

# How Big Data and IP Intersect

## Big Data is big business—but who owns it?

BY PAVEN MALHOTRA

**A**ccording to IBM, enterprises and individuals create 2.5 quintillion bytes of data every day. The explosive growth in data means that an estimated 90 percent of all data that exists today was created in just the last two years. The advent of the internet and now the Internet of Things—the integration of connected computers into all manners of devices, from automobiles to refrigerators to hospital equipment—has left companies sitting on vast troves of information concerning how their products and services are being used.

Big Data—as this information is often called—has become big business. International Data Corp. estimates that by 2019, spending on big data technology and services may reach nearly \$49 billion. Of course, the vast majority of information that businesses collect is worthless and ultimately discarded. But for the data that does yield useful insights, executives have been looking to extract as much commercial value as possible. For corporate counsel, Big Data raises other concerns: What legal tools are available to protect and regulate the data collected?

Although academics and policy-makers have proposed new forms of intellectual property rights to protect big data, those efforts have stalled in legislatures, and companies must look to more pedestrian legal avenues to protect their



informational assets. Each has its limitations, but copyright, trade secret and contract law offers the most promising avenues for corporate counsel to consider.

## Copyright

Copyrights attach to creative works that are fixed in a tangible medium. Traditionally, copyright has been used to protect writings, musical compositions and artistic pieces. Copyright holders are entitled to prohibit others from reproducing, distributing or selling the copyrighted work as well and to prohibit others from preparing derivative works based on the copyrighted work.

Important in the Big Data context, copyrights are also available for data compilations. Copyright law defines a compilation as “a work formed by the collection and assembling of pre-existing materials or of data that are selected, coordinated or arranged in such a way that the resulting work as a whole constitutes an original work of authorship.” Critically, the underlying data in a compilation is not copyrightable—only the compilation itself. Additionally, not every compilation is entitled to copyright protection. The Supreme Court in *Feist Publications v. Rural Telephone Services* held that the compilation must reflect some originality,

# Intellectual PROPERTY

## COLUMN

although the amount of originality can be minimal. A listing of data chronologically, sequentially or alphabetically would likely not suffice, but another grouping that reflected the exercise of subjective judgment would.

When initially collected, Big Data is often raw and unstructured.

### **The compilation of data into a format that reflects the judgment and efforts of a corporation might be copyrightable.**

As a result, much of the spending on Big Data is for software and for personnel, who then organize the data into something more useful. Under the framework of copyright, the compilation of data into a format that reflects the judgment and efforts of a corporation may be copyrightable. Importantly, though, the individual pieces of data that form the compilation are not—a significant shortcoming of copyright law. As the U.S. Supreme Court noted in *Feist Publications*, “the raw facts may be copied at will.”

#### **Trade Secrets**

While copyright is somewhat limited in the protections it offers the individual data that goes into a compilation, trade secret law offers more robust protections. All the states and, as of May 2016, the federal government have enacted trade secret acts. Although the details differ, many states and the federal government define trade secrets as information and compilations of information that are not generally known, that confer a competitive advantage, and that have been the subject of efforts to maintain their

confidentiality. Trade secret laws prohibit the misappropriation of trade secrets—meaning the acquisition of information through improper means or through a breach have a duty to keep it confidential. Notably, acquisition of a trade secret through reverse engineering is not prohibited.

Because trade secret laws protect not only the compilation of data but also the underlying data itself, it offers companies a potent tool. To avail themselves of these protections, however, companies must put in place procedures to ensure that the data itself stays confidential. Prudent confidentiality measures include securing the data within the company by erecting IT procedures to guard against a data breach, restricting the data within a company to employees who have a need to know the data, and sharing data with third parties only under a nondisclosure agreement.

#### **Contract Law**

Contractual rights offer another critical form of legal protection for businesses sitting on Big Data. These rights must be addressed on two fronts—contracts with parties that may be sources of data and with parties that may be the users of data.

For many consumer-facing businesses, the largest source of data is often derived from consumer interactions with a product or service. A business worried that consumers

(or consumers in a class action) may claim that they retain ownership in data would be well advised to evaluate their terms of service. The terms may be structured to specify that any data derived from consumer transactions belong to the corporation.

If a company elects to share a data set with a licensee, the license agreement must specify the terms of ownership rights not only over the data itself but also any derivative works derived from the data. Because a licensee may spend substantial time and resources mining a data set to derive critical business insights, any licensor or licensee must address how ownership over derivative works is structured.

#### **Conclusion**

Every company that generates electronic information in any significant way—and that is most businesses today—must confront the challenges and opportunities posed by Big Data. For corporate counsel concerned with protecting their data, trade secret and contract law offer the two most promising avenues of protection. Although copyright law can protect a data compilation, it offers little protection for the individual data.

---

*Paven Malhotra is a partner at Keker & Van Nest in San Francisco, where his practice focuses on intellectual property disputes and complex commercial litigation matters.*

---

Reprinted with permission from the Fall 2016 edition of INTELLECTUAL PROPERTY an ALM Supplement to CORPORATE COUNSEL © 2016 ALM Media Properties, LLC. All rights reserved. Further duplication without permission is prohibited. For information, contact 877-257-3382 or reprints@alm.com. # 016-08-16-05