The evolving battle against illegal file-sharing: some data protection observations

An article on the data protections issues raised by the collection and processing of internet protocol addresses for the purpose of copyright enforcement.

Robin Hopkins, 11 KBW, London

News of a data security breach at law firm ACS:Law in September 2010 once again highlighted the practice of collecting internet protocol (IP) addresses to facilitate litigation against suspected illegal file-sharers. As a result of a number of related legislative and judicial developments, the tension between copyright protection and data protection has recently come to a head.

The Digital Economy Act 2010 (DEA 2010) gives statutory force to co-operation between internet service providers (ISPs) and owners of music and film copyrights in identifying internet subscribers from whose IP addresses files appear to have been illegally shared, and then taking action against them.

The DEA 2010 is shortly to be judicially reviewed on the application of two ISPs: BT and TalkTalk (see *Legal update, High Court agrees to judicial review of Digital Economy Act (www.practicallaw.com/2-504-1674)*). One ground of challenge is that the DEA 2010 breaches Council Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data (Data Protection Directive), upon which the UK's Data Protection Act 1998 (DPA) is based.

Rather than providing a commentary on that case, this article will consider more generally the data protection issues arising from UK law's evolving attempts to tackle online copyright infringement. In particular, it will explore whether data protection law, as currently interpreted in the UK (note both of those qualifiers), can and should accommodate most of the practices employed by copyright owners to enforce their rights.

Copyright enforcement and the DEA 2010: past, present and future

The UK's legal framework for combating illegal file-sharing is, like that of the EU, a work-in-progress. In the recent past (that is, before the DEA 2010), action against illegal file-sharers was taken through the common law by the copyright owners themselves.

With the aid of purpose-built detection software, owners or their agents would prowl the internet in search of IP addresses from which unlawful uploads and downloads take place. Without additional information,

however, an IP address (comprising a string of numbers linked to a single internet subscription) does not reveal the identity of the relevant subscriber associated with that IP address. A Norwich Pharmacal order would therefore be used to obtain the subscriber's personal details from the ISP.

Legal action for copyright infringement would duly be taken or threatened, often by firms such as ACS:Law, whose practice of sending copyright enforcement letters to thousands of alleged infringers asking them to pay a fee of £500 to avoid legal action has been somewhat controversial. In a recent decision relating to 27 cases where the firm eventually issued claims, the London Patents County Court criticised it for its methods arguing that the letters it sent to persons accused of file-sharing materially overstated the untested merits of the claimants' approach (see *Legal update, Patents County Court refuses permission to discontinue claims in ACS:Law case (www.practicallaw.com/5-504-7556)*).

Many observers, including the European Data Protection Supervisor (EDPS), have also argued that these practices are highly invasive of the individual subscribers' private sphere as they entail the generalised monitoring of internet users' activities, including perfectly lawful ones (see *Legal update, EDPS issues opinion on proposed Anti-Counterfeiting Trade Agreement (www.practicallaw.com/2-501-5496)*).

Initial obligations under the DEA 2010

The DEA aims at, among other things, tidying up this messy and controversial situation. Action against copyright infringement remains a matter for owners and ISPs rather than for a state regulator, but this is now to be governed by the new sections 124A to M of the Communications Act 2003 (CA 2003), as inserted by sections 3 to 16 of the DEA 2010. These provisions began trickling into force since Royal Assent was granted to the DEA 2010 in April 2010, but they do not bite immediately.

Some, namely sections 124A to E, will only be effective upon the promulgation of an Initial Obligations Code, currently being finalised by Ofcom (draft Initial Obligations Code) (see *Legal update, Ofcom publishes draft code on ISPs' initial obligations under Digital Economy Act (www.practicallaw.com/0-502-4369)*), and thus represents the imminent approach to copyright enforcement. Others, namely sections 124G and H, will come into effect if and when the Secretary of State so orders, once he has considered another statutory report to be prepared by Ofcom. This evolutionary stage is yet to begin, and thus represents the (potential) future of copyright enforcement.

Under the imminent approach introduced by the DEA 2010, the copyright owner continues to gather suspicious IP addresses through its own means. It then presents to the ISP a "copyright infringement report" listing IP addresses at which infringements appear (note the statutory wording) to have taken place (*sections*

124A(1)-(3), CA 2003, as revised by the DEA 2010 and *section 4, draft Initial Obligations Code*). The ISP must then do two things:

- It must notify the subscribers of the allegations against them, how to appeal against those allegations and how to avoid online copyright infringement in future (*section 124A(4)-(9), CA 2003* and *section 5, draft Initial Obligations Code*).

- It must keep a record of the number of reports about each subscriber, and compile, on an anonymous basis, a copyright infringement list of those who have been the subject of at least three valid allegations from an ISP over the preceding 12 months (*sections 124B, CA 2003* and *section 6, draft Initial Obligations Code*).

The copyright owner may then seek a court order for disclosure of the personal details of those subscribers who, through this three-strikes process, appear on the ISP's list.

Technical measures under the DEA 2010

The future, as envisaged by sections 124G and H of the CA 2003, retains the three-strikes approach, and supplements it with a novel type of sanction to which the DEA gives the enigmatic title of "technical measures". The offending subscriber faces not only potential prosecution by the copyright owner, but also the potential restricting, slowing down or disconnection by the ISP of his internet connection. A subscriber so punished would then need to open a new internet account with a different ISP if he wanted fully-functional private access to the internet.

This is copyright enforcement at its most muscular. In the UK, it currently exists as a mere **possibility**, to be pursued only if the letter-writing approach fails to curtail copyright infringement. Sections 124G and H of the CA 2003 authorise only research reports and other preparatory steps, with further legislation being needed to empower the Secretary of State to order the taking of technical measures.

At an international level, progress in this direction is rather more advanced. The Anti-Counterfeiting Trade Agreement (ACTA), a multi-lateral agreement currently being negotiated by the EU and others, proposes to combat illegal file-sharing by encouraging signatories to adopt precisely this regime, commonly known as graduated response (see *Legal update, European Commission publishes final ACTA text (www.practicallaw.com/0-503-9196)*). It is this proposal which elicited from the EDPS the unfavourable opinion alluded to earlier. His stance represents data protection law at its most restrictive.

The polar opposite is embodied in two judgments of the Irish High Court in 2010, both handed down by Charleton J. In *EMI Records (Ireland) Ltd and others v Eircom Ltd [2010] IEHC 108*, the court held that "three strikes followed by technical measures" (here implemented not under statute but pursuant to a compromise agreement between a copyright owner and an ISP) was data protection-compliant. Similarly, in *EMI Records (Ireland) Ltd v UPC Communications Ireland Ltd [2010] ECDR 17*, the court held that, while it lacked the power to order an ISP to impose technical measures, there were no data protection barriers to such measures.

To date, UK courts have not been called upon to take a position on this spectrum between the EDPS and Charleton J.

Data protection concerns

All of the models summarised above (be they past, present or future) raise privacy and data protection concerns, in part because they utilise somewhat blunt instruments from the outset.

The process of detecting suspicious file-sharing involves a private company (rather than a judicial or regulatory body) monitoring the online activity of thousands of potentially innocent and unknowing subscribers. Even if a suspicious IP address is identified, the subscriber associated with that IP address might not actually be the perpetrator of the alleged offence: their connections might have been used by, say, other members of their households, or, with unsecured connections, someone next door. Even a subscriber who admits the relevant acts might very well escape a finding of guilt in court, given the intricacies of copyright law.

At the same time, aside from potential legal proceedings (and the costs associated with them), suspects face the prospect of accusatory letters and, in the future, the potential restriction or termination of their internet access.

These concerns underpin the stance taken by the EDPS and digital campaigners like the Open Rights Group against the evolving legal framework for combating copyright infringement (see *Open Rights Group*). However, for the purposes of this article, the ultimate question to be answered in this context is whether the actions of copyright owners and ISPs amount to breaches of their duties under the UK's DPA.

Does the DPA apply?

Before answering this question, it must first be established to what extent the duties under the DPA are in fact engaged?

DPA duties are only engaged where the requisite definitions under section 1 of the Act are met; namely, where **data controllers** process the personal data of **data subjects**.

A data controller is defined as the person who (either alone, jointly or in common with other persons) determines the purposes for which, and the manner in which, any personal data is, or is to be, processed. A data subject is any individual about whom personal data is processed. Processing is broadly defined to include obtaining, recording, holding, using, disclosing or erasing data (*section 1(1), DPA*). In effect, any activity involving personal data will fall within its scope.

This case involves copyright owners, ISPs and subscribers, each of whose accounts with ISPs is linked to a unique IP address. For present purposes, it is useful to divide subscribers into bystanders (those whose online activity the owner sees during its hunt for infringers, but who do not appear to be doing anything wrong and who are therefore passed over), and suspects (those against whom the owner seeks to take action for alleged wrongdoing).

Broadly, the copyright owners approach to enforcement involves the processing of three types of data:

- IP addresses.

- Event data, that is, information about IP address A (apparently unlawfully) having uploaded or downloaded file B at date and time C.

- Personal details (names and contact information) for the individual subscriber matching a given IP address.

Under all of the copyright enforcement approaches summarised above (past, present or future), the copyright owner carries out the following activities:

- For **bystanders**, it arguably obtains but, as far as we are aware, certainly does not record IP addresses and event data.

- For **suspects**, the owner obtains, records and holds both IP addresses and event data. These it discloses to the ISP. Later, it can (under the DEA 2010) obtain, upon request from the ISP, a copyright infringement

list, again containing IP addresses and event data. If it wishes to sue suspects, it can apply to have the suspects' personal details disclosed to it by the ISP.

The ISP's activities depend on the legislative regime under which they act:

- Before the DEA 2010, ISPs carried out almost no discernible processing until a Norwich Pharmacal order was made for disclosure of personal details to the owner.

- Under the imminent DEA 2010 regime, it first obtains IP addresses and event data from the owner in the form of copyright infringement reports. It holds this information, combines it with the associated personal details, discloses it to the suspects and, after three warning letters and upon the application of the owner, discloses personal details to the latter.

- The future regime of "three strikes followed by technical measures" would involve additional consequences, but no additional processing worth analysing here.

It is clear from this that the requisite definitions under section 1 of the DPA will be met in respect of **most** of the activities involved: copyright enforcement is likely to involve the processing by data controllers (owners and ISPs) of the personal data (personal details, for example) of data subjects (subscribers) at some point during the process.

It is not clear, however, that this holds for **all** of the activities with which we are concerned. In particular, it is not clear that all of the data involved meets the definition of personal data upon which DPA duties are premised.

The limits of personal data

Section 1 of the DPA tells us that data is personal if it relates to a living individual "who can be identified (a) from those data, or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller". We are also told that personal data "includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual".

In the present context, this latter part of the definition will encompass the owner's opinion that an infringement has been committed and that the owner intends to take action for this infringement, but only if

such data concerns the individual, by which is meant the subscriber who can be identified in one of the two ways mentioned above.

In the UK, the definition of personal data has been applied relatively restrictively: the data must, for example, be "biographical in a significant sense", such that the data subject is one of its focuses (*Durant v Financial Services Authority [2003] EWCA Civ 1746*). This need not detain us here as these thresholds would certainly be met by the data with which we are concerned. For present purposes, the crucial point about the statutory definition of personal data is that the essence of personal data is its utility in identifying someone in one of two ways (and, it seems to follow as a matter of construction, only two ways).

When one considers this point in the context of the ISP's actions, the matter is clear: all three types of data with which we are concerned here (IP addresses, event data and personal details) are personal data, because the ISP is a data controller for the personal details of subscribers. The ISP is therefore bound, by section 4(4) of the DPA, to abide by that Act.

Can the answer differ if one considers matters in the context of the copyright owner's actions? In other words, can the same data be personal in the hands of one party but not in the hands of the other?

Unfortunately, there does not seem to be any unequivocal authority on these questions in the jurisprudence of EU member states. In at least some cases, an IP address has been treated as personal data regardless of who holds it (see, for example, the decision of the Berlin Court of Appeal of 6 September 2007 (*23 S 3/07*). On the other hand, the EDPS, in its opinion on ACTA, considers IP addresses and event data to be personal data "in any case under the relevant circumstances" (see *EDPS: Opinion on the ACTA*). The implication is that it is not permissible to answer the question "is this personal data?" by saying "it depends on who is holding it".

UK jurisprudence does, however, seem to sanction precisely such an approach. The phrase "in the hands of" comes from the speech of Hope LJ in *Common Services Agency v Scottish Information Commissioner [2008] UKHL 47*. It has been applied by the Information Tribunal in *Department of Health v Information Commissioner and another (EA/2008/0074)* where the Tribunal held that the disputed abortion statistics, although entirely anonymised, did constitute personal data because they were not anonymous in the hands of the data controller. EU case law does not appear to rule out this "in the hands of" approach to the question of personal data. Nor does there appear to be any reason, in principle, why the legal status of data should not be a context-dependent matter.

Take a hypothetical example: suppose that, pursuant to a pay transparency initiative, an employer publishes a list of figures consisting only of the salaries it pays, and suppose also that there is only one employee on each of these salaries. These figures would constitute personal data in the hands of the employer, but it does not seem to follow that they are, for that reason, personal data in all circumstances.

To continue the hypothetical example, suppose a pressure group performs data analysis on the published salary points: it would not seem correct to conclude that this pressure group was thereby processing personal data. In other words, despite the passive wording used in the definition under section 1 of the DPA ("can be identified"), personal data is not necessarily an all-or-nothing classification.

If this is right, the upshot is significant: just because all relevant data is personal in the hands of the ISP, it does not follow that it is personal in the hands of the copyright owner. If it is not, then the copyright owner is not bound by the duties imposed by the DPA.

Of the three types of data with which we are here concerned (IP addresses, event data and personal details), only the latter is intrinsically personal data, regardless of its context. From the other two types alone, nobody (other than possibly the subscriber himself) could identify any living individual. There are, therefore, two arguments in support of IP addresses and event data constituting personal data in the hands of the copyright owner. One is based on self-identification, the other on the likelihood of the owner obtaining the corresponding personal details so as to facilitate legal action against the suspect.

The first argument asserts that data is personal if the data subject is able to identify himself from that data. It is not yet established whether or not this argument is correct. On this point, the wording of the definition contained in section 1 of the DPA is unhelpful (employing the passive "can be identified") and authority is scant.

In the *Department of Health* case, self-identification was not treated as relevant to the question of whether data was personal or not. Rather, the high probability of self-identification was a factor to be considered when assessing the consequences of disclosure. The Tribunal observed that self-identification "does not enable a data subject to be identified by another" and therefore, the Tribunal implied, this factor did not carry much weight in its verdict on fairness.

This approach has much to recommend it. If self-identification suffices to engage DPA obligations, then genuine anonymisation would, in a great many cases, be rendered impossible. For example, with statistics involving low cell counts, the data subject would almost always be able to self-identify.

On the basis of the distinction developed by Hope LJ, IP addresses and event data would, however, be personal data in the hands of the copyright owner where other information enabling the identification of the data subject is likely to come into the owner's possession. This raises the question of the required standard of likelihood. In an opinion on the concept of personal data, the EU's Article 29 Working Party's answer appears to be that unless the risk of identification is non-existent or negligible, this threshold will be met (see *Article 29 Working Party: Opinion 4/2007 on the concept of personal data* and *Legal update, EC Working Party issues opinion on concept of personal data (www.practicallaw.com/4-370-0030)*).

However, one could argue that a UK court is unlikely to put the matter quite so high, in the light of the ordinary meaning of likely. The Irish High Court has taken likely in this context to mean probable (*EMI Records (Ireland) Ltd and others v Eircom Ltd [2010] IEHC 108*). As a matter of language and common sense, this interpretation commends itself.

Admittedly, Recital 26 of Data Protection Directive states that "to determine whether a person is identifiable [from certain data] account should be taken of all the means likely reasonably to be used by either the controller or by any other person to identify the said person". Here, it is at least "reasonably likely" that some "other person", namely the ISP, will use the IP address passed to it by the owner to identify the subscriber. This, however, is merely something of which "account should be taken".

In any event, the DPA takes the data controller, rather than "any other person" as its reference point, that is, the definition turns on whether additional information is likely to come into the possession of the data controller. This means that, in the context of the imminent DEA 2010 regime under which most processing is done without personal details passing to the copyright owner, there is scope for concluding that IP addresses and event data are not always personal data in the hands of the owner.

The probability of an owner obtaining personal details will need to be considered on a case-by-case basis, by reference to the owner's apparent intentions. For example, where an owner has habitually sought Norwich Pharmacal orders in the past, one could reasonably assume a similar course of action in future to be probable. The opposite conclusion could reasonably be drawn where the owner has stated in its correspondence with the ISP (for example, when submitting a copyright infringement report) that it does not intend to take legal action against subscribers.

In summary, it is likely that copyright enforcement will, in almost all cases, engage the ISP's duties under the DPA. However, owners may only be so bound where they intend to sue subscribers suspected of copyright infringement. If the owner chooses only to report apparent infringements and then leave the matter in the hands of the ISP (to issue warning letters under the present approach, or potentially to take "three strikes

followed by technical measures" action in future), then its activity falls outside the DPA. So too will the owner's passing observation of the innocent online activity of bystanders.

These conclusions are, of course, far from settled. At first glance, they may seem alarmingly permissive. It should be remembered, however, that the detection exercise undertaken by the copyright owner in its hunt for online infringement is one that can be carried out by anyone, with the aid of readily available software. In this light, the reach of the DPA and its duties would be practically limitless if IP addresses and event data, without additional information identifying the data subject, were to constitute personal data.

Copyright enforcement and compliance with the DPA

Having suggested the boundaries within which duties under the DPA are engaged, it remains to consider whether, in taking the kinds of copyright enforcement action described in this article, owners and ISPs are complying with those duties, and in particular the data protection principles under Schedule 1 of the DPA. A number of these principles will be relevant, but easily complied with under the auspices of the DEA 2010 and the draft Initial Obligations Code. For example, personal data must only be obtained for and processed compatibly with specified lawful purposes (the second principle), and must be kept for longer than is necessary for those purposes (the fifth principle).

As is so often the case, however, the first principle will be decisive. This provides that "personal data shall be processed fairly and lawfully and, in particular, shall not be processed unless at least one of the conditions in Schedule 2 is met". Lawfulness will again be supplied by the DEA 2010. The condition relied upon from Schedule 2 will almost certainly be that at paragraph 6(1), which applies where:

"The processing is necessary for the purposes of legitimate interests pursued by the data controller or by the third party or parties to whom the data are disclosed, except where the processing is unwarranted in any particular case by reason of prejudice to the rights and freedoms or legitimate interests of the data subject."
The processing with which we are concerned here is plainly reasonably necessary (the established standard) for the purposes of the owner's **legitimate** interest in enforcing its copyright over illegally downloaded content. A subscriber who has in fact infringed copyright is unlikely to have a right, freedom or legitimate interest in concealing that fact. In contrast, a suspect who is not in fact guilty of wrongdoing will, like a bystander, have a right, freedom or legitimate interest in the privacy of his lawful online activity.
The outstanding questions are closely entwined: is the processing carried out by the owner and the ISP **fair**, and is any resultant prejudice to the innocent suspect's privacy **unwarranted**?

The factors tending to suggest an answer of no to both questions are those articulated by the EDPS and summarised earlier in this article. These are serious and complex. See, for example, the concerns noted by Birss J in his decision of 8 February 2011 in *Media CAT Ltd v Adams and others [2011] EWPCC 6*, in particular concerning the difficulties of linking an IP address to the individual responsible for the file-sharing, and of establishing that copyright has been infringed. Together, these problems are likely to mean that the past approach (whereby owners pursued subscribers directly) would not comply with the DPA.
In contrast the imminent DEA 2010 regime is likely to mean that processing is generally fair and does not result in unwarranted interferences. This incorporates safeguards that aim to minimise the prejudice to affected subscribers.

For example, ISPs are required to ensure that:

- Allegations made by owners are supported by credible evidence gathered in a robust manner (*paragraph 1.3 and section 5, draft Initial Obligations Code*).

- A detailed (and anonymous) appeal process will be prescribed (*section 7, draft Initial Obligations Code*).

- Legal action will be taken only against repeated (and therefore generally more serious) infringers.

This regime appears to be a proportionate means of achieving the aim of protecting copyright and other intellectual property rights. It should be remembered that EU legislation endorses these aims. For example, Council Directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society (Copyright Directive) states that "member states shall ensure that rights-holders are in a position to apply for an injunction against intermediaries whose services are used by a third party to infringe copyright or a related right" (*Article 8.3*).

EU Court of Justice (ECJ) jurisprudence, for example in the case of *Productores de Musica de Espana (Promusicae) v Telefonica de Espana SAU (Case C-275/06) [2008] ECR I-00271*, also acknowledges rights-holders' interests and the fact that they are legally entitled to adequate remedies for the breaches they suffer. Although that decision did not result in ISPs being obliged to provide owners with details of suspected file-sharers, it does illustrate the importance the ECJ places on the objective of protecting copyright.
If the "three strikes followed by technical measures" approach were to be implemented in future, the outcome would be more finely balanced (given the penalty of restricted internet access), but the ultimate outcome would be likely to be unchanged. In many cases, the subscriber's unlawful online activity will place

him in breach of his contract with his ISP. In any event, a subscriber would be free to enter into a new contract with an alternative ISP.

Consequently, UK law as it currently stands leans more towards the Irish approach than that of the EDPS on the question of how data protection accommodates copyright protection. The last word on this may well be spoken by the ECJ in the future, but that is unlikely to be the case for some time. Until then, uncertainty persists. In the interim, this means there is ample scope for the administrative court to shape the domestic data protection landscape as it sees fit.

Robin Hopkins is a barrister at 11 KWB, Temple, London.