

INFORMATION OWNERSHIP IN THE CLOUD⁺

*Chris Reed**

ABSTRACT

Uncertainty regarding ownership of information stored, created, processed, and distributed in cloud environments presents a potential obstacle to the development of the cloud market. This paper addresses these issues in the context of information flows between cloud providers and their customers. It distinguishes between content that is stored and processed by users, as compared with information generated by cloud providers. Although intellectual property law may determine ownership rights in relevant information, IP rights do not necessarily achieve what parties to cloud computing transactions expect or need. Moreover, in practice, ownership may be less important than control over, and access to, information. The paper concludes with an examination of the emerging concept of ‘data trusts’ as a potential approach to managing complex interests in data.

[February 2022]

⁺ This paper is an author-edited version of Chapter 5, “Information Ownership in the Cloud”, in C. Millard (ed.) *Cloud Computing Law*, (2nd edn, OUP 2021). The book is available from Oxford University Press at: <https://global.oup.com/academic/product/cloud-computing-law-9780198716679?lang=en&cc=uk>. The author is grateful to Microsoft for the generous financial support that has made this research possible. Responsibility for views expressed, however, remains with the author.

^{*} Professor of Electronic Commerce Law at the Centre for Commercial Law Studies, Queen Mary University of London.

TABLE OF CONTENTS

1. What is ownership?	2
1.1 Introduction	2
1.2 Owning digital information.....	3
1.3 Psychological and emotional claims to ownership	5
1.4 Ownership as rights to control.....	8
2. Ownership rights in information.....	9
2.1 Introduction	9
2.2 Copyright.....	10
2.3 Database right	14
2.4 Confidential information	17
3. Information provided or generated by the customer.....	19
3.1 Introduction	19
3.2 Information placed in the cloud	20
3.3 Information generated in the cloud	23
3.4 Confidentiality and contracts.....	27
4. Information provided by the cloud provider	28
5. Information generated by the cloud provider which relates to the customer's information.....	31
5.1 Introduction	31
5.2 Information generated for the purpose of service provision.....	32
5.3 Disclosing information about the customer.....	32
5.4 Metadata	33
5.5 Information generated from aggregated data analysis	35
6. Data sharing and data trusts	37
7. Concluding remarks	41

1. WHAT IS OWNERSHIP?

1.1 INTRODUCTION

A major obstacle to understanding ownership is that it is both a social and a legal construct. Children are not born with an inherent concept of ownership but gradually learn it from others around them, a constant

source of playground disputes.¹ In everyday use of the term ‘ownership’ the social and legal concepts are not distinguished, and so a person's concept of what is ‘mine’ may not match exactly the legal rights which that person possesses. We therefore need to separate out the legal concept of ownership in order to understand how it applies to information in the cloud.

In law, ownership consists of a collection of legal rights over property. These are primarily the right to exclusive possession of the property, which allows the owner to exclude others from using it or to reclaim possession of it, and the right to dispose of the property, which includes both the transfer of ownership and more limited disposals such as the right to use the property, to take it as a security, etc.

It is helpful to take a concrete example here. Let us suppose that you, the reader, have purchased a physical copy of the *Cloud Computing Law* book (OUP, 2021).² Because you own it you have legal rights to reclaim it if someone takes it away from you. You can sell it to another reader, or lend or hire it to them. If you owe money to someone you can let them have the book temporarily on terms that they can sell it if you don't pay your debt. All these are quite easy to understand, and closely match the social construct of property as well.

However, things become much more complicated when we consider the information contained in the book. What rights does your ownership of the physical book give you in respect of that information? Here, the existence of your property rights in the physical book confuses things. You clearly have the *power* to read the information in the book because no other person has a legal right to take the book away from you. You also have the power to make mental use of that information because the law has no power to control your mental processes. But, as all readers are probably aware, your ownership of the book does not give you unrestricted legal rights to copy the information or to make that copy available to others.

1.2 OWNING DIGITAL INFORMATION

The law has struggled with ownership of digital information (though not necessarily with digital *assets*³) precisely because it is no longer recorded in

¹ Susan Isaacs, *Social development in young children* (Routledge & Kegan Paul 1993).

² If you purchased this as an e-book, the position is far more complex and interesting. See further, Johan David Michels and Christopher Millard, ‘Mind the Gap: The Status of Digital Files Under Property Law’ (2019) Queen Mary School of Law Legal Studies Research Paper No. 317/2019 <<https://ssrn.com/abstract=3387400>> accessed 7 February 2022.

³ In *Swift v Dairywise Farms Ltd* [2001] EWCA Civ 145 the UK courts held that a milk quota could be trust property, and in *Armstrong v Winnington* [2012] EWHC 10 that a carbon emissions allowance would be the subject of a

permanent form on a physical object which can be owned. Thus even if the digital information is recorded on physical property which you own, for example your laptop's hard disk, the fact that it is technologically necessary to make copies of the information in order to use it produces potential conflicts between your claims to unrestricted use and the claims of the person who produced or supplied you with the information.

These problems arise because information is non-rivalrous. Physical property is rivalrous – only one person can possess your laptop at any one time – and so for physical property it makes sense to define rights in terms of possession. But information can be possessed by many persons simultaneously, without possession by one of them in any way diminishing the ability of any other possessor to make use of the information. This is legally unsatisfactory because information can be valuable. If a book contained only blank pages you would not pay nearly as much for it.

The law attempts to solve this problem by creating legal rights in some types of information, in the form of intellectual property rights. However, these rights cannot be exclusive in the same sense as the ownership rights in physical property. Society recognises that useful information is generated by using others' information – this paper could not have been written if the author had no access to the texts of laws and judicial decisions and the writings of other scholars. As a consequence, intellectual property rights are highly limited in scope in order to prevent the monopolisation of information. We will examine them further in section 0.

proprietary tracing claim. Most recently in *AA v Persons Unknown* [2019] EWHC 3556 Bryan J held that Bitcoins, procured through criminal activity and held in an account by an intermediary, could be the subject of an interim proprietary injunction requiring the identities of the holders of those Bitcoins to be revealed. The judge adopted the reasoning of the UK Jurisdiction Task Force in its 'Legal statement on crypto assets and smart contracts' (*Lawtech Delivery Panel*, November 2019) para 59 and held that a crypto asset such as Bitcoin was property, at least for the purposes of granting the injunction, because it was: 'definable, identifiable by third parties, capable in [its] nature of assumption by third parties, and having some degree of permanence.'
<https://35z8e83m1ih83drye280o9d1-wpengine.netdna-ssl.com/wp-content/uploads/2019/11/6.6056_JO_Cryptocurrencies_Statement_FINAL_WEB_111119-1.pdf> accessed 7 February 2022. See further Johan David Michels and Christopher Millard, "Digital Assets in Clouds", in: C. Millard (ed.) *Cloud Computing Law*, (2nd edn, OUP 2021). In each of these cases the digital asset was functionally unique – its digital information could be copied, but could not be used to dispose of the asset. These digital assets are very different from the kinds of digital information discussed in this paper, for all of which copies are as useful as the 'originals' and which therefore cannot be the subject of this kind of property-based remedy – see *Your Response Ltd v Datateam Business Media Ltd* [2014] EWCA Civ 281.

1.3 PSYCHOLOGICAL AND EMOTIONAL CLAIMS TO OWNERSHIP

Disputes over information ownership, and negotiations involving transfer of digital information, can be surprisingly difficult to resolve. This is because property is not merely a factual description, but also a psychological concept. Ownership is a feeling as well as a fact. Our physical possessions often manifest our personalities and can feel like extensions of ourselves,⁴ so that wanton destruction of those possessions feels also like a psychological assault.⁵ This can equally be true of information, and particularly so for information about ourselves or which we have created or generated. Businesses can feel equally strongly about the information which they believe they 'own'.

The endowment effect, which has been studied extensively by economists,⁶ leads individuals to value more highly something which they already possess than some equivalent for which they might exchange it. Interestingly, the endowment effect becomes stronger if the item in question is owned, and not merely possessed. Psychological research reveals that individuals who believe they own an item exhibit the effect even more strongly, resulting in a value being placed on that item which cannot be

⁴ Thus Hegel argues that our intellectual constructions become part of our persona, Georg Wilhelm Friedrich Hegel, *Elements of the Philosophy of Right* (1821) and Spiekermann et al state that:

...possessions also serve as a symbolic expression of the self, and...there is a close connection between possessions, self-identity, and individuality.'

Sarah Spiekermann, Jana Korunovska, and Christine Bauer, 'Psychology of Ownership and Asset Defense: Why People Value their Personal Information Beyond Privacy' (2012) Thirty Third International Conference on Information Systems, 5 <https://epub.wu.ac.at/3630/1/2012_ICIS_Facebook.pdf> accessed 7 February 2022 (hereafter Spiekerman and others, 'Psychology of Ownership')

See also RP Abelson and DA Prentice, 'Beliefs as Possessions: A Functional Perspective' in Anthony R Pratkanis, Steven J Breckler, and Anthony G Greenwald (eds), *Attitude Structure and Function* (Erlbaum 1989) 361; Helga Dittmar, *The Social Psychology of Material Possessions: To Have Is to Be* (St Martin's Press 1992); J Douglas Porteous, 'Home: The Territorial Core' (1976) 66 *Geographic Review* 383.

⁵ 'Possessions may even afford a sense of security. If they are lost or taken away, individuals may experience an erosion of the sense of self.' Jon L Pierce and others, 'The State of Psychological Ownership: Integrating and Extending a Century of Research' (2003) 7 *Review of General Psychology* 84, 90 (hereafter Pierce and others, 'The State of Psychological Ownership').

⁶ Richard Thaler, 'Toward a positive theory of consumer choice' (1980) 1 *Journal of Economic Behavior and Organization* 39; Daniel Kahneman, Jack L Knetsch, and Richard H Thaler, 'Experimental Tests of the Endowment Effect and the Coase Theorem' (1990) 98 *Journal of Political Economy* 1325.

explained by economic theories.⁷ This valuation is increased further if the individual considers that he or she has earned, or deserves, the owned thing,⁸ so that creators of intellectual property value their creations even higher than intellectual property which was purchased from another creator.⁹ This sense of ownership applies to information as well as to physical property.¹⁰

This psychological aspect of property produces strong emotional feelings in information 'owners'. If they perceive the information in question as 'mine', then this influences how far they are prepared to allow others, such as cloud providers, to make use of it. A likely reason for this strong desire for control is that legal ownership of physical property confers control of use by others, and this might explain the strength of our emotional need to control the information which we think we own. Psychologists certainly find that a key psychological element of ownership is the right, or at least the perceived right, to control the thing which is owned.¹¹

Legal scholars have suggested that this explains why property owners are highly resistant to damages awards when their property is interfered with, and will go to great lengths to secure injunctions and mandatory orders as a remedy. Being forced to accept damages is effectively

⁷ James K Beggan, 'On the Social Nature of Nonsocial Perception: The Mere Ownership Effect' (1992) 62 *Journal of Personality and Social Psychology* 229, suggesting that this is explained by the role which property ownership plays in one's self-image.

⁸ George Loewenstein and Samuel Issacharoff, 'Source Dependence in the Valuation of Objects' (1994) 7 *Journal of Behavioral Decision Making* 157, 165.

⁹ Christopher Buccafusco and Christopher Sprigman, 'Valuing Intellectual Property: An Experiment' (2010) 96 *Cornell LR* 1.

¹⁰ Spiekermann and others conducted experiments on personal information stored in Facebook accounts which found that the more strongly an individual believed they owned that information, the more they would be willing to pay to preserve it from deletion. They also found that knowledge that a third party was interested in acquiring the information led individuals to value it even more highly. Interestingly, and perhaps counter to what lawyers have thought in this field, a belief in information ownership was demonstrated experimentally to be far more influential in valuing personal information highly than whether the information was considered private or not. Spiekermann and others, 'Psychology of Ownership' (n 4) 6, 9.

¹¹ FW Rudmin and JW Berry, 'Semantics of ownership: A free-recall study of property' (1987) *The Psychological Record*, 37. See also L Furby, 'Possession in Humans: An Exploratory Study of Its Meaning and Motivation' (1978) 6 *Social Behavior and Personality* 49, 60:

Possessions have an instrumental function – they make possible certain activities and pleasures. In other words, they enable one to effect desired outcomes in one's environment ... a central feature of possession is the ability to affect and control the object in whatever way one wishes.

a compulsory disposal of the property or part of it, and denies owners the control which they think ownership ought to confer.¹²

In the realm of physical property, the legal constituents of ownership are a very close match to its psychological conception. Differences arise only in special cases, such as credit or leasing purchases of cars, but even here there is likely to be a psychological recognition that the purchaser is not fully the owner. By contrast, as we shall see in the following sections, the legal ownership status of much information is very far from a match to its psychological or emotional status, and the legal rights which information ownership provides will not confer the level of control which information owners need, psychologically or emotionally. This is a regular cause of disagreement between those engaged in information transactions, and it may be helpful to recognise that it has a psychological cause.¹³

This should tell us that a purely legal understanding of information ownership will be unable to explain the demands which information owners are likely to place on cloud providers, and which they will expect to be reflected in the cloud Terms of Service (‘ToS’). In addition to the legal/psychological mismatch, it is both commercially and technologically necessary for cloud providers to use customer information in ways which would infringe their ownership rights, and thus to take permission to do so in the cloud ToS.

By now, then, it should be clear that legal ownership rights are not the most important factor in defining cloud relationships. The primary mechanism for apportioning control between cloud providers and users will always have to be the cloud ToS, rather than ownership rights in the information. This is comparatively uncontroversial in commercial cloud relationships, partly because commercial cloud users are used to agreeing the limits of their control over ‘their’ information and also because those

¹² Jeffrey J Rachlinski and Forest Jourden, ‘Remedies and the Psychology of Ownership’ (1998) 51 Vand L Rev 1541.

¹³ See Pierce and others, ‘The State of Psychological Ownership’ (n 5) 96:

...legal ownership may facilitate and speed-up the emergence of psychological ownership because it allows the individual to explore the three routes leading to this state. Legal ownership provides the right to control or change the target, more or less at one’s own will, the right to explore and to come to intimately know the target, and the right to invest the self in the target. Lack of legal ownership may in some cases provide a more precarious form of ownership, in that an individual has to avoid violation of the law (eg, physical barriers, customs and social practices) to exercise one or more of the three routes to psychological ownership. In the absence of legal ownership, one may also have to contend with a greater fear of separation, claim of ownership by the legal owner, and loss of the object.

who manage commercial operations tend to be less psychologically invested (though not completely so) in ownership.

However, ToS remain highly problematic in relation to consumer cloud services. Although consumers increasingly understand that their ‘free’ services such as social media are in fact paid for through use of their information, the psychological sense of ownership is highly influential in defining those consumers’ attitudes to cloud ToS which they may feel take too many liberties with ‘their’ property.

1.4 OWNERSHIP AS RIGHTS TO CONTROL

The analysis above indicates that when considering information in digital form, the most important aspect of ownership (in its widest sense) is the right to control how others use that information.

Legal rights are not the only way of achieving control. One self-help option is by using technology to change the *non-rivalrous* nature of that particular information by making it *excludable*, thus ensuring that it cannot be used, or at least that its content cannot be used, by others. A cloud user might do this via technical restrictions, such as encrypting the information so that only he or she can use it. But encryption has drawbacks as well as advantages, in particular that some aspects of the cloud service might be unusable in respect of the encrypted information. Another option for sophisticated customers is to use a Virtual Machine (‘VM’) running on the cloud provider’s infrastructure, which will only allow the customer to access and use the information.

These self-help options are not appropriate for many cloud users, and so they need alternative control mechanisms. These might be conferred by the cloud service contract, and will be supplemented by any ownership rights in the information. Intellectual property ownership confers some rights of control, but so too does the law of confidence, and so we will include that in our discussion.

Arguably, the law pertaining to privacy and data protection also confers rights of control, but these rights are outside of the scope of this paper.¹⁴ That said, it is worth mentioning one recent development here, because in *Lloyd v Google LLC*¹⁵ the UK Court of Appeal held that data protection law confers upon data subjects some rights of control over their personal data. In that case Google had circumvented the cookie controls in

¹⁴ To learn more about privacy and data protection in a cloud-related context, see Dimitra Kamarinou, Christopher Millard and Felicity Turton, “Protection of Personal Data in Clouds and Rights of Individuals” and “Responsibilities of Controllers and Processors of Personal Data in Clouds” and Ulrich Wuermeling and Isabella Oldani, “Regulation of International Data Transfers in Clouds”, all in: C. Millard (ed.) *Cloud Computing Law*, (2nd edn, OUP 2021).

¹⁵ *Lloyd v Google LLC* [2019] EWCA Civ 1599 (hereafter ‘*Lloyd v Google*’).

the Safari browser, allowing it to collect browser-generated information ('BGI') even though the user's browser setting ostensibly prevented Google from doing so. The court held:

Even if data is not technically regarded as property in English law, its protection under EU [data protection] law is clear. It is also clear that a person's BGI has economic value: for example, it can be sold. It is commonplace for EU citizens to obtain free wi-fi at an airport in exchange for providing their personal data. If they decline to do so, they have to pay for their wi-fi usage. The underlying reality of this case is that Google was able to sell BGI collected from numerous individuals to advertisers who wished to target them with their advertising. That confirms that such data, and consent to its use, has an economic value.

Accordingly, in my judgment, a person's control over data or over their BGI does have a value, so that the loss of that control must also have a value.¹⁶

On appeal, the UK Supreme Court¹⁷ agreed that data protection law did confer rights on control over personal data, and also that the ability to make use of that personal data had economic value. However, it also held that data protection law did not give a remedy in damages for such misuse because the wording of section 13 of the UK Data Protection Act 1988 limited damages claims to those for distress or material damage.

Article 82(1) of the GDPR now provides that 'Any person who has suffered material or non-material damage as a result of an infringement of this Regulation shall have the right to receive compensation from the controller or processor for the damage suffered.' Non-material damage is wider than just distress, and would surely cover the claim made in *Lloyd v Google LLC*. It will be interesting to see whether in future the courts develop this line of reasoning into some wider, quasi-property conception of rights in data. Currently, though, it is only personal data over which rights of control exist otherwise than via contractual or property rights. Our next step will therefore be to examine information ownership rights in more detail, to identify the levels of control they provide and their limits.

2. OWNERSHIP RIGHTS IN INFORMATION

2.1 INTRODUCTION

Intellectual property rights are extremely varied. Patents protect inventions, design rights protect both decorative and functional designs,

¹⁶ *Lloyd v Google* (n 15) paras 46-47.

¹⁷ *Lloyd (Respondent) v Google LLC (Appellant)* [2021] UKSC 50].

trade marks protect signs which indicate the origin of products, performance rights protect the way in which works are performed, and so on. Any or all of these rights might be applicable to particular information in the cloud, but their primary focus is not the information itself but some other aspect of the protected thing. Thus, for example, a patent confers the right to prevent others from making or using the invention – although the patent is described using information, that information is not itself protected by the right.

This section will therefore concentrate on the two intellectual property rights which *do* focus, albeit indirectly, on information, and on the law of confidence which is directly about information. There is only space for the briefest description of the nature of these rights and how they can be infringed, sufficient we hope to enable the reader to understand the issues which arise in the cloud.

2.2 COPYRIGHT

National copyright laws across the world recognise that copyright subsists in *works* in digital form. A work is more than a collection of information – to achieve protection it must be a work of authorship. Copyright laws tend to set out lists of types of work, such as literary and artistic works, films, computer programs, etc. These lists can be exhaustive, or can extend protection to any work similar to the list. But in practice, if a court finds a sufficient level of authorship it seems always able to fit the work into an appropriate copyright category.

However, these laws differ in their conceptions of what degree of authorship is necessary to constitute a work. English copyright law, and the law of those countries whose law originated from England, started from the position that copyright should protect all works¹⁸ where sufficient labour, skill or judgment had been used in their creation.¹⁹ There was no requirement for creativity per se, but information was not to be protected if its creation required minimal effort²⁰ or if what was created was too

¹⁸ With the exception of databases in England – Copyright, Designs and Patents Act 1988, section 3(1)(a) as amended by Copyright and Rights in Databases Regulations 1997, SI 1997/3032, regulation 5 (hereafter ‘Copyright, Designs and Patents Act 1988’).

¹⁹ *Ladbroke (Football) Ltd v William Hill (Football) Ltd* [1964] 1 WLR 273. A similar view is taken in many other common law jurisdictions – *Desktop Marketing Systems Pty Ltd v Telstra Corporation Limited* [2002] FCAFC 112 (Federal Court of Australia).

²⁰ *GA Cramp & Sons Ltd v Frank Smythson Ltd* [1944] AC 329.

minimal to be recognised as a work.²¹ So uncreative works such as a spreadsheet or a table would qualify for copyright protection.

However, EU jurisprudence from the CJEU has influenced this approach, and in *SAS Institute Inc v World Programming Ltd*²² the Court of Appeal held that it should follow the CJEU decision in *Infopaq International A/S v Danske Dagblades Forening*²³ by requiring that a work should be the authors ‘intellectual creation’ in order to qualify for protection:

The essence of the term is that the person in question has exercised expressive and creative choices in producing the work. The more restricted the choices, the less likely it is that the product will be the intellectual creation (or the expression of the intellectual creation) of the person who produced it.²⁴

What this means is that although there is no requirement for a work to show artistic or literary merit, the mere effort required to produce it is not enough. Thus a work which can only be expressed in a small number of ways, for example because of technical constraints, may not qualify as an intellectual creation.

Other common law countries might still protect such information by copyright on the basis of the labour and skill required to create it, although the US Supreme Court has rejected this ‘sweat of the brow’ test, and now requires a minimal level of creativity for copyright protection, expressed as ‘intellectual creation’.²⁵ That said, the level of creativity required in the US is still low, so that if the creator is able to exercise any choice at all in how it should be expressed, copyright can be held to subsist. Thus, for example, the choice of headings for a ‘Yellow Pages’ directory has been held to be a copyright work.²⁶

Civil law countries, which protect authors’ rights rather than copyright, always tended to demand a minimum level of creativity to qualify

²¹ *Exxon Corporation v Exxon Insurance Consultants International Ltd* [1982] ch 119 (no copyright on the word ‘Exxon’), *Hitachi Ltd v Zafar Auto and Filter House* [1997] FSR 50, 58 (Copyright Board, Karachi, Pakistan; no copyright in the word ‘Hitachi’).

²² *SAS Institute Inc v World Programming Ltd* [2013] EWCA Civ 1482.

²³ *Infopaq International A/S v Danske Dagblades Forening* [2009] ECR I-6569 (hereafter ‘*Infopaq*’).

²⁴ *Infopaq* (n 23) para 32.

²⁵ *Feist Publications Inc v Rural Telephone Service Company Inc* 499 US 340 (1990) (hereafter ‘*Feist*’).

²⁶ *BellSouth Advertising and Publishing Corp v Donnelly Information Publishing Inc* 933 F 2d 952 (11th Cir 1991) (holding that copyright subsists in Yellow Pages telephone directories due to the minimal creativity in devising the business categories under which listings are set out).

for protection. Germany is perhaps the clearest example, requiring there to be a creative step (*Gestaltungshöhe*) to distinguish a work from mere information.²⁷ Prior to the EU Software Directive²⁸ the German courts had held that some types of software (such as operating systems) were functional, rather than creative, and were thus not protected by author's right.²⁹ This is still likely to be the position for purely functional information such as data tables.

Taking a holistic view of copyright laws around the world, it seems that a substantial proportion of the information which business customers upload to or generate by using the cloud service is unlikely to qualify for copyright protection. Documents and emails will of course be copyright works, but the requirement for intellectual creation means that works which consist primarily of facts cannot, without more, satisfy the required level of creativity. Where creativity is possible lies in the selection and arrangement of those facts, and in the EU this is stated expressly in the Databases Directive, Article 3.³⁰ If the selection and arrangement show sufficient intellectual creativity (and the level required will vary between countries) then that aspect of the work will attract copyright protection. Similarly, if the facts can be expressed in narrative form in several ways, the choice of how to express them could be sufficiently an intellectual creation to merit protection. However, this will confer no copyright protection on the facts themselves – if they can be copied without also taking the protected selection and arrangement or expression, no infringement will have occurred.

One uncertainty under UK law is the status of computer-generated works, which are specifically granted copyright protection by the Copyright, Designs and Patents Act 1988, section 9. The person who made the arrangements necessary for the work to be created is deemed to be the author. This is necessary because such works are defined in section 178 as works 'generated by computer in circumstances such that there is no human author of the work'. Some works uploaded to or generated in the cloud will be computer-generated, such as periodic database reports produced by running a script. By definition, the author will have undertaken no

²⁷ *Brombeer Muster*, BGH decision of 27 January 1983, 1983 GRUR 377.

²⁸ Directive 91/250/EEC on the legal protection of computer programs [1991] OJ L 122/42, Article 1(3) which requires a computer program to be the author's 'own intellectual creation' to qualify for protection by copyright.

²⁹ *Inkasso-Programm*, BGH decision of 9 May 1985, 1986 IIC 681; *Betriebssystem*, BGH decision of 4 October 1990, 1991 IIC 723.

³⁰ Directive 96/9 of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L77/20 (hereafter 'Databases Directive 1996').

intellectual creation to produce the work,³¹ and so the current copyright status of such works remains in doubt in the light of the later cases discussed above.

Control of use is what copyright is about, and copyright owners have a wide range of rights which vary according to the nature of the work. The main rights which are likely to be relevant to a cloud relationship are the right to control copying and the right to control communication to the public.

All processing of digital information takes place by making and operating on copies. Initially, copyright law treated this as meaning that even temporary copies required permission of the copyright owner. However, in recent years EU law has developed on the basis of the provisions of the Information Society Directive,³² Article 5(1), which provides:

1. Temporary acts of reproduction referred to in Article 2, which are transient or incidental, which are an integral and essential part of a technological process and whose sole purpose is to enable:
 - a) a transmission in a network between third parties by an intermediary, or
 - b) lawful use

of a work or other subject-matter to be made, and which have no independent economic significance, shall be exempted from the reproduction right provided for in Article 2.

In the *Meltwater* case³³ the CJEU interpreted Article 5 as meaning that temporary and incidental copies which were made in the course of a normal and intended use of the work would not be infringements. Thus the

³¹ Though, if the computer was used as a tool to assist the human author's creative process then the work will of course be an intellectual creation and copyright will subsist – *Express Newspapers plc v Liverpool Daily Post & Echo* [1985] FSR 306. But such works fall outside the s 178 definition because there *is* a human author, and so are not computer-generated.

³² 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 [2001] OJ L167/10.

³³ *Public Relations Consultants Association Ltd v Newspaper Licensing Agency Ltd and Others* [2014] C-360/1, ECLI:EU:C:2014:1195.

temporary copies made when viewing a website did not require permission from the copyright owner.³⁴

Similar limitations apply also to the right to control communication to the public. In *Svensson and Others*³⁵ the court decided that although providing clickable links to another's website was an act of communication to the public, it did not infringe because it was a communication, by the same means, to the same public as communication via the original website. By contrast, in *Renckhoff*³⁶ it was decided that copying information (in this case a photograph) from a website and then making it available from its new location was an infringing act of communication because the photograph was not longer viewed by users in its original context, and it was thus a communication to a different public.

Cloud providers will inevitably be copying information which is 'owned' by their customers in the course of providing the service, and in the case of some cloud services (such as social media) will be communicating it to the public. It would seem that these might not even amount to infringements because they are all incidental to the customer's intended use of the work, so that there will be an implied licence to make the necessary copies. In most cases the cloud ToS will grant the cloud provider an express licence to do these things.

2.3 DATABASE RIGHT

If the law applicable to creation is that of an EU Member State then information in the form of a database receives *sui generis* protection under the Database Directive. A database is 'a collection of independent works, data or other materials arranged in a systematic or methodical way and capable of being individually accessed by electronic or other means' (Article 1(2)).

The basic scheme of the Directive is as follows:

1. Databases are protected by the *sui generis* right if 'there has been qualitatively and/or quantitatively a substantial

³⁴ In that case, Meltwater made a press monitoring service available to its customers under a licence from the body representing the press companies. Meltwater's customers reviewed these reports, which consisted of licensed extracts from press websites, via the Meltwater service rather than directly from the original website. The argument, which was rejected by the court, was that by viewing these reports Meltwater's customers themselves made unauthorised copies of that material and therefore also required a licence from the representative body.

³⁵ *Nils Svensson and Others v Retriever Sverige AB* [2014] C-466/12, ECLI:EU:C:2014:76.

³⁶ *Land Nordrhein-Westfalen v Dirk Renckhoff* [2018] C-161/17, ECLI:EU:C:2018:634.

investment in either the obtaining, verification or presentation of the contents' (Article 7(1)). Surprisingly, this means that a large number of databases receive no protection because almost all the investment went into the *creation* of the data, and not its obtaining, verification or presentation. In *British Horseracing Board Ltd and Others v William Hill Organization Ltd*³⁷ the British Horseracing Board (BHRB) maintained a comprehensive database relating to British horse racing, and William Hill took information from third party news feeds from the BHRB database and used that information in its online betting service. On reference to the ECJ, the court held that all investment in creating the database, i.e. receiving and verifying data about race entries, should be disregarded in deciding whether the database met the definition.³⁸ When the case returned to the UK Court of Appeal³⁹, it was found on the evidence that almost all the investment related to creation, and that so little had been invested in obtaining, verification or presentation that the BHRB database did not qualify for protection under database right. More recently though, in *Football Dataco Ltd v Stan James Ltd (No2)*⁴⁰ the UK courts decided that database right *did* subsist in a database of information about live football matches, even though the main element of investment was the cost of human observers who recorded events occurring during the matches, holding that the investment was made for the purpose of recording the data for use in the database. The important distinction (although this is not clearly stated in either decision) appears not to be between creation and recording, but rather the *purpose* for which the investment was made. In the *BHRB* case the purpose of the investment was to enable horse races to be organized, whilst in *Football Dataco (No2)* the purpose was solely to make the database. Given that the database right is intended to protect investment in making databases, asking the purpose of the investment seems an appropriate way of drawing a dividing line.

³⁷ *British Horseracing Board Ltd and Others v William Hill Organization Ltd* [2001] EWHC 517 (Pat) (High Court); [2001] EWCA Civ 1268 (CA); [2004] Case C-203/02, ECLI:EU:C:2004:695 (hereafter '*British Horseracing Board* C-203/02'); [2005] EWCA Civ 863 (CA) (hereafter '*British Horseracing Board* [2005] EWCA Civ 863').

³⁸ *British Horseracing Board* C-203/02 (n 37) paras 30-34.

³⁹ *British Horseracing Board* [2005] EWCA Civ 863 (n 37).

⁴⁰ [2013] EWCA Civ 27.

2. The maker of such a database has the right 'to prevent extraction and/or re-utilization of the whole or of a substantial part evaluated qualitatively and/or quantitatively, of the contents of that database' (Article 7(1)) subject to the lawful user's rights (Article 8(1)).
3. Protection under the *sui generis* right lasts for 10 years from first making the database available to the public or 15 years from its creation, whichever is the shorter (Article 10).
4. Copyright protection is restricted to those databases which 'by reason of the selection or arrangement of their contents, constitute the author's own intellectual creation',⁴¹ but this protection does not extend to the database contents because they are not sufficiently creative (Article 3).

It is important to note that database right does not enable the right-holder to control copying of data. Rather, it protects the investment which the maker of the database has made, through preventing the extraction and reutilisation of such a large amount of the database that its economic value is impaired. This assessment can be made on a quantitative or qualitative basis, or both, but in either case must be based on how much of the investment was thereby appropriated. This point was also established by the ECJ in *British Horseracing Board Ltd and Others v William Hill Organization Ltd*, where BHRB argued that weight should also be given to the importance of the data in question; in other words, even if all data elements required the same investment, some would be of more interest to users of the database and thus be qualitatively more substantial than other elements. This argument was rejected. Accepting it would, in effect, grant a right in individual content elements rather than the database as a whole, and this was not the purpose behind the Directive.⁴² The correct test was whether the material taken by William Hill represented a substantial part of BHB's investment in the database as a whole, and on the facts what was taken was an insubstantial part of that investment.⁴³

⁴¹ See further *Football Dataco Ltd v Yahoo! UK Ltd* [2013] FSR 1 paras 37 - 39 (references omitted):

...as is apparent from recital 16 of Directive 96/9, the notion of the author's own intellectual creation refers to the criterion of originality. As regards the setting up of a database, that criterion of originality is satisfied when, through the selection or arrangement of the data which it contains, its author expresses his creative ability in an original manner by making free and creative choices. By contrast, that criterion is not satisfied when the setting up of the database is dictated by technical considerations, rules or constraints which leave no room for creative freedom.

⁴² *British Horseracing Board* C-203/02 (n 37) para 72.

⁴³ *British Horseracing Board* C-203/02 (n 37) paras 78–80.

2.4 CONFIDENTIAL INFORMATION

The third tool for controlling use of information is the law relating to the protection of confidential information, including trade secrets. Much of the information which cloud users upload to or generate by using the cloud service is likely to be information which the cloud user wishes to keep confidential, whether or not it is also protected by copyright or as a database. There is, of course, an important distinction between information which someone wishes to remain confidential, in the sense that it should not be disclosed to others, and information which the law treats as sufficiently confidential to provide a remedy in the case of disclosure.

An international consensus on a minimum level of protection is set out in Article 39(2) of the World Trade Organisation Agreement on Trade-Related Aspects of Intellectual Property Rights 1994, which provides that protection must be given to information which:

1. is secret in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question;
2. has commercial value because it is secret; and
3. has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret.⁴⁴

Trade secrets are thus, in simple terms, information which another business might use to gain a commercial advantage.

Businesses will also possess information which is not a trade secret but is still confidential – for example, the reason why an employee was dismissed – and non-businesses such as individuals or charities will also possess confidential information. This confidential information usually receives legal protection as well. English law protects all information which is confidential in nature and has been received subject to an obligation of confidence, express or implied.⁴⁵

Those who ‘own’ confidential information will want to protect it from unauthorised disclosure, and from use by others even if they do not disclose it. But where that information is held in the cloud, the cloud provider also has some degree of control over the information. The concern is therefore to

⁴⁴ This wording is adopted without modification in EU Directive 2016/943 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L157/1 15.

⁴⁵ *Coco v AN Clark (Engineers) Ltd* [1969] RPC 41.

ensure that the cloud provider preserves its confidentiality, both against accidental disclosure and against misuse by, for example, an employee. The cloud provider may also make confidential information available to the customer, and will want similar protection. In either case the continued maintenance of confidence in the information is important because once the information becomes known outside the confidential relationship it loses its protection⁴⁶ except, to some extent, as against a wrongful discloser or user who is acting in breach of confidence.⁴⁷

Strictly speaking, the law of confidence does not protect the information itself, but rather protects against breach of the confidential relationship. In practice, though, such breaches occur through disclosure or misuse, and these are the very things which the owner wishes to control. The most important remedy is an injunction to prevent disclosure or misuse, which can sometimes be granted if a breach of confidence is anticipated but has not yet occurred.⁴⁸ Additionally, if a breach has occurred, damages can also be awarded.

However, a cloud customer cannot just do nothing and expect the law to protect the customer's confidential information. An obligation of confidence will only exist if the customer has a relationship of confidence with the cloud provider in respect of the information in question.⁴⁹ That relationship can arise automatically, which at a minimum requires that the recipient should have been aware that the information is confidential. In practice cloud providers will have no way of deciding which parts of the customer's information are confidential, nor will customers necessarily understand how far the information made available by the provider is also confidential.

To resolve this uncertainty it is almost universal for cloud ToS to contain confidentiality obligations.⁵⁰ In business relationships it is likely

⁴⁶ *Attorney-General v Guardian Newspapers Ltd (No 2)* [1990] 1 AC 109; *Public Systems Inc v Towry and Adams* (Ala. 1991) 587 So. 2d 969 (Alabama, US).

⁴⁷ *Seager v Copydex* [1967] RPC 349.

⁴⁸ See e.g. *Personnel Hygiene Services Ltd & Ors v Rentokil Initial UK Ltd (t/a Initial Medical Services) & Anor* [2014] EWCA Civ 29. See also EU Directive 2016/943 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L157/1, Article 10.

⁴⁹ *Saltman Engineering v Campbell* [1948] 65 RPC 203.

⁵⁰ If there are no express terms about confidentiality, or other aspects of information ownership, a court is likely to find implied terms. The difficulty with implied terms is that although the court's analysis of their content is derived from the nature of the relationship between the parties, this analysis occurs after the event. By definition, the court would not be asked to determine the terms if the parties had a common understanding of the permitted uses of information within that relationship, and so one of them is certain to be disappointed – see

that the cloud provider will agree that all information provided or generated by the customer is confidential, which means that if a breach occurs the only question is whether the information at question was in fact confidential. In consumer cloud relationships the position is much more complicated because in many of them, social media services in particular, disclosure of information to others is an important part of the service. Just as one example, if a post is made on Facebook and marked for disclosure to friends only, does that create an obligation of confidence on the part of Facebook so that if it were made public Facebook would be in breach? Does it make a difference if the customer has 10 friends, or 10,000 friends? It is far from clear how the law of confidence would decide these questions, and Facebook's ToS make no attempt to do so. Instead, Facebook promises its UK customers that it will only use information in accordance with the customer's user settings,⁵¹ so that in our example the claim against Facebook would probably be dealt with as a breach of contract, rather than a breach of confidence.

3. INFORMATION PROVIDED OR GENERATED BY THE CUSTOMER

3.1 INTRODUCTION

Customers will often use cloud platforms to store information which they have created or acquired, and to generate new information. But these platforms are owned and operated by the cloud provider, not the customer. As a result, customers may be concerned that they may lose 'ownership' of, or at least control over, such information.

The overvaluing of ownership, probably based on its psychological and emotional aspect as discussed in section 0, may lead to an excessive focus on maintaining intellectual property rights, even though there is no appreciable likelihood of any loss of ownership here. It also means that insufficient attention is paid to the question of control over that information, which *is* an area where cloud use presents real risks which need to be guarded against. Achieving a solution is challenging.

A conceptual map of the information flows in the cloud provider / cloud customer relationship might look something like Figure 5.1.

the Australian decision in *Trumpet Software Pty Ltd v OzEmail Pty Ltd* [1996] 34 IPR 481 (Federal Court of Australia).

⁵¹ Facebook UK Terms of Service (4 January 2022 version) clause 3.3.1, <<https://en-gb.facebook.com/legal/terms>> accessed 7 February 2022.

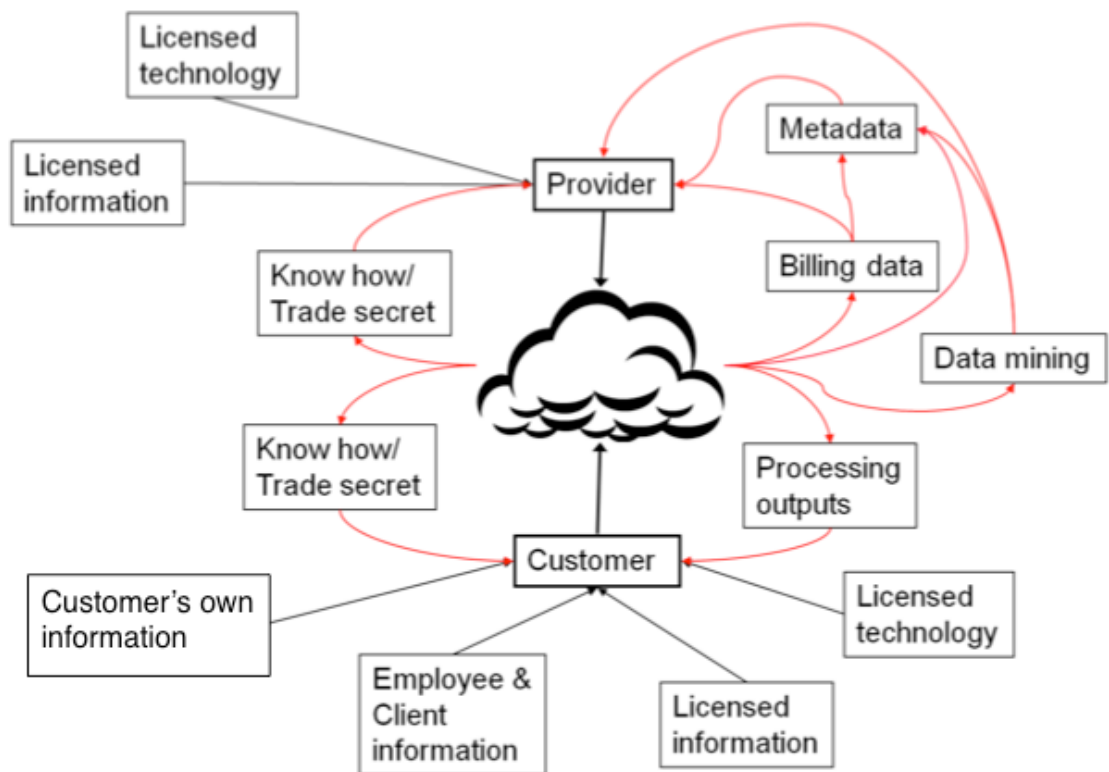


Figure 5.1 Information flows in the cloud

Although this map is complex, it can be simplified by making a distinction between information which originates outside the cloud and is brought in by the customer and information which the customer generates in the cloud via the cloud service, and by dealing with the cloud provider's 'own' information separately.

3.2 INFORMATION PLACED IN THE CLOUD

Where information has originated outside the cloud and is brought into the cloud service by the customer, that information will already have an established ownership status. This information will most commonly take the form of data, which the customer intends to process using the cloud service, or software which will run on the cloud infrastructure or platform, or documents which the customer wishes to store and perhaps modify. The likely expectation of all parties is that merely placing this information in the cloud will not per se change its ownership.

The map in Figure 5.2 shows only the information which an enterprise cloud customer is likely to bring in, or the information the cloud provider brings in. The ownership issues are the same for both.

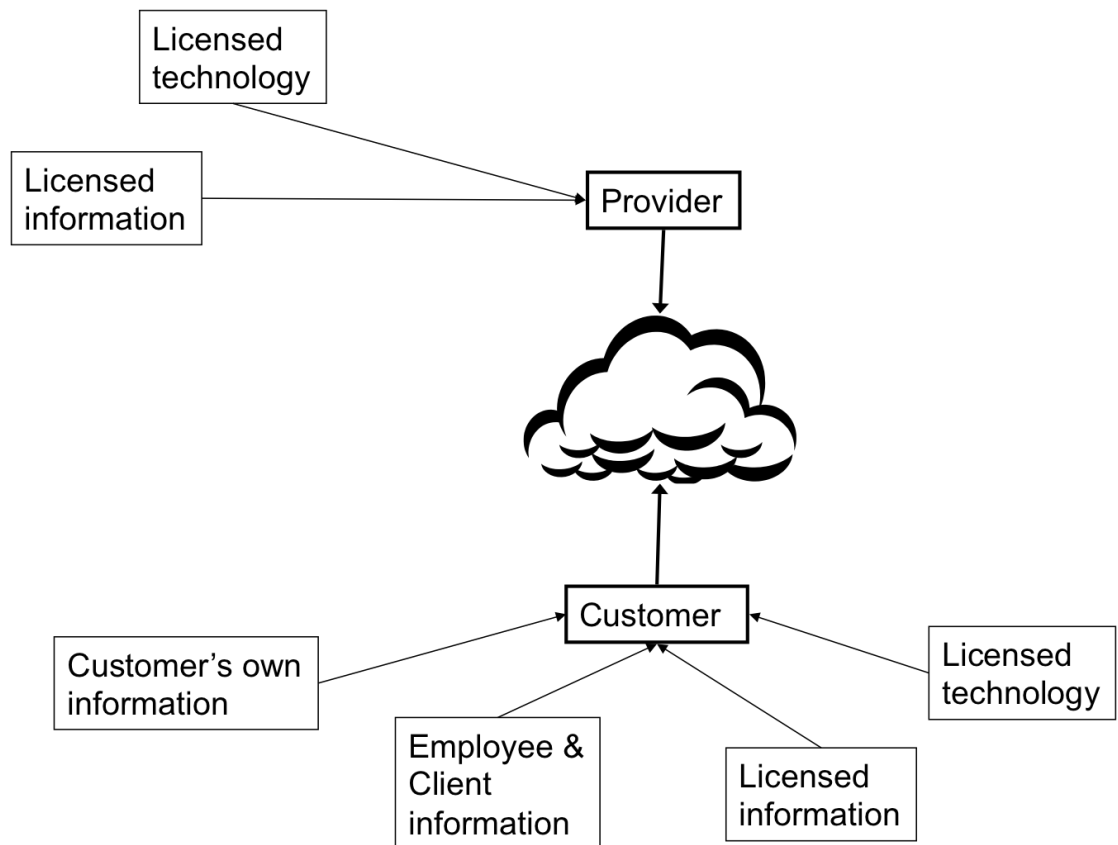


Figure 5.2 Information brought into the cloud

If the customer generated the information itself, then the customer will normally own any intellectual property rights in it. If the information was generated by a third party and licensed for use by the customer, again its intellectual property rights will already have been established.

Although customers may fear that uploading information to the cloud creates a risk to their ownership, that fear is entirely irrational. The only way ownership of IP can be lost is via a formal assignment of ownership or, in some countries, a dedication of the work to the public domain. The mere uploading of information to a cloud platform or service does neither of these things.

Of course, the customer will be handing over at least some level of control over the information to the cloud provider, and so is likely to want the service provider to ensure that the customer's ownership rights are respected, along with the rights of any third party owner. It is almost universal practice that the cloud ToS state that the cloud provider will do

so,⁵² probably because the expectation of all parties is that the mere fact of placing the information in the Cloud should not alter its ownership status.

Consumer cloud use is more complex, because millions or even billions of sound recordings, images and videos are stored and shared on a daily basis via social media services. The complex norms of information sharing between consumers largely ignore questions of intellectual property ownership, even though sharing does not change that ownership, and so of course many of these acts of storage and sharing infringe third party rights. Until now, cloud providers in this field have been protected by liability immunities, such as that under the EU E-Commerce Directive⁵³ Article 14, so that they only have liability for infringing content if they fail to take action after receipt of notice from the rights owner. As a consequence, consumer cloud ToS say little on this topic – for example, Facebook’s UK ToS provide only that:

‘You may not use our Products to do or share anything ... that infringes or breaches someone else's rights.’⁵⁴

However, the EU Directive 2019/790 on copyright and related rights in the Digital Single Market⁵⁵ imposes liability on online content-sharing service providers for communicating infringing material uploaded by their users. In addition to the notice and take-down regime already established by the E-Commerce Directive, the larger providers will be required to make best efforts to secure licences from content owners to communicate works to the public, and also to make best efforts to block communication of works specified in advance by content owners.⁵⁶ This will focus their attention

⁵² Johan David Michels, Christopher Millard, and Felicity Turton, “Standard Contracts for Cloud Services”, in: C. Millard (ed.) *Cloud Computing Law*, (2nd edn, OUP 2021). See also Johan David Michels, Christopher Millard and Felicity Turton, “Contracts for Clouds, Revisited: An Analysis of the Standard Contracts for 40 Cloud Computing Services” 11 June 2020 <<https://ssrn.com/abstract=3624712>> accessed on 14 February 2022.

⁵³ Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (‘Directive on electronic commerce’) [2000] OJ L 178/1.

⁵⁴ Facebook UK Terms of Service (4 January 2022 version) clause 3.2.1, <<https://en-gb.facebook.com/legal/terms>> accessed 7 February 2022.

⁵⁵ EU Directive 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (text with EEA relevance) [2019] OJ L130/92 (hereafter ‘EU Directive 2019/790’). The Directive came into force on 6 June 2019 and the deadline for EU Member States to implement the Directive in their national laws is 7 June 2021.

⁵⁶ EU Directive 2019/790 (n 53), Article 17.

more closely on infringing conduct by users, and ToS are likely to be amended accordingly.

In the same way that ownership of intellectual property rights is not changed by uploading information to the cloud, so also an owner's rights of confidentiality will not be affected so long as all those who thereby have access to the information are under an obligation to maintain its confidence. The nature of commercial cloud computing relationships would seem to suggest that the service provider impliedly undertakes to maintain confidence in the customer's information, and this may be stated expressly in the ToS.⁵⁷

Again, the issue is more complex in relation to information-sharing services such as Facebook or LinkedIn, which are used to maintain a one-to-many relationship with 'friends', network members and others. Here the nature of the provider's confidentiality obligations is less obvious, because of the more open nature of the relations between users. This makes drafting comprehensive contractual terms very difficult – for example, the LinkedIn ToS merely provide that users will not:

Disclose information that you do not have the consent to disclose (such as confidential information of others (including your employer)).⁵⁸

Of course, any consent to disclose will sometimes be express, sometimes granted implicitly through the relationship between the user and the information 'owner', and sometimes granted implicitly via the context in which the information was received. Any attempt to capture the complexity of the normative expectations of users in more detailed terms is probably a hopeless task.

3.3 INFORMATION GENERATED IN THE CLOUD

The fact that the customer generates information in the cloud by using the cloud provider's facilities, software tools and possibly even data, does not affect the customer's status as copyright author or database creator. If parts of those creations incorporate information which the service provider owns, the cloud ToS could in theory make it clear what, if any, restrictions this places on the customer's use of the new creation. However, a survey of cloud ToS in 2019 suggests that cloud providers do not attempt to do this,⁵⁹ but rather simply state that each party (or, where applicable, third parties)

⁵⁷ 'Salesforce Master Subscription Agreement' clause 7, <https://www.salesforce.com/content/dam/web/en_us/www/documents/legal/salesforce_MSA.pdf> accessed 7 February 2022.

⁵⁸ LinkedIn User Agreement (11 August 2020 version) clause 8.2.5, <<https://www.linkedin.com/legal/user-agreement>> accessed 7 February 2022.

⁵⁹ Johan David Michels, Christopher Millard and Felicity Turton, "Standard Contracts for Cloud Services" (n 52).

owns all the intellectual property rights in the information which it brings within the service. This leaves the question of ownership of creations which mix the two to be determined after the event. The survey shows that use of all this information is normally cross-licensed, meaning that a customer who makes a mixed creation will not infringe the provider's or any third party's rights by doing so, but this does not mean that the customer is sole owner of the mixed creation.

It should not be surprising that cloud ToS do not attempt to define the ownership position for mixed creations. Some of the provider information will be protected by copyright, but only its expression and not the information content per se. So if protected expression is copied, the mixed work will be owned jointly by the customer and the owner of the included expression as most likely a derivative work.⁶⁰ The information in a database does not receive any protection via database right, so incorporating that in the mixed creation means the customer will still be the owner of that creation. If what is incorporated is confidential information, then the customer will be able to use the mixed creation for its internal purposes because of the cross-licence but will not be able to disseminate or exploit it in ways which would breach the obligation of confidence. All this is so complex, and so dependent on what information is incorporated and how it is used, that cloud ToS concentrate on ownership rights in the information which customers and cloud providers bring within the service,⁶¹ and do not attempt to deal with the mixed creation issue.

The author of any copyright work will be the customer, or more likely, an employee acting in the course of employment which means that copyright in the information will usually belong to the employer.⁶² So long as the applicable law is that of a Berne Convention⁶³ country then copyright comes into existence at the moment the work is created, which under English law is when it is recorded,⁶⁴ and that copyright subsists in all other Berne countries without the requirement for further formalities.

The lack of a fixed geography for many cloud services means, however, that there may be unexpected answers as to *where* the information was first recorded. This is relevant if the work is published online, because the place of publication is its country of origin for copyright purposes. This

⁶⁰ See *Kogan v Martin* [2019] EWCA Civ 1645 paras 33-35.

⁶¹ Johan David Michels, Christopher Millard and Felicity Turton, "Standard Contracts for Cloud Services" (n 52).

⁶² Copyright, Designs and Patents Act 1988, s 11(2) for the UK. Under other national copyright laws employers seem universally to own the economic rights in works created in the course of employment, although the position for moral rights may vary.

⁶³ Berne Convention for the Protection of Literary and Artistic Works (9 September 1886) (hereafter 'Berne Convention').

⁶⁴ Copyright, Designs and Patents Act 1988 s 3(2).

place is most probably the location of the server from which it was first made available to others, if that can be discovered.⁶⁵ As we have seen in section 0, different countries have different standards of creativity, and so for some functional works the place of publication will determine if the work attracts copyright protection at all.

If it is unpublished, as is likely for documents produced for internal purposes, the country of origin is that of which the author is a national. Thus the place of recording is immaterial for the question whether copyright subsists.⁶⁶

The place of recording is also irrelevant to the question of infringement. Provided the work is in fact protected by copyright, national law differences from its place of origin are unimportant because under the Berne Convention a work receives ‘national treatment’, ie any claims for infringement are decided by the law of the country where the infringement occurs, so long as the work has a country of origin in a Convention member state.

However, the potential uncertainty about country of origin of published works in the cloud can lead to litigation procedure issues. The occurred in the US case *Moben v 335 LLC*,⁶⁷ where the defendant was sued for copyright infringement of a graphic work uploaded to a server in Germany from which it immediately became accessible world-wide. There was no dispute that the defendant had made copies of the work in the US, without licence, but the defendant argued that the immediate accessibility from the German server also amounted to first publication in the US (and, presumably, in every other country). If this were so, the action could not be brought until the plaintiff had registered copyright in the work in the US. The court rejected this argument, holding that first publication had occurred on the server to which the work was uploaded, that is in Germany.

The EU database right is very different because it was created only to provide rights for EU databases, and only against infringement in the EU. Thus database right only subsists if the database’s ‘makers or right

⁶⁵ This point has not definitively been decided anywhere and raises difficult conceptual issues – see ALAI Country of Origin Study Group, *Determination of Country of Origin When a Work Is First Publicly Disclosed Over the Internet* (2011), <<https://www.alai.org/en/assets/files/resolutions/country-of-origin.pdf>> accessed on 7 February 2022. If publication is via a cloud platform the precise server may be unknown, unpredictable and quite possibly merely fortuitous. A pragmatic approach, such that of the CJEU in *Football Dataco Ltd and Ors v Sportradar GmbH and Anor (Legal protection of databases)* [2012] Case C-173/11 ECLI:EU:C:2012:642 (hereafter ‘*Football Dataco C-173/11*’), is likely to be the only solution, and the place of establishment or habitual residence of the publisher would be an attractive choice.

⁶⁶ Berne Convention (n 63), Art 5(4)(c).

⁶⁷ Unreported, D Delaware 6 October 2009.

holders are nationals of a Member State or ... have their habitual residence in the territory of the Community.’⁶⁸ This wording has some unexpected consequences – for example, a French national who had been resident in Hong Kong for several years and created a database there would still benefit from database right if unauthorised extraction or reutilisation occurred in an EU country – but anomalies of this kind are to be expected from legislation which substantially pre-dates the arrival of the cloud. At least, though, it will be reasonably clear whether a database benefits from the right.

More challenging is the question of infringement if either the hosting of the database itself, or the location of the person extracting or reutilising its contents, is outside the EU’s geographical territory.

The most likely scenario involves hosting or extraction in the US, as the major cloud providers with a significant European presence are all US corporations and much of their cloud infrastructure is located there. US law provides no *sui generis* protection for databases which consist of factual information, and only limited copyright protection for any elements of creativity in a database’s structure.⁶⁹ Thus if a database protected by EU database right is hosted on a US-located cloud server, the right will only be infringed if the acts of extraction or reutilisation take place in the EU, rather than at the server. A similar though reverse problem arises if a database is hosted on an EU-located cloud server but acts of extraction or reutilisation are undertaken by a person located in the US.

In *Football Dataco Ltd & Ors v Sportradar GmbH & Anor (Legal protection of databases)*⁷⁰ the CJEU declined to decide where cross-border acts of extraction took place. It did, though, consider the question of reutilisation by means of making the database contents available to the public.

Categories based on concepts, such as time and space, the meaning of which becomes highly ambiguous in the world of virtual reality, are rendered ineffective by the networked configuration of a global communication medium, the content of which is constantly being renewed and which even today remains highly resistant to the discipline of a legislative framework that can be effective and efficient only if it is set up with the support of the international community of States as a whole.⁷¹

The Court therefore decided to adopt a pragmatic approach to the question. In that case the Austrian defendant Sportradar hosted extracts

⁶⁸ Databases Directive (n 30), Article 11(1).

⁶⁹ See *Feist* (n 25) and accompanying text.

⁷⁰ *Football Dataco* C-173/11 (n 65).

⁷¹ *Football Dataco* C-173/11 (n 65) para 55.

from the claimant's database in its home country, from which it made them available to persons in the UK. The court held that the process of making available consisted of a number of connected acts, some of which occurred in Austria and others where the targeted recipients were located, in the UK. Thus reutilisation occurred in both countries.

It is unclear whether this reasoning should be applied to acts of extraction, but a plausible case can be made. Copying parts of a database hosted in country A has a higher potential impact on the database owner's economic interests⁷² if those parts are recorded to permanent storage for subsequent reutilisation. Thus, a recording of the extracted parts in country B is a continuing part of the extraction process, so that extraction should be considered to have occurred in both A and B.

3.4 CONFIDENTIALITY AND CONTRACTS

None of the legal difficulties about infringement caused by the geographical uncertainty which cloud introduces affect the basic conclusion that IP rights subsist in information generated within the cloud. The ownership of those rights is no different than if the information had been generated outside.

The same is true of confidential information. Placing or generating it in the cloud does not change its confidential status. And here there are fewer geographical uncertainties about infringement, because a breach of confidence occurs when information is either disclosed or used when it ought not to have been. The place of infringement will therefore be the place where that disclosure or use occurs, and the applicable law of confidence will be the law of that place. If disclosure is between parties in different countries, the reasoning in *Football Dataco*⁷³ is plausible in suggesting that disclosure requires communication both from one person and to another, and therefore takes place in both locations.

Contract law in the form of terms and conditions can do little to resolve these uncertainties. The cloud ToS usually provide only that each party retains ownership of its intellectual property rights, which leaves the ownership question to be decided by the relevant law. The ToS could go further and could state that certain categories of information produced by the customer are protected by copyright or database right, but this would not actually confer these IP rights if, as a matter of law, they did not subsist in the information. Such a contract term would, of course, be binding as between provider and customer, but its legal effect would be uncertain. The applicable law might treat this term as a promise to treat the customer's information *as if* the stated IP rights subsisted, in which case a provider who did not do so would be in breach of contract. Alternatively, the term

⁷² Infringement requires the extraction or reutilisation to have appropriated a substantial part of the economic value of the database – see Section 0 above.

⁷³ *Football Dataco* C-173/11 (n 64).

might merely act as an estoppel, preventing the provider from denying that the IP rights subsisted, but this might not be enough to give the customer a positive right of action against the provider if in fact the provider's actions would not have infringed the right in any event.⁷⁴

In theory, contract law could be used to remove any uncertainties, though only as between customer and provider, if the ToS were drafted so as to replicate the desired IP rights as contractual obligations, for example, by deeming any infringement by the provider to have occurred in a specified country.⁷⁵ However, such a contract would be complex to draft and would probably need to be individually negotiated for each customer/provider relationship.

4. INFORMATION PROVIDED BY THE CLOUD PROVIDER

All that has been said in Section 0 above applies equally to the information which the cloud provider makes available to customers. The provider will retain all its ownership rights in that information, as will any third party whose information is made available in the form of software or data, and the customer's use of the cloud poses no real threats to any of those ownership rights.

Often, information originating from third parties is provided as part of the cloud service for use by the customer. Not only will the provider need a licence to provide this information, but the restrictions in that licence will also need to be replicated in the cloud ToS. This was standard in all of the cloud ToS that were surveyed in 2019, with the ownership rights of third parties preserved and respected.⁷⁶

There is, though, one category of information provided by cloud providers to their customers whose ownership remains controversial. Application Programming Interfaces ('APIs') are the technical mechanisms through which different software applications interoperate. If the customer wishes, for example, to use the cloud provider's analytics software to interrogate the customer's databases and generate useful information, then the customer needs access to the API for that analytics software. The customer (or a third party developer acting on the customer's behalf) can

⁷⁴ For example, even if the contract stated that the provider acknowledged that database right subsisted in the customer's databases, this would not give the customer any claim if the provider's unauthorised extraction and reutilisation took place on a non-EU server.

⁷⁵ The jurisdictional problems when considering infringement of IP rights are complex and uncertain – see Paul Edward Geller, 'Rethinking the Berne-plus framework: from conflicts of laws to copyright reform' (2009) EIPR 391.

⁷⁶ Johan David Michels, Christopher Millard and Felicity Turton, "Standard Contracts for Cloud Services" (n 52).

then write code which passes data from the databases to the analytics software and issues commands to that software.

There is no doubt that most of the detailed code which implements an API is a copyright-protected work. Where the uncertainty arises is in respect of those parts of an API which need to be copied in order to make use of it, that is the commands which it makes available, its overall structure, and the label names used to communicate with it. In our example there are two different types of copying which might give rise to copyright infringement allegations:

- Copying by, or on behalf of the customer, to create ‘middleware’ which would allow the customer’s code to communicate with analytics software from a different cloud provider. The customer would wish to create this middleware if it wanted to change to a new cloud provider’s service seamlessly without needing to rewrite major parts of its existing code.
- Copying by a competitor cloud provider, so that it could offer potential customers seamless migration to its own services without their needing to write middleware to replicate the functionality of that API.

It seems clear that the first of these is unlikely to infringe copyright. In the EU, the 2012 decision in *SAS v World Programming*⁷⁷ held that purely functional elements of software such as its basic functionality, programming language, or data formats were not a form of expression of the program, and therefore were not protected by copyright. It is these elements which the customer needs to copy, and doing so will not infringe copyright. Under US law, as we will see, such copying will almost certainly amount to fair use and will therefore not infringe US copyright either.

The second is more controversial under US law, though *SAS v World Programming* suggests that, in the EU, it too will not amount to infringement. US copyright law attempts to draw clear lines between the ideas contained in a work, which are unprotected, and their expression, which is protected. The early litigation over interfaces focused mainly on the human-computer interface, seeking protection for such matters as menu and command structures⁷⁸ and graphic user interfaces.⁷⁹ Eventually, the courts converged on a consensus that because there was limited choice in expressing these at the highest level, because of both technical constraints

⁷⁷ *SAS Institute Inc v World Programming Ltd* [2012] C-406/10 ECLI:EU:C:2012:259.

⁷⁸ *Lotus Dev Corp v Borland International Inc*, 49 F 3d 807 (1st Cir 1995), 516 US 233 (1996).

⁷⁹ *Apple Computer Inc. v Microsoft Corp*, 799 F Supp 1006 (ND Cal 1992), 35 F 2d 1435 (9th Cir 1994).

and user expectations about how software would work, those high level elements were unprotected ideas.

Following the resolution of the first API copyright war, the software engineering community came to view high-level functions, labeling conventions, and the functional specifications of APIs as unprotectable under copyright law.⁸⁰

In recent years, a long-running dispute between Oracle and Google over the incorporation of the Java programming interface into Google's Android operating system has reopened discussion of copyright and APIs. The factual background is immensely complicated,⁸¹ but in simplified terms Google copied the interface for the Java programming language, including a substantial number of lines of code, in order to enable developers of apps for the Android operating system to do so using the Java language. It is important to the decision that neither Sun Microsystems (who originally developed Java) nor Oracle, who acquired Sun, were competing with Google in the operating systems market.

The final decision on subsistence of copyright was given by the Federal Circuit in 2014, holding that the structure, sequence, and organisation of the 37 Java APIs which Google had copied were protected by copyright.⁸² This did not dispose of the litigation, however, because Google had asserted that its copying amounted to fair use⁸³, which had to be determined by jury trial. This trial was held in 2016⁸⁴ and the jury found that Google's copying was fair use, though without giving reasons. All that we can say for certain is that Google's commercial success with Android did not prevent it from asserting fair use successfully. Because much of the trial evidence focused on the fact that there was no competition between the parties in the operating systems market, it seems likely that this will have influenced the jury's decision.

The case is not over though. In 2018, the Court of Appeals for the Federal Circuit heard Oracle's appeal against the jury verdict.⁸⁵ Because

⁸⁰ Peter S Menell, 'Rise of the API Copyright Dead: An Updated Epitaph for Copyright Protection of Network and Functional Features of Computer Software' (2017) 31 Harv JL & Tech 308, 343 (hereafter Menell, 'Rise of the API Copyright Dead').

⁸¹ For a detailed explanation see Menell, 'Rise of the API Copyright Dead' (n 80) 343-410.

⁸² *Oracle Inc v Google Inc*, 750 F 3d 1339 (Fed Cir 2014).

⁸³ It is worth noting that fair use is an affirmative right under US copyright law, and not merely a defence to copyright infringement - *Lenz v Universal Music Corp*, 801 F 3d 1126 (9th Cir 2015).

⁸⁴ For a full account see Menell, 'Rise of the API Copyright Dead' (n 80) 391-410.

⁸⁵ *Oracle America Inc v Google Inc*, 886 F 3d 1179 (Fed Cir 2018) (hereafter '*Oracle v Google* Fed Cir 2018').

the jury verdict was, as a matter of law, purely advisory except as to findings of 'historical fact',⁸⁶ the court substituted its own view of fair use for that of the jury and decided in favour of Oracle.

Unsurprisingly, the US Supreme Court agreed to hear an appeal from that decision.⁸⁷ It overturned the Court of Appeals verdict, holding that Google's activities amounted to fair use because it had not copied any more of the Oracle code than was necessary to create its new, transformative, software. The fact that some code was copied verbatim was still fair use, because doing so was the only practical way to create the new software.⁸⁸

From that judgment, it seems plausible that if a cloud customer wishes to replicate the relevant parts of an API in order to move to a new cloud provider, the copying required to do so has a reasonable chance of amounting to fair use under US copyright law. Whether a cloud provider who wishes to reproduce a competitor's API, for the purpose of attracting that competitor's customers to its service, could plead fair use successfully, remains an open question.

5. INFORMATION GENERATED BY THE CLOUD PROVIDER WHICH RELATES TO THE CUSTOMER'S INFORMATION

5.1 INTRODUCTION

Merely in order to provide the service, a cloud provider will need to generate information. Some cloud services are funded by using the customer's information to generate new information or to enable targeted advertising. And the aggregate of customer information is a resource from which data mining tools and other data analysis techniques can discover valuable information. In all these cases the customer might feel some degree of psychological ownership, at least in the sense that customers will think they should have some right of control over the uses which the provider makes of it. But will the customer actually have any legal ownership claims?

There are four broad categories of information which the cloud provider might generate, and each needs to be analysed separately. The dividing line between them is, inevitably, blurred.

⁸⁶ *Oracle v Google* FedCir 2018 (n 85) 1196.

⁸⁷ 141 S. Ct. 1183 (2021).

⁸⁸ The Court of Appeals for the Federal Circuit placed great weight on its finding that Google's success with Android had prevented Oracle from expanding its Java SE operating system into the smartphone market, *Oracle v Google* Fed Cir 2018 (n 85) 1207-10.

5.2 INFORMATION GENERATED FOR THE PURPOSE OF SERVICE PROVISION

In order to provide and manage the service, and to improve and upgrade it, the provider needs information about the customer's activities. Such information will be captured in the form of data usage logs, billing data, and so on. If any rights in this information subsist at all, they will clearly be owned by the provider. But as we have seen, ownership does not necessarily confer exclusive control, and so the cloud customer might still assert some rights in respect of that information.⁸⁹

This provider information may contain confidential information about the customer's activities, and this may restrict any disclosures by the provider. In practice, disclosure of such information could also identify some of the provider's commercially useful trade secrets, and so the provider will not wish to disclose it. This category of data therefore seems unproblematic.

5.3 DISCLOSING INFORMATION ABOUT THE CUSTOMER

Some cloud services, particularly social media services, are specifically set up to share information between users. Much of this information is provided by the users themselves, but some of it is generated by the service provider. Thus, for example, Facebook alerts users to postings by friends and events they might be attending, while YouTube tells followers of channels when a new video has been posted to that channel. Users know in advance that these disclosures of information will be made when they undertake particular activities, and they have consented to them in the ToS.

But there are disclosures of information about the customer which are not known to them in advance, and will remain unknown to them unless they seek information about such disclosures (e.g. under data protection law). Some of these disclosures might reasonably be expected, such as disclosure for the performance of third-party payment services or credit checks. Others might not be expected at all unless the customer is familiar with the cloud ToS. None of these disclosures are likely to be of information in which the customer owns intellectual property rights, but the information may well be private or confidential. It is this element of privacy or confidentiality, together with the psychological feeling that one 'owns'

⁸⁹ One controversial example is the extent to which the right to data portability in the EU General Data Protection Regulation (GDPR) requires the cloud provider to make some of its own data available to facilitate the customer's move to a new provider – see Section 5.4 below. Regulation 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 (hereafter 'GDPR').

information about oneself, which might lead customers to expect some right to control disclosure.

The type of disclosure which customers are most likely to wish to control is disclosure to a third party to be used for marketing that third party's products or services to the customer. A survey of ToS⁹⁰ shows that service providers broadly took two approaches to this issue.

Some stated simply that they would share such information for marketing. This is what customers might expect from 'free' social media services such as Pinterest, but a number of cloud storage services also included such a provision in their ToS.

Others took a more complicated approach. Facebook (owned by Meta but still marketed as Facebook), for example, which generates much of its revenue by serving its users with targeted advertising for third-party products and services, is unlikely to want to share information with advertisers which would allow them to market their products and services directly. However, this does not mean that no sharing of personal data takes place. Indeed Facebook stated that it would share personal information about customers with companies that aggregate it, and then provide advertisers with such aggregate data.⁹¹ Other service providers say that they will share information with third parties, presumably under confidentiality terms, to market those service providers' own services.

Beneath these broad descriptions of the two approaches lies a complex nuancing of how disclosure for marketing is undertaken and explained to customers.⁹²

Because all the information sharing discussed here has the potential to disclose confidential information, we would expect a cloud provider to secure customer agreement to it via the ToS or the privacy policy.

5.4 METADATA

Metadata is data about data, in our context information about the relationships between the different elements of the customer's data, the relationships between the customer's data and provider or third party applications or data, and thus indirectly includes data about how the customer has used and processed the customer's own data. Data about the customer's use of the cloud service is also usually considered to be metadata. Metadata will be generated by the service provider, and will include or be derived in part from logs of data accesses or application usage by the

⁹⁰ Dimitra Kamarinou, Christopher Millard and Felicity Turton, "Protection of Personal Data in Clouds and Rights of Individuals" (n 14).

⁹¹ Ibid.

⁹² Ibid.

customer or its individual employees or end users, so there is some overlap with the information discussed in Section 02.

This information is generated by the provider, and so any ownership rights in it will belong to the provider. It is unlikely that any confidential information is contained in the metadata, except in the hands of the service provider, because the implicit information about how the customer has used and processed its data only makes sense in the context of that data and the applications used. However, the close connection between metadata and the customer's own information will create a strong feeling that the customer should have some rights over it.

This is most relevant to the question of data portability, if the customer wishes to move its information to a different cloud provider. Without at least some of the metadata the customer may find it difficult to replicate its current information activities on a new platform.

If EU data protection law applies, the customer has the 'right to data portability' under Article 20(1) of the EU General Data Protection Regulation ('GDPR'), but this means only that the customer must receive it 'in a structured, commonly used and machine-readable format'. The right only applies to personal data, not the rest of the customer's information in the cloud, and there is some doubt whether the right extends to metadata, which might not be personal data. The Article 29 Working Party guidelines⁹³ provide that personal data should be provided in 'commonly used open formats (e.g., XML, JSON, CSV, etc.) along with useful metadata at the best possible level of granularity, while maintaining a high level of abstraction', with the aim that this should 'allow the effective re-use of the data'. This guideline implicitly argues that metadata becomes part of the personal data and so should also be disclosed. But, as has been pointed out, the Regulation only makes the format mandatory, rather than providing that the aim of re-use should be achieved, and so this is merely recommended practice.⁹⁴

A right to receive some or all of the metadata can therefore only be found in contract via the ToS. Two-thirds of the cloud providers' ToS that were surveyed in 2019 included some provisions on data portability, either promising to make personal data available in a structured format or promising that all the customer's information would be made available.⁹⁵ However, none of these terms made any reference to metadata, so it seems that, at present, its ownership remains completely with the service

⁹³ Article 29 Working Party, 'Guidelines on the right to data portability' (2017) WP 242 rev.01, 17.

⁹⁴ Paul De Hert and others, 'The right to data portability in the GDPR: Towards user-centric interoperability of digital services' (2018) 34 Computer Law & Security Review 193.

⁹⁵ Dimitra Kamarinou, Christopher Millard and Felicity Turton, "Protection of Personal Data in Clouds and Rights of Individuals" (n 14).

provider, subject to any rights of use by the customer which might derive from the GDPR.

5.5 INFORMATION GENERATED FROM AGGREGATED DATA

ANALYSIS

This is new information generated by the provider through analysis of customer data or metadata. The fact that service providers have possession of and technological control over the customer's information makes it possible for them to use data mining⁹⁶ tools to trawl through customer information, either individually or on a collective basis, and thereby generate new and potentially valuable information. How widespread such data mining is, only the cloud providers know. As a hypothetical example, a service provider with a number of motor insurers as customers could mine their data to extract information on the accident rates and types for different makes and models of vehicle. As a real example, Facebook and Google mine customer information to generate new information about their customers' interests, activities and preferences. This new information can be exploited commercially, often for marketing purposes, and is part of the 'price' which customers (to an extent unknowingly) pay for their use of ostensibly 'free' services.

Although this kind of data analysis makes use of customer-owned information, the new information generated will not copy protected expression or contain multiple extracts from a protected database. Thus, the cloud provider will be its author, and will own any intellectual property rights in it.

This use of customer information might well infringe a customer's sense of psychological ownership, and the process of data mining and analysis could require some copying by the provider. Whether the customer might use this copying as a basis to object on legal grounds depends on the licences which the customer has granted in the ToS.

The terms of these licences are likely to be influenced very strongly by the cloud service provider's business model. The main distinction is between those services where the fees paid by the customer are sufficient to allow the provider to make a profit, and those services which are provided wholly or partly without charge. In the latter case, the provider needs to generate revenue somehow, and an obvious source is by making use of customer information.

The difference is illustrated strikingly by comparing Dropbox's ToS with those of Facebook. Dropbox for Business is a paid for service, and thus:

⁹⁶ Jiawei Han and Micheline Kamber, *Data Mining: Concepts and Techniques* (2nd edn, Morgan Kaufmann 2006).

The Customer grants Dropbox only the limited rights that are reasonably necessary for Dropbox to provide the Services.⁹⁷

This wording matches the customer's expectation, that Dropbox is acting as a paid custodian of the customer's information and thus will not use the information for its own benefit.⁹⁸

Facebook's business model is expressly based on analysing customer information, and so its ToS confer extensive use rights over its customers' information:

...you grant us a non-exclusive, transferable, sub-licensable, royalty-free and worldwide licence to host, use, distribute, modify, run, copy, publicly perform or display, translate and create derivative works of your content (consistent with your privacy and application settings) ... This licence will end when your content is deleted from our systems.⁹⁹

This term allows Facebook to make any use of the information, for whatever purpose it desires, without limitation of any kind except after the information or the customer's account has been deleted.

None of these activities present any threat to a customer's ownership of the IP rights in its information, but they do present potential threats to the information's confidentiality. The risk of a breach of confidentiality is small if the provider is only using this derived information for its internal activities of running the service, and in any event the provider will have an express or implied licence to do so. The danger of breach of confidentiality is increased, however, once the provider discloses to third parties the information it has derived.

There is, unsurprisingly, no known instance of a cloud provider deliberately disclosing derived information which identifiably links to its customers without first obtaining the customer's consent. Rather, providers aggregate customer information, analyse that data, and exploit the results. In theory the data are anonymised, so there is no risk of a breach of confidence.

Unfortunately, anonymisation of data is not a complete safeguard against confidentiality breaches. Paul Ohm has alerted lawyers to advances

⁹⁷ Dropbox Business Agreement 2022 clause 8(2), <https://www.dropbox.com/business_agreement> accessed 7 February 2022.

⁹⁸ Note the parallels with the obligations of fiduciary relationships, which might have the potential to create new forms of ownership. These issues are discussed further in Johan David Michels and Christopher Millard, "Digital Assets in Clouds" (n 3).

⁹⁹ Facebook UK Terms of Service clause 3.3.1, <<https://en-gb.facebook.com/legal/terms>> accessed 7 February 2022.

in reidentification technology, which uses the recombination of separate databases to build connections between items of data and thereby identify the person to whom they relate.¹⁰⁰ A service provider's obligation is probably only to take reasonable care to preserve confidentiality¹⁰¹ and not a strict liability for any breach however caused. Thus the provider will have fulfilled this obligation if the derived information cannot, in the light of the state of technology at the time, foreseeably identify the customer via the use of reidentification technology and thereby lead to a disclosure of confidential information. However, a wise provider will need to keep alert to new developments in reidentification, as a failure to adapt to new technologies can also amount to a failure to take reasonable care.¹⁰²

6. DATA SHARING AND DATA TRUSTS

The story of this paper has been that ownership of information does not confer on the owner the levels of control which the owner would desire if that information is to be stored and processed by others. No matter how strong the psychological feeling that particular types of information are, or ought to be, owned by one or other of the parties to a cloud relationship, focusing on legal ownership distracts attention from the real issues.

Before we conclude the story it might be instructive to examine briefly the topic of data trusts. These are legal mechanisms or structures which enable data to be shared, usually for research purposes, and thus have many similarities to cloud computing. Indeed, it is likely that any data trust which holds a substantial quantity of data will outsource the actual storage and sharing processes to a specialist cloud provider. The main difference is that a cloud provider normally stores and processes data in accordance with the instructions of the cloud customer, whilst a data trust stores and processes the data in accordance with its own aims and objectives. Indeed, one of the main points of a data trust is that those who supply data to it hand over control of their data to the trust.¹⁰³

The origin of this concept seems to lie in a 2004 article by Edwards,¹⁰⁴ which suggested that trust law might be used to impose legal duties on those who make use of personal data. The holder of the personal data would

¹⁰⁰ Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) UCLA LR 1701. For an update on reidentification risks, see Jane Henriksen-Bulmer and Sheridan Jeary, 'Re-identification attacks—A systematic literature review' (2016) 36 International Journal of Information Management 1184.

¹⁰¹ *Weld-Blundell v Stephens* [1919] 1 KB 520.

¹⁰² *The TJ Hooper* [1932] 60 F 2d 737 (ship-to-shore radio).

¹⁰³ Though note that the ceding of control might be temporary, for a defined period of time or until notice to return control is given.

¹⁰⁴ Lillian Edwards, 'The problem with privacy: A modest proposal' (2004) 18 International Review of Law, Computer, and Technology 309, 326.

be a trustee, and the data subject would be a beneficiary of the trust. Trust law would oblige the holder to use the data only in the interest of the beneficiary, and would thus produce a form of privacy and 'ownership' protection for that data.

The concept remained dormant for many years, most likely because the EU data protection directive of 1995 and the subsequent expansion of the scope of data protection law through regulatory guidance and court decisions such as that in *Google Spain*,¹⁰⁵ created an alternative and more comprehensive legal regime to protect personal data. However, recent breakthroughs in Artificial Intelligence technology, largely based on machine learning, have led to a revival of interest in the concept in this new context.

Machine learning analyses data sets to generate new insights into decision making. Examples of its successes are found in diverse areas, such as voice and face recognition, medical diagnosis, driving autonomous vehicles, and much more. The richer the data set on which the learning is based, the more likely it is to produce high quality and useful results. But much potentially useful data is, at present, held in silos by public organisations and by private entities. If that data could be shared for research purposes, real public and private benefits would likely be achieved.

The main obstacle to such sharing is that multiple entities have interests in those data sets which need to be protected. These go far beyond the privacy and data protection concerns of individuals, if the data set contains personal data. Those who currently possess the data set may have a range of ownership interests (in the wide sense discussed in this paper), as may any entities from whom parts of the data set were acquired. There may also be public interests to be considered, most obviously but not exclusively in the field of medical research.¹⁰⁶

If data sharing takes place between a small group of entities then these interests can be dealt with in a bipartite or multi-party data sharing agreement, a legal structure which is well-understood. But data sharing agreements become unworkable if (a) new sharers of data are continually being added, and (b) the aims of the sharing are likely to change over time as machine learning technology advances or researchers target new problems.¹⁰⁷ Each of these requires all parties to a data sharing agreement

¹⁰⁵ *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González* C-131/12, ECLI:EU:C:2014:317.

¹⁰⁶ As a further example, machine learning is being used in the criminal justice system to assist in the making of probation and sentencing decisions. See *State of Wisconsin v Loomis* (2016) WI 68 for a detailed analysis of the public and individual interests which are at stake in such an application of the technology.

¹⁰⁷ As a simple example, medical data shared for cancer research might later be useful for research into the improved management of hospital admissions more generally.

to consent to changes to the agreement, which could rapidly become unworkable as sharing grows.

To solve this problem, Hall and Pesenti proposed in their 2017 independent review of AI for the UK government¹⁰⁸ that data trusts could be used as the sharing mechanism. They argued that if data could be conceptualised as property for trust law purposes, then owners of data sets could transfer ownership of them to a legal trust. The trust deed would set out the aims of the trust, in accordance with which data could be shared with others, and those running the trust would owe fiduciary duties as trustees to those identified as beneficiaries. These beneficiaries could include individuals whose data was incorporated in a data set and, of course, all those with ownership interests in the data.

However, the law of trusts is far from a perfect mechanism for achieving these ends. Its underlying basis is that the trustees become the legal owners of property which is the subject of the trust, and there are substantial doubts whether data can constitute property in the legal trust sense.¹⁰⁹ And of course, not every legal system has a concept of trust law, which creates uncertainty for cross-border data sharing. Although the problem of data as property might be overcome by taking contractual licences from the data owners, because some (though not all) contractual rights are choses in action and thus capable of being trust property,¹¹⁰ there are other incompatibilities between trust law and data trusts:

- first, a legal trust must be run for the benefit of the beneficiaries, so that trustees are required only to consider the collective interests of the beneficiaries when dealing with trust property. This means that they cannot allow data to be used for some socially beneficial purposes if that use does not also benefit the legal trust's beneficiaries, that is those with legal interests in the data. The exception to this is a charitable trust, which is likely to be incompatible with data sharing for the purposes of commercial research.
- second, the trustees are obliged not to use the property of the legal trust in a way which generates benefits for themselves¹¹¹ unless the

¹⁰⁸ See Wendy Hall and Jérôme Pesenti, 'Growing the artificial intelligence industry in the UK' (*UK DCMS and BEIS*, October 2017) 45-47 <<https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk>> accessed 7 February 2022.

¹⁰⁹ Expert commentators agree that in the current state of the law it is highly unlikely that a court will classify information or data as a species of conventionally understood property – for an overview and analysis see Jeffrey Ritter and Anna Mayer, 'Regulating Data as Property: A New Construct for Moving Forward' (2018) 6 *Duke Law & Technology Review* 220, 247-52.

¹¹⁰ Halsbury's Laws of England (5th edn, 2017) vol 13, para 6.

¹¹¹ *Aberdeen Town Council v Aberdeen University* [1877] 2 App Cas 544.

trust deed specifies otherwise.¹¹² This may well deter data providers from contributing data to the trust, because they are likely to want representation on the controlling body of the data trust but may be unwilling to accept the burden of fiduciary duties.

Research conducted for the Open Data Institute in 2019¹¹³ ('ODI Report') suggests that a more useful approach to data trusts is to abandon trust law, and thus the focus on property rights in data, and to use company law and contract law to build a legal structure which focuses on permitted uses of the shared data and on balancing the interest of all stakeholders against the aims of the data trust.

The ODI Report explains in more detail than previously what a data trust might be trying to achieve:

From a legal perspective, a data trust is a mechanism for achieving a defined set of aims. At the highest level, those aims are:

- to enable data to be shared;
- for the benefit of those sharing the data, and possibly also for some, broadly conceived, public benefit purpose;
- respecting the interests of those with legal rights in the data; and
- ensuring the data is used ethically and in accordance with the rules established by the data trust; and
- ensuring that whoever holds data which is subject to the data trust rules does so safely and securely, and that data is dealt with appropriately (for example by deletion) if the data trust comes to an end.

Specific data trusts may also have one or more of the following aims and characteristics:

- collective management of individual rights and interests (including any sharing of benefits received by the data trust);
- standard set of rules etc. to govern all data sharing;

¹¹² The difficulty here is defining the permitted benefits in advance, and ensuring that the trustee's overriding fiduciary duties are not compromised.

¹¹³ Chris Reed, BPE Solicitors, and Pinsent Masons, 'Data Trusts: legal and governance considerations' (*Open Data Institute*, April 2019), <<https://theodi.org/article/data-trusts-legal-report/>> accessed 7 February 2022 (hereafter 'ODI Report 2019').

- custodian/steward makes decisions on behalf of data providers/data users; and
- ability to evolve to have new purposes, governance and working methods.’¹¹⁴

Because each data trust will have different aims and stakeholders, there is no single legal structure which is appropriate for all. Trusts will be able to choose between a range of possible structures, the most useful of which are likely to be incorporation, a purely contractual arrangement, or setting up as some kind of co-operative.¹¹⁵

The most important issue facing a data trust is how to ensure that it is actually trusted by its stakeholders, those whose information ownership or other interests might be affected by the data sharing. Legal rights secure a degree of trust, at least if they can be enforced effectively, but a system which relies on enforcement after the event is unlikely to lead owners of data sets to entrust them to the trust in the first place. This places the focus squarely on the data trust’s constitution and its governance structure and processes. The mechanism for generating and updating its aims and objectives and the rules for data sharing is likely to require representation of stakeholders, if trust is to be achieved. Similarly, those responsible for the trust’s day-to-day operations and decisions will need to be trusted, and the governance structure will need to provide for appropriate oversight. Achieving these ends is a complex task, but can be achieved using well-established constitutional and governance structure principles from other fields.¹¹⁶

As with cloud computing, information ownership within a data trust is comparatively unimportant, except as an interest which needs to be protected. The way to protect this, and other interests, is (as for cloud) by focusing on controlling the use of information.

7. CONCLUDING REMARKS

This paper has shown that owning property rights in information, most likely copyright and database right, has little importance in terms of the cloud relationship.¹¹⁷ Neither customer nor provider has any real interest in acquiring property rights over the other's information, and the ways in

¹¹⁴ ODI Report 2019 (n 113) s 1.4.

¹¹⁵ ODI Report 2019 (n 113) s 2.1.3.

¹¹⁶ ODI Report 2019 (n113) s 6.

¹¹⁷ Though of course, the possibility that some other form of proprietary right in information which is part of an individual's digital persona might be recognised by law would be an important development – see Johan David Michels and Christopher Millard, “Digital Assets in Clouds”, (n 3).

which information is used and processed in the cloud present few risks to those rights.

Cloud computing does, though, create some risks to confidential information because each player in the cloud is handing over some element of control to other players. The law of confidence, together with the express confidentiality obligations of the cloud ToS, provides some protection here. However, the most important protections probably lie in the cloud provider's mechanisms to preserve the security and integrity of information. Once confidentiality is lost, the best that legal mechanisms can attempt to do is to limit further dissemination of the information and to provide some level of compensation. Compensation will usually be inadequate because lost confidentiality can never fully be recovered.¹¹⁸

A customer's psychological desire to control how the cloud provider uses the customer information for its own benefit is not, as we have seen in section 0, something with which even the wider concept of ownership discussed in this paper can assist. A customer who is seeking this kind of control can only achieve it through ToS which restrict the ways in which the cloud provider can use customer information.

In order for a cloud service to work, the service provider must necessarily make some uses of the customer's information. The ToS will grant permission for this, in the form of a licence. So how can a cloud customer ensure that the licence does not permit uses which the customer would wish to constrain? There are several reasons why this is an intractable problem.

First, the actual uses of information which the service provider makes are largely invisible to the customer. This is because of information's non-rivalrousness; no matter what the provider is doing with the information, the customer can continue to use it unhindered. This is very different from physical property – for example, if I lend my car to another person I cannot use it myself, and can to some extent monitor the borrower's use, perhaps by checking the distance driven and inspecting the car for damage. All that most cloud customers can do to discover the likely uses of their information is to read the disclosures made by the provider, either in the ToS, the privacy policy or the service description.

These disclosures are unlikely to be very informative because, outside negotiated cloud relationships, the service provider is likely to have

¹¹⁸ The same is, of course, true of loss of control of personal data, and for this reason the Court of Appeal in *Lloyd v Google* recognised that calculation of damages for breach of a data subject's rights might need to be on the basis of a nominal tariff per data subject, unless specific loss of damage were proved by an individual. *Lloyd v Google* (n 15) para 77. Following the UK Supreme Court decision, discussed above, damages were not claimable in that case, but the different wording of the GDPR suggests that they might now be.

little knowledge of what limitations the customer would wish to place on use of its information. Indeed, each customer is likely to desire different restrictions, or at least would do so if it considered the matter. Further, the provider may well be unsure about its future plans and how they will require use of customer information; cloud business models are continuing to evolve rapidly. All this leads to very open drafting of licence terms, under which the provider is granted a wide range of permissions to use customer information.

Finally, there is little or no scope to negotiate ToS. The vast majority of cloud services are sold as commodities, which customers can either take or leave. Even in major commercial cloud transactions, the scope for negotiating limitations on information use is very limited.¹¹⁹

All of this tells us that the gap between psychological feelings of ownership and the legal conception of ownership is as wide as ever in the cloud, and is unlikely to be closed via cloud ToS. Fortunately, the gap is only a real concern so far as use of customer information by the provider for its own purposes, or those of third parties, is concerned. The solution to this is not really a legal one. The best that cloud customers can do is to seek transparency from cloud service providers about such uses, and make their choice of cloud service accordingly.

¹¹⁹ W Kuan Hon, Christopher Millard and Ian Walden, 'Negotiating Cloud Contracts - Looking at Clouds from Both Sides Now' (2012) 16 *Stan. Tech. L. Rev.* 79.