpers v Barry Henry & Co (1892) 20 R 133

HM Stationary Office v Green Amps Ltd [2007] EWHC 2755

Hollinrake v Truswell [1894] 3 Ch 420

Innoweb v Wegener (C-202/12) [2014] Bus. L.R. 308

IPC Magazines Ltd v MGN Ltd [1998] 2 WLUK211

ITV Broadcasting Ltd v TV Catch Up Ltd [2011] EWHC 2977

ITV Broadcasting Ltd v TV Catchup Ltd [2015] EWCA Civ 204

Kabushiki Kaisha Sony Computer Entertainment Inc v Owen (t/a Neo Technologies) [2002] EWHC 45

Lucasfilm Ltd v Ainsworth [2011] UKSC 39

Media CAT Ltd v Adams [2011] EWPCC 10

Murray v Express Newspapers Plc [2007] EWHC 1908

National Guild of Removers and Storers Ltd v Milner (t/a Intransit Removals and Storage) [2014] EWHC 670 (IPEC)

Navitaire Inc v EasyJet Airline Co Ltd (No.3) [2004] EWHC 1725 (Ch)

Newspaper Licensing Agency Ltd v Marks & Spencer Plc [2001] Ch 257

Newspaper Licensing Agency Ltd v Meltwater Holding BV [2011] EWCA Civ 890

NT1 v Google [2018] EWHC 799

Perfect 10 v. Amazon.com. 508 F.3d 1146 (CA9 2007)

Phonographic Performance Ltd v South Tyneside MBC [2001] 1 WLR 400

Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd [2013] UKSC 18

Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd (C-360/13) [2014] WLR 1025

R (on the application of the Office of Communications) v Information Commissioner [2009] EWCA Civ 90

R V Registered Designs Appeal Tribunal Ex P. Ford Motor Co Ltd [1994] RPC 545

Racing Partnership Ltd v Done Brothers (Cash Betting) Ltd [2019] 3 W.L.R. 779

Ryanair Ltd v PR Aviation BV (C-30/14) EU:C:2015:10; [2015] 2 All E.R. (Comm) 455 (CJEU (2nd Chamber))

Sietech Hearing Ltd v Borland [2003] 2 WLUK 568 *Taylor v Macguire* [2013] EWHC 3804 (IPEC)

Stunt v Associated Newspapers Ltd [2017] EWHC 695

University Court of the University of St Andrews v Student Gowns Ltd [2019] CSOH 86

University of London Press v University Tutorial Press [1916] 2 Ch 601

Universities UK Ltd v Copyright Licensing Agency Ltd [2001] 12 WLUK 37

Vitof Ltd v Altoft [2006] EWHC 1678,

77M Ltd v Ordnance Survey Ltd [2019] EWHC 3007 (Ch)

III. Books and Journal Articles

Aplin., T., & Bently, L., *Global Mandatory Fair Use: The Nature and Scope of the Right to Quote Copyright Works* (Cambridge: Cambridge University Press, 2020)

Austin, W.L. "A Thoughtful and Practical Analysis of Database Protection under Copyright Law, and a Critique of Sui Generis Protection" (1997) *Journal of Technology Law & Policy*, Vol. 3, p35

Bloch, F., "Aspects of the protection of electronic databases and data banks in Europe, in view of the draft Directive. Towards a synthesis of European laws?" (1994*) Illinois Business Law Journal,* Vol. 4, pp457-467

Bonadio, E., & Lucchi, N., "How Far Can Copyright Be Stretched? Framing the Debate on Whether New and Different Forms of Creativity can be Protected" (2019) *Intellectual Property Quarterly* Vol 2, 115

Borghi, M., & Karapapa, S., "Contractual Restrictions on the lawful use of information: sole-source databases protected by the back door?" (2015) *European Intellectual Property Review*, Vol. 37(8), pp505-514

Bygrave, L.A., "Information Concepts in Law: Generic Dreams and Definitional Daylight" (2015), *Oxford Journal of Legal Studies*, Vol.35(1), pp91-120

Bygrave, L. A., "The Data Difficulty in Database Protection" (June 20, 2012). *University of Oslo Faculty of Law Research Paper No. 2012-18*, Available at SSRN: https://ssrn.com/abstract=2088018

Calabresi, G., & Melamed, A.D. "Property Rules, Liability Rules and Inalienability: One view of the Cathedral" (1972) *Harvard Law Review* Vol. 85, p1089

Carroll, M., "Copyright and the Progress of Science: Why Text and Data Mining is Lawful" (2019) *University of California Davis Law Review*, Vol.53, p893

Cerina, P., "The Originality Requirement in the Protection of Databases in Europe and the United States", (1993) *International Review of Intellectual Property and Competition Law* p579

Comande, G., & Schneider, G., "Regulatory challenges of data mining practices: the case of the never-ending lifecycles of "health data"" (2018) *European Journal of Health Law*, Vol.25(3), pp284-307

Cooper, E., and Burrow, S., "Photographic Copyright and the Intellectual Property Enterprise Court in Historical Perspective," (2018), *Legal Studies*, Vol.39(1), pp143-165

Davison, M.J., & Hugenholtz, P. B., "Football fixtures, horseraces and spin-offs: the CJEU domesticates the database right," (2005) *European Intellectual Property Review*, pp113-118

De Franceschi, A., & Lehmann, M., "Data as tradable commodity and new measures for their protection", (2015) *Italian Law Journal*, Vol. 1, pp51-72

Derclaye, E., "What is a database? A critical analysis of the definition of a database in the European Database Directive and suggestions for an international definition", (2002) *Journal of World Intellectual Property* Vol. 5, p981

Derclaye, E., "Do sections 3 and 3A of the CDPA violate the Database Directive? A closer look at the definition of a database in the U.K. and its compatibility with European law" (2002) *European Intellectual Property Review*, Vol.24(10), pp466-474

Derclaye, E., "Databases sui generis right: should we adopt the spin off theory?" (2004) *European Intellectual Property Review*, Vol.26(9), pp402-413

Derclaye, E., "Database sui generis right: what is a substantial investment? A tentative definition" (2005) *International Review of Intellectual Property and Competition Law*, Vol. 36(1), pp2-30

Derclaye, E., "The Court of Justice interprets the database sui generis right for the first time" (2005) *European Law Review* Vol.30, p420

Derclaye, E, *The Legal Protection of Databases: A Comparative Analysis* (Edward Elgar Publishing, 2008)

Derclaye, E., "Recent French decisions on database protection: towards a more consistent and compliant approach with the Court of Justice's case law?"(2012) *European Journal of Law and Technology*, Vol 3(2)

Devici, H.A., "Databases: Is Sui Generis a Stronger Bet than Copyright" (2004) *International Journal of Law & Information Technology*, Vol 12(2), pp178-208

Doldirina, C., et al. "Legal Approaches for Open Access to Research Data." *LawArXiv*, 1 Apr. 2018. Web. https://doi.org/10.31228/osf.io/n7gfa

Drexl, J., "Designing Competitive Markets for Industrial Data: Between Propertisation and Access" (2016) *Max Planck Institute for Innovation and Competition Research Paper No. 16-13* available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2862975

Drexl, J., "Connected Devices – An Unfair Competition Law Approach to Data Access Rights of Users" (2020) *Max Planck Institute for Innovation & Competition Research Paper, No. 20-22*, Available at SSRN: https://ssrn.com/abstract=3746163

Ducato, R., & Strowel, A., "Limitations to text and data mining and consumer empowerment: making the case for a right to "machine legibility"" (2019) *International Review of Intellectual Property and Competition Law*, Vol.50(6), pp649-684

Floridi, L., *Philosophical Conceptions of Information*, in Sommaruga G. (eds) *Formal Theories of Information. Lecture Notes in Computer Science* (Springer, Berlin, Heidelberg, 2009)

Floridi, L., "Is Semantic Information Meaningful Data?" (2005) *Philosophy and Phenomenological Research,* Vol 70(2)

Flynn, S., Geiger, C., Quintais, J., Margoni, T., Sag, M., Guibault, L. & Carroll, M.W., "Implementing User Rights for Research in the Field of Artificial Intelligence: A Call for International Action" (April 20, 2020) *European Intellectual Property Review* 2020, Issue 7, *American University, WCL Research Paper No. 2020-12*, Available at SSRN: https://ssrn.com/abstract=3578819

Gaster, J,. "The EC sui generis right revisited after two years: a review of the practice of database protection in the 15 EU Member States" (2000) *Communications Law, Vol.* 5(3), pp87-98

Geiger, C., Frosio, G., & Bulayenko, O., "Text and data mining in the proposed copyright reform: making the EU ready for an age of big data? Legal analysis and policy recommendations (Legislative Comment)" (2018) *International Review of Intellectual Property and Competition Law*, Vol.49(7), pp814-844

Geiger, C., "The future of copyright in Europe: striking a fair balance between protection and access to information" (2010) *Intellectual Property Quarterly* Vol. 1, pp1-14

Geiger, C., Frosio, G., & Bulayenko, O., "Opinion of the CEIPI on the European Commission's Proposal to Reform Copyright Limitations and Exceptions in the European Union" (2017) *Centre for International Intellectual Property Studies Research Paper No. 2017-09* available at http://infojustice.org/wp-content/uploads/2017/10/Opinion-of-the-CEIPI-on-the-European-Commissions-Proposal-to-Reform-Copyright-Limitations-and-Exceptions-in-the-European-Union.pdf

Ginsburg, J., "No Sweat Copyright and Other Protection of Works of Information and *Feist v Rural Telephone"*, (1992) *Columbia Law Review*, Vol. 92(2), pp338-388

Ginsburg, J., "Creation and Commercial Value: Copyright Protection of Works of Information", (1990), *Columbia Law Review*, Vol. 90 (7), pp1865-1938

Gitelman, L., (ed). *Is Raw Data an Oxymoron*, (Cambridge:MIT Press, 2013)

Gurry, F., *Breach of Confidence* (Oxford: Clarendon Press, 1984)

Han, J., Kamber, M., Pei, J., *Datamining: concept and techniques*, 3rd edn. (Morgan Kaufmann, Waltham, USA, 2011)

Handke, C., Guibault, L., Vallbé, J.J., "Is Europe falling behind in data mining? Copyright's impact on data mining in academic research" in Schmidt, B., Dobreva, M., (eds) *New avenues for electronic publishing in the age of infinite collections and citizen science: scale, openness and trust* (IOS Press, Amsterdam, Netherlands, 2015)

Hansuker, G.M., "The European Database Directive: Regional Stepping Stone to an International Model?" (1997) F*ordham Intellectual Property, Media and Entertainment Law Journal*, Vol.7, p697

Hicks, J.B., "Copyright and Computer Databases: Is Traditional Compilation Law Adequate?" (1987) *Texas Law Review* Vol. 65, p993

Hilty, R.M., & Richter, H., "Text and data mining" in Hilty, R.M., & Moscon, V., (eds) "Modernisation of the EU copyright rules. Position Statement of the Max Planck Institute for Innovation and Competition", *Max Planck Institute for Innovation and Competition Research Paper No. 17-12*, available at https://pure.mpg.de/rest/items/item_2470998_12/component/file_2479390/content

Höppner, T., Kretschmer, M., Xalabarder, R., "CREATe public lectures on the proposed EU right for press publishers", (2017) *European Intellectual Property Review* Vol. 39(10), pp607-622

Hoeren, T. "Big Data and the Ownership in Data: Recent Developments in Europe" (2014) *European Intellectual Property Review* pp751-754

Hoeren, T., & Bitter, P., "Data ownership is dead: long live data ownership" (2018) *European Intellectual Property Review* Vol. 40(6), pp347-348

Hugenholz, P.B., "The New Database Right: Early Case Law from Europe" *Fordham University School of Law Ninth Annual Conference on International IP Law & Policy*

New York, 19-20 April 2001 available at https://www.ivir.nl/publications/intellectual-property/database-right/

Hugenholtz, P.B., "Program Schedules, Event Data and Telephone Subscriber Listings under the Database Directive The 'Spin-Off' Doctrine in the Netherlands and elsewhere in Europe", Paper presented at *Fordham University School of Law Eleventh Annual Conference on International IP Law & Policy* New York, 14-25 April 2003, available at https://www.ivir.nl/publicaties/download/spinofffordham.pdf

Hugenholtz, P.B., "Abuse of Database Right: Sole-source information banks under the EU Database Directive" Paper presented at Conference 'Antitrust, Patent and Copyright', École des Mines/UC Berkeley, Paris, January 15-16, 2004 available at https://www.ivir.nl/publicaties/download/abuseofdatabaseright.pdf

Hugenholtz, P.B., "Data Property: Unwelcome Guest in the House of IP" in J. Reda (ed.), *Better Regulation for Copyright: Academics meet Policy Makers:* Wed 6 Sept 2017 15:00-18:30 University of Southampton, (pp. 65-77). https://juliareda.eu/events/better-regulation-for-copyright/ available at https://www.ivir.nl/publicaties/download/Data_property_Muenster.pdf

Hughes, J., & Weightman, E., "E.C. Database Protection: Fine Tuning the Commission's Proposal", (1992) *European Intellectual Property Review*, Vol.14, p147

Kretschmer, M., Derclaye, E., Favale, M., & Watt, R., "The Relationship between Copyright and Contract Law: A Review commissioned by the UK Strategic Advisory Board for Intellectual Property Policy" (SABIP), 2010, Available at SSRN: https://ssrn.com/abstract=2624945

Koscik, M., & Myska, M., "Database authorship and ownership of sui generis database rights in data-driven research" (2017), *International Review of Law, Computers & Technology,* Vol.31(1), p43

Lipton, J., "Databases as intellectual property: new legal approaches", (2003) *European Intellectual Property Review*, Vol.25(3), pp139-145

Lessig, L., *Code and other laws of cyberspace* (Basic Books, 1999) now *Code 2.0*

Margoni, T., & Dore, G.,"Why we need a text and data mining exception (but it is not enough)", 1 April 2016. 10th edition of the Language Resources and Evaluation Conference (LREC 2016), Portorož (Slovenia), 23-28 May 2016 available at https://zenodo.org/record/248048#.WXdf2oiGNEY

Margoni, T., "The Harmonisation of EU Copyright Law: The Originality Standard" (June 29, 2016) published in Perry, M. (ed), *Global Governance of Intellectual Property in the 21st Century* (Springer International Publishing: Switzerland, 2016) pp. 85-105. Also available at SSRN: https://ssrn.com/abstract=2802327

Maurer, S., Hugenholtz, P., Onsrud, H., "Europe's Database Experiment", (2001) *Science*, Vol.294, pp789-790

Metaxas, G, "Protection of databases: quietly steering in the wrong direction?" (1990) *European Intellectual Property Review*, Vol.12(7), pp227-234

Muller, A., ""Scouting" for unofficial data: SIS pays the price - adding 'breach of confidence' to the list of potential claims by rightsholders" (2019) *Computer and Telecommunications Law Review* Vol.25(8), pp187-191

Myska, M., & Harasta, J., "Less is More? Protecting Databases in the EU After Ryanair", (2016) *Masaryk University Journal of Law and Technology*, Vol.10(2), p170

Powell, M., "The European Union's Database Directive: An International Antidote to the Side Effects of Feist", (1997) *Fordham International Law Journal* Vol.20, p1215

Raue, B., "Free flow of data? The friction between the Commission's European data economy initiative and the proposed Directive on Copyright in the Digital Single Market (Editorial)" (2018) *International Review of Intellectual Property and Competition Law* Vol.49(4), pp379-383

Reichman, J.H., & Samuelson, P., "Intellectual Property Rights in Data" (1997) *Vanderbilt Law Review* Vol. 50, p51

Reichman, J.H., & Uhlir, P.H., "Database Protection at the Crossroads: Recent Developments and Their Impact on Science and Technology" (1999) *Berkeley Technology Law Journal* Vol.14, p793

Rosati, E., "An EU text and data mining exception for the few: would it make sense?" (2018) *Journal of Intellectual Property Law & Practice*, Vol.13(6), pp429-430

Rosler, D. B., "The European Union's Proposed Directive for the Legal Protection of Databases: A New Threat to the Free Flow of Information" (1995) *High Technology Law Journal*, Vol. 10, p105

Sag, M., "The New Legal Landscape for Text Mining and Machine Learning" (2019) *Journal of the Copyright Society of the USA*, Vol.66, p3

Sanks, T., "Database Protection National and International Attempts to Provide Legal Protection for Databases," (1998) *Florida State University Law Review* Vol. 25, p991

Sheils, P.T., & Penchina, R., "What's all the Fuss About Feist – The Sky Is Not Falling on the Intellectual Property Rights of Online Database Proprietors" (1992) *University of Dayton Law Review*, Vol.17, p563

Spence, W.C., & Pedersen, G.L., "Scraping Electronic Data from websites", (2009) *Copyright World*, Vol.193, pp18-21

Stepanov, I., "Introducing a property right over data in the EU: the data producer's right – an evaluation" (2019) *International Review of Law, Computers and Technology*, Vol.34(1), p65

Synodinou, T., "Databases and screen scraping: lawful user's rights and contractual restrictions do not fly together (Case Comment)" (2016) *European Intellectual Property Review,* Vol. 38(5), pp312-315

Taylor, L., "The Ethics of Big Data as a Public Good: Which Public? Whose Good?." *LawArXiv*, 19 June 2018. Web. https://doi.org/10.1098/rsta.2016.0126

Terry, N., "Protecting Patient Privacy in the Age of Big Date" (2012) *UMKC Law Review* Vol. 81, p385

Tombal, T., "Economic dependence and data access" (2020) *International Review of Intellectual Property and Competition Law* Vol.51(1), pp70-98

von Simson, C., "Feist or Famine American Database Copyright as an Economic Model for the European Union" (1994) *Brooklyn Journal of International Law* Vol.20, p729

Waisman, A., "Revisiting Originality" (2009) *European Intellectual Property Review* p370

Wiebe, A., "Protection of industrial data – a new property right for the digital economy?" (2017) *Journal of Intellectual Property Law & Practice*, Vol.12(1), pp62–71

Westkamp, G., "Protecting Databases Under US and European Law – Methodical Approaches to the Protection of Investments Between Unfair Competition and Intellectual Property Concepts", (2003) *International Review of Intellectual Property and Competition Law*, Vol.34, pp772-778

Wolken, J.C., "Just the Facts, Ma'am – A Case for Uniform Federal Regulation of Information Databases in the New Information Age", (1998) *Syracuse Law Review*, Vol. 48, p1263

Zech, H., "A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data", (2016) *Journal of Intellectual Property Law & Practice* Vol. 11, pp460-470

IV. Institutional Publications

European Commission, *Green Paper on copyright and the challenge of technology Copyright Issues requiring immediate action* dated 7 June 1988 Catalogue number CB-CO-88-267-EN-C ISSN 0254-1475

European Commission, *First evaluation of Directive 96/9/EC on the legal protection of databases*, 12 December 2005

European Commission, *Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market*, 2016/0280(COD)

European Commission, *Evaluation of Directive 96/9/EC on the legal protection of databases*, 25th April 2018

European Commission (2018) *Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases*, *Final Report*. A study prepared for the European Commission DG Communications Networks, Content & Technology by JIIP and Technopolis, European Union, 2018

European Copyright Society (2017) *General opinion on the EU copyright reform package*, 24 January 2017 available at https://europeancopyrightsocietydotorg.files.wordpress.com/2015/12/ecs-opinion-on-eu-copyright-reform-def.pdf

Hargreaves, I, *Digital Opportunity: A Review of Intellectual Property and Growth*, Intellectual Property Office, May 2011

Hargreaves, I., Guibault, L., Handke, C., Valcke, P., & Martens, B. (2014). *Standardisation in the area of innovation and technological development, notably in the field of Text and Data Mining: report from the expert group*. (Studies and reports). Publications Office of the European Union. https://doi.org/10.2777/71122

JISC (2012) *The Value and Benefit of Text Mining to UK Further and Higher Education*. Digital Infrastructure. Available at: http://bit.ly/jisc-textm Programme: Digital Infrastructure www.jisc.ac.uk/whatwedo/programmes/di_directions.aspx

JISC (2014) *Data protection and research data* available at https://www.jisc.ac.uk/full-guide/data-protection-and-research-data

JISC (2019) *Data Protection Regulation* available at https://rdmtoolkit.jisc.ac.uk/manage-store-and-preserve/data-protection-regulation/

OECD, *Recommendation of the Council concerning Access to Research Data from Public Funding*, 2006

OECD, *Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information*, 2008

OECD *Recommendation of the Council on Health Data Governance*, 2016

OECD, *The OECD Ministerial Declaration on the Digital Economy: Innovation, Growth and Social Prosperity* (the Cancun Declaration), OECD, Paris, 2016 https://www.oecd.org/internet/Digital-Economy-MinisterialDeclaration-2016.pdf

OECD, *Data in the Digital Age*, OECD Publishing, Paris, 2019 https://www.oecd.org/going-digital/data-in-the-digital-age.pdf

Opinion of Advocate General Stix-Hackl delivered on 8 June 2004 regarding Case C-338/02 *Fixtures Marketing Ltd v Svenska Spel AB*, para 38

Triaille, J., de Meeûs d'Argenteuil, J., & de Francquen, A., "Study on the legal framework of text and data mining", 2014, Publications Office of the European Union DOI:10.2780/1475

V. Other sources

Ballsun-Stanton, B., "Asking about data: Experimental philosophy of Information Technology", (2010) *Computer Science: 5th International Conference on Computer Sciences and Convergence Information Technology* accessed at *https://www.researchgate.net/publication/224218081_Asking_about_data_Experimental_philosophy_of_Information_Technology*

Barbero, M., Cocoru, D., Graux, H., Hillebrand, A., Linz, F., Osimo, D., Siede, A., Wauters, P., "Study on emerging issues of data ownership, interoperability, (re-)usability and access to data, and liability", 25 April 2018. Available at https://ec.europa.eu/digital-single-market/en/news/study-emerging-issues-data-ownership-interoperability-re-usability-and-access-data-and

Borghi, M., "Text and Data Mining", Copyright User - https://www.copyrightuser.org/understand/exceptions/text-data-mining/

Brook, M., Murray-Rust, P. & Oppenheim, C., "The Social, Political and Legal Aspects of Text and Data Mining (TDM)" (2014) D-*Lib Magazine,* 20(11/12), doi:10.1045/november2014-brook accessed at https://openaccess.city.ac.uk/id/eprint/4784/1/D-Lib%20Magazine.pdf

Cocoru, D., & Boehm, M., "An analytical review of text and data mining practices and approaches in Europe", (2016) *Open Forum Europe* available at https://www.creative-destruction.org/publication/ofe-tdm/

Dietrich, N., Guibault, L., Margoni, T., Siewicz, K., Spindler, G., & Wiebe, A., "Safe to Be Open: Study on the Protection of Research Data and Recommendations for Access and Usage" (2013) Göttingen, Germany: Universitätsverlag Göttingen, available at: http://eprints.gla.ac.uk/129335

Hearst, M., "What is Text Mining." Online. Available: http://people.ischool.berkeley.edu/~hearst/text-mining.html last accessed 14 December 2020

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung Byers, A., *Big Data: The Next Frontier for Innovation, Competition and Productivity*, Mckinsey Global Institute, 1 May 2011 - https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation#

Margoni, T., & Kretschmer, M., (2018) "The text and data mining exception in the proposal for a Directive on Copyright in the Digital Single Market: why it is not what EU copyright law needs." Blog post available at https://www.create.ac.uk/blog/2018/04/25/why-tdm-exception-copyright-directive-digital-single-market-not-what-eu-copyright-needs/

Creativecommons.org

https://ukscl.ac.uk/

UK Copyright and Creative Economy Centre
School of Law / University of Glasgow
10 The Square
Glasgow G12 8QQ
www.create.ac.uk

2021/02 DOI: 10.5281/zenodo.4635759

In collaboration with:

University of Glasgow

Urban Big Data Centre