Client Alert

12 March 2015

Data for the Taking: Using Website Terms and Conditions to Combat Web Scraping

By Susan McLean and Mercedes Samavi

Is it stealing to take data without permission from a public website, or is it simply making use of resources that are made available to you? "Web scraping" or "screen scraping" is the practice of extracting large amounts of data from public websites using bots.

A recent case in the European Court of Justice has focused attention both on the intellectual property infringement aspects of scraping practices and on the potential for website owners to use their sites' contractual terms and conditions to combat the scrapers.

Scraping is not new, but it has become increasingly widespread in recent years, fuelled by the rise in big data analytics and the popularity of price comparison websites. Indeed, in 2013, it accounted for 18% of site visitors and 23% of all Internet traffic. Scraping is not inherently bad: it can have legitimate uses, spur innovation and give companies with limited resources access to large amounts of data. However, unsurprisingly, many website operators do not like it. Not only are operators keen to protect their proprietary rights, but repeated scraping can also take a heavy toll on websites by using up bandwidth and leading to network crashes.

In the U.S., website operators have asserted various claims against scrapers, including copyright claims, trespass to chattels claims and contract-based claims alleging that scrapers violated their websites terms of use. In the EU, operators have tended to rely on intellectual property infringement claims against scrapers, but there has been little case law to provide guidance.

However, in January 2015, in a much anticipated decision, the European Court of Justice (CJEU) held that where a website operator cannot establish intellectual property rights in its database, an operator may still be able to rely on its website terms and conditions to prohibit scraping. This ruling may impact an increasing number of companies whose business models rely on mining data from websites and social media platforms without permission. On the other hand, it will be viewed positively by those data-rich businesses keen to protect and/or monetise their data.

RYANAIR LTD V PR AVIATION

The CJEU case involved PR Aviation which operates a price-comparison website for low-cost airlines. Consumers can book a flight on the website and PR Aviation receives a commission. The website relies on information obtained by screen scraping publicly available data from the websites of low cost airlines, including data from Ryanair's website.

Ryanair sued the defendant for infringement of database rights under the Database Directive (96/9/EC), and breach of its website terms and conditions. It sought an order against PR Aviation to refrain from any further infringement on pain of a financial penalty and for PR Aviation to pay damages.

WHAT ARE DATABASE RIGHTS?

Database rights are a form of unregistered intellectual property rights introduced by the Database Directive in 1996 and implemented into national law across the European Union.

The aim of the Database Directive was to harmonise the rules that applied to copyright protection of databases across the EU, safeguard the investment of database makers and secure the legitimate interests of database users. In essence, the Directive sought to create a legal framework appropriate to the use of databases in the information age. It did so by ensuring copyright protection to those elements of databases possessing protectable expression and introducing a new form of "*sui generis*" protection to those elements of databases which are not "original" in the sense of being the author's own intellectual creation.

Accordingly, the Database Directive provides two forms of protection. Article 3(1) establishes the first of these rights: "*databases which by reason of the selection or arrangement of their contents, constitute the author's own intellectual creation shall be protected as such by copyright*". The second form of protection (established by Article 7) provides protection where there "*has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents [of a database]*".

Importantly, the Database Directive includes certain limited exceptions to the rights created. In particular, Article 6 allows lawful users to make a copy of a copyright-protected database without consent where it is necessary to do so in order to access its contents. Further, Article 8 permits lawful users of a publicly available database to extract and/or reuse insubstantial parts of its contents, as long as this use does not conflict with normal exploitation of the database or unreasonably prejudice the legitimate interests of the database's author.

DUTCH SUPREME COURT

The Ryanair dispute ended up in the Dutch Supreme Court where PR Aviation successfully argued that Ryanair could not rely on copyright protection because Ryanair's database was not sufficiently original to attract copyright protection, and also that there had been insufficient investment by Ryanair, in compiling its database, for it to claim the *sui generis* right.

However, the court still faced the question whether Ryanair could assert a claim that PR Aviation had breached Ryanair's website terms and conditions by scraping and re-using data from the Ryanair site. Significantly, Ryanair's website terms and conditions contain the following express prohibition on the use of screen scraping: "*The use of automated systems or software to extract data from this website or www.bookryanair.com for commercial purposes ('screen scraping') is prohibited unless the third party has directly concluded a written licence agreement with Ryanair which permits access to Ryanair's price, flight and timetable information for the sole purpose of price comparison.*"

Ryanair sought to enforce this term. PR Aviation argued that the prohibition against screen scraping was not enforceable because, under Article 15 of the Database Directive, any contractual provisions which are contrary to

Articles 6 and 8 are rendered null and void. The Dutch Supreme Court was unsure whether Article 15 of the Directive applied to a database which did not attract copyright protection or the *sui generis* right and therefore it sought a preliminary ruling from the CJEU.

CJEU DECISION

In a logical ruling, the CJEU ruled that the limitations on rights introduced by the Database Directive do not apply to databases that are not protected by the Directive. Accordingly, Articles 6, 8 and 15 of the Database Directive do not preclude a website operator from laying down contractual limits on the use of a database, without prejudice to applicable national law.

The case has now been sent back to the Dutch courts, which must decide on the enforceability of the Ryanair website terms and conditions.

For a website owner, it is not simply a question of prohibiting scraping in its terms and conditions; an operator also needs to ensure that those terms and conditions are enforceable. We have written before about the issues involved, particularly in Europe, in ensuring that online terms are deemed fair and reasonable.

Ideally, a website operator would require any user of its site to accept the website terms and conditions before allowing the user to access the website. However, the majority of sites are reluctant to enforce this rule because it is not considered user-friendly. It is therefore more common to ensure that a link to the terms is displayed prominently on the site. The problem with this method is that there is no active acceptance of the terms (e.g., clicking a box). As a result, there is a risk that the website owner will be unable to demonstrate that there is a contract in place with the user. This is a question for national law, as indicated in this case.

There is no binding case law on the issue in the UK. Unfortunately, although the issue was touched on in the recent high-profile *Newspaper Licensing Agency v Meltwater [2011] EWCA Civ 890*, the Court of Appeal did not consider whether an end-user was bound by the website terms of use because, given the nature of the case, it said that it was unnecessary "to enter into that controversy".

Lastly, the *Ryanair* judgment does not answer the question of whether a screenscraper would ever be able to rely on the lawful use exceptions set out in Article 6 or 8 of the Database Directive if the database owner were able to establish copyright protection or the *sui generis* right in the database.

OTHER CLAIMS

It is worth pointing out that, in addition to intellectual property rights infringement and contract breach claims, website owners may have other legal arguments against scraping. For example, as in the U.S., in the UK, a website operator may try to bring a claim for trespass to chattels, a common law tort. In addition, an operator may seek to rely on the Computer Misuse Act 1990 which prohibits unauthorised access to, or modification of, computer material. To date, as with database rights, neither of these arguments have been tested in the UK courts in connection with web scraping. (However, similar legislation has formed the basis of claims elsewhere, for example, in the U.S., as described in our previous alert, "Data for the Taking: Using the Computer Fraud and Abuse Act to Combat Web Scraping".)

Of course, when dealing with scraped data, issues of privacy and security loom large and web scrapers and users of scraped data will also need to tread extremely carefully in order to avoid problems under applicable privacy laws. For a detailed discussion on privacy and big data, see our previous Alert.

CONCLUSION

The CJEU's *Ryanair* decision appears to give a rather contrary result – in certain circumstances, a database owner may have broader, albeit contractual, rights to prevent scraping if it does not actually have proprietary rights in the database. However, in any event, in light of this decision, website owners based in the EU may be encouraged to amend their website terms and conditions to expressly prohibit screen-scraping in order to try to protect their valuable data.

Whether the impact of this decision will truly disrupt those companies with business models that rely on the use of data mined from websites and social media platforms remains to be seen. Certainly, any business that carries out screen-scraping activities should consider where it sources its data from and identify whether such data are bound by contractual limitations or other restrictions. It can then make a reasoned decision on whether or not it should approach the database owner for a commercial licence to ensure that the data keeps flowing.

Contact:

Susan McLean
+44 (0)20 7920 4045
smclean@mofo.com

Mercedes Samavi
+44 (0)20 7920 4170
msamavi@mofo.com

About Morrison & Foerster:

We are Morrison & Foerster—a global firm of exceptional credentials. Our clients include some of the largest financial institutions, investment banks, Fortune 100, technology and life science companies. We've been included on *The American Lawyer*'s A-List for 11 straight years, and *Fortune* named us one of the "100 Best Companies to Work For." Our lawyers are committed to achieving innovative and business-minded results for our clients, while preserving the differences that make us stronger. This is MoFo. Visit us at www.mofo.com.

Because of the generality of this update, the information provided herein may not be applicable in all situations and should not be acted upon without specific legal advice based on particular situations. Prior results do not guarantee a similar outcome.