WINTER 2014

New Matter

Official Publication of the Intellectual Property Law Section of the State Bar of California

Volume 39, Number 4

Contents

Thomas A. Ward
Editor-in-Chief
Arris
6450 Sequence Drive
San Diego, CA 92121
tel: (858) 404-3542
tom.ward@arrisi.com

Anthony G. Craig, Ph.D.
Production Editor
Sci-Law Strategies, PC
P.O. Box 1729
Solana Beach, CA 92075
tel: (858) 848-6115
agc@sci-law.com

Amanda Nye, Acquisitions Editor
Lawyers At Large LLC
6 Davis Drive, #E
Tiburon, CA 94920
tel: (415) 937-5244
anye@lawyersatlargellc.com

Robert Dabney Eastham
Senior Articles Editor
Think Tank Photo, Inc.
1105 N. Dutton Avenue, Suite C
Santa Rosa, California 95401
direct: (443) 867-2827
office: (866) 558-4465 ext. 204
dabneyeastham@yahoo.com

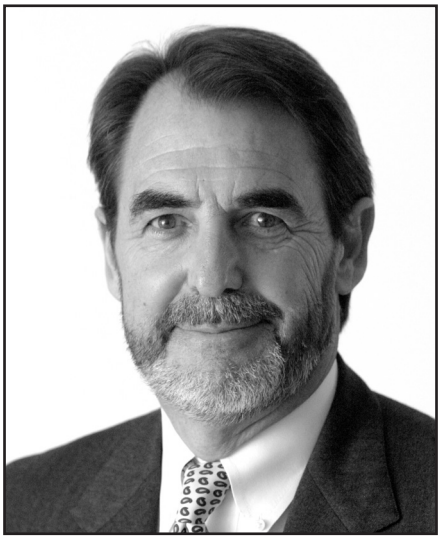
Articles Editor:

Yumi K. Nam
Idell & Seitel LLP

Alutha Jamancar
Publication Coordinator
tel: (408) 507-3770
alutha@gmail.com

For replacement copies:
Raven Ogden
Intellectual Property Section
State Bar of California
tel: (415) 538-2091
raven.ogden@calbar.ca.gov

The Licensing Corner



James C. Roberts III
Global Capital Law Group PC

FARMERS, GOOGLE DRIVE AND NEW JARGON

Q: So what do farmers and Google Drive users have in common?

A: In this column, each of them provides data access to another party in exchange for explicit (the farmers) or implicit (Google Drive users) benefits while also granting those other parties very valuable rights to create separate, and massive, datastreams with substantial commercial value.

* * *

Let's start with a hypothetical. Your client is one of hundreds of customers of an online service provider. Like all the other customers, your client uploads its databases for whatever reasons (storage, access, archiving, data manipulation). Now, assume that the provider uses pattern recognition software and other data mining techniques to look at all of those databases (including your client's) in an aggregate way. That analysis yields interesting results that are themselves data about multiple databases—data that can be made available to third parties as datastreams.

Is your client OK with that? In fact, who owns—and who should own—those datastreams (or at least who has rights in them)?

FARMERS AND GOOGLE DRIVE USERS

These questions came to my mind in two ways. First, I recently heard about a multinational corporation launching a program to provide US farmers with substantial technical assistance if the farmers uploaded certain of their data to a server managed by that company. Second, I had to review the TOS for Google Drive.

Don't worry that this column is a "Chicken Little" piece about a falling sky. There is nothing necessarily unsavory in these two examples. They simply give me an opportunity to explore what is a burgeoning business: Datastreams derived from thousands of databases at a time, well beyond our current understanding of, for example, data mining. These are data well beyond the concept of metadata: They are data analyzing the performance of and patterns in those thousands of datastreams provided by the farmers, including the metadata. This "stuff" is no new it does not even have a name (as to which, more later).

In the farmer database example, each farmer collects and uploads data as he or she farms, using GPS-based inputs to provide everything from yield per square meter, to irrigation per square meter, number of seeds and so forth. The chemical company can provide technical assistance based on that information, whether it is adjustments to fertilizer amounts, irrigation or other variables. In turn, the company can use the data to alter its own products—from the seeds to the fertilizer to pricing.

In Google's case, their creation of this "secondary" datastream (pattern recognition from multiple databases) is not explicit; I have tried to infer it from the TOS. While I don't know much about how Google works, I do know that they provide access to certain datastreams. Thus, it struck me as a model that is similar to the farmer model above, but without explicit benefits that the data providers (farmers) receive by participating in the program.

THE GOOGLE TOS

Google's TOS first acknowledges the user's ownership of the data (called content in the TOS) uploaded by the user:

> Some of our Services allow you to submit content. You retain ownership of any intellectual property rights that you hold in that content. In short, what belongs to you stays yours.

But the next paragraph takes it a step further. There, you grant Google and others a license to use your data in multiple ways:

> When you upload or otherwise submit content to our Services, you give Google (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display and distribute such content. *The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones.* [emphasis added]

Putting aside any ambiguities in that license grant, let's look at the "limited purpose" of that grant. Google could improve its "Services" or develop new ones based upon the datastreams it develops from looking at all the data (content) uploaded to Google Drive. No one is suggesting that Google violates any laws in doing so; it is, simply, that the license grant in the TOS could enable Google to mine those databases in Google Drive—within the confines of the law—and derive datastreams about those data.

Before going further, let's first look at categories of data.

DATA, METADATA, DATA MINING & DMO

Your Own Data. What is data? (I am "Old School" in viewing "data" as a plural word.) While computer scientists will argue that all content is data, for our purposes, data are information you would typically find in databases such as customer lists, price lists, website or app metrics and so forth.

In principle, your client owns the data (assuming legal compliance in the data collection). For example, your client would own a database of the metrics of its website use. If your client uploaded that database to the cloud, your client would still own the database, although cloud providers usually receive a license to make one or more copies (*e.g.,* for backup, security and load-balancing).

Metadata. Then there are metadata. If the on-going revelations about the NSA collection policies have taught us anything, it is the definition of "metadata." Put simply, it is "data about data," such as the file type and size, its path from sender to recipient, date and time of creation, keywords (metatags) and so forth.

It is not exactly clear who owns the metadata to your client's data. If the principle is that the person who creates IP owns it (with obvious exceptions), then ownership could depend on how the data were created. Authorship, length and other attributes are probably created by your client through its software. It gets less clear when the metadata are created by the path through telecommunications networks.

Data Mining. Wikipedia defines data mining as:

> […] the computational process of discovering patterns in large data sets […]. The overall goal of the data mining process is to *extract information from a data set and transform it into an understandable structure for further use.* […] [emphasis added]

For example, your client creates and uploads a massive database about website use that incudes user demographics, geographic location, time of day and type of transaction. (We are assuming all such data are non-PII.) Your client hires a data mining company that produces results that tell your client that, say, 14.7% of the users are 35–38 year-old females with college degrees who read the article in the center of the page between 8 and 9 PM PST in certain zip codes.

Data Mining Output. However, data mining is not data: Data mining *produces* data. If we follow the logic of the definition of metadata above, then these results are "data about datasets." It develops patterns from your client's data, rather than just metadata about those data. Looking at the sentence above in bold, the results are the information extracted from the dataset (or perhaps about the dataset) transformed into another use and another database.

Consider the example above about females with college degrees. Those are new data. So what if this data mining company offers the same services to your competitors and aggregates all of the *results* (not the actual data in those databases) from all of the databases it has mined? Who owns those data? (OK, I know that some of you would argue that it depends on what the contract says and some of you would argue that the person that reduced the results to a tangible medium would own them.)

We do not yet even have a name for the output, or results, of data mining—the results of mining more than one dataset. It certainly is not "metadata" in the commonly accepted definition because it is not about the origin, path, etc., of those data but data derived from the data in the original dataset when certain queries are posed. So, for our purposes, let's call it DMO for "Data Mining Output."

DMO are the data about which we need to ask questions about ownership: Who owns DMO and who has the right to transform DMO into another use? Presumably, your client signed an agreement with the data mining provider specifying that your client owned all the DMO, not the provider. If not, then it is entirely possible that the provider can take the DMO that *it* has generated (about your client's data in its databases) and use it elsewhere.

AGGREGATING DATABASES AND DATA MINING

What happens when you multiply this scenario times the number of clients of an online service that explicitly or implicitly conducts data mining? In other words, what happens when data mining takes place across hundreds, thousands or millions of databases? (Google was reported to have 120 million Google Drive users in November 2013.) And what if many of those database providers (who are providing the grist for the data mining mill) are competitors?

First, perhaps we have to create another name: in this case "Aggregated Data Mining Output" or ADMO. Second, it is almost axiomatic that a larger population (in this case lots and lots of databases) generally yields more accurate results, which is a good thing.

Well, for whom is it a good thing? It *could* be a good thing for your client if your client were granted access to those data—what we have called ADMO; it could be a *great* thing if the ADMO compared results across your client's industry.

But who has access to those datastreams aggregated from thousands or millions of databases that have been mined? Where is the grant of the right to access an analysis that exists, in no small part, because your client provided data? In the case of the farmer model, I don't know, because I have not seen the terms of the deal. It looks like the farmers do get a good deal, exchanging their collection and uploading their data in exchange for technical assistance. In the Google

Drive case, I see no such access granted. Google Drive users, too, get a benefit, at least implicitly: Google's rights in your data are to be used to improve their services or develop new ones.

SO WHAT?

So, licensing lawyers think about subsequent uses of data—and access to the results of those uses, both the DMO and the ADMO. What does "think about" it mean? Certainly, one should negotiate for ownership of all DMO from a provider, denying use of the DMO by the provider in subsequent engagements. Or, perhaps the provider can use such data *if* you are provided access to DMO from their later engagements with other customers or on an aggregated basis. As for ADMO, it would be a longshot when market power is essentially asymmetrical (as it would probably be with Google). But, in other circumstances, it's worth a shot. ◂◂

The views expressed in this article are personal to the author and do not necessarily reflect the views of the author's firm, the State Bar of California, or any colleagues, organization, or client.

James C. Roberts III is the Chair of Licensing Interest Group. He manages Global Capital Law Group and serves as the CEO of its consulting group, Global Capital Strategies. He advises venture capital funds, startups and corporations on domestic and international transactions, including mergers & acquisitions, startup funding and international distribution. You can reach him at jcr@globalcaplaw.com.

EDITOR'S LETTER
continued from page 4

the column are as follows: The Ninth Circuit Report by Anne Marie Dao, International IP Developments by Aurelia Shultz, The Law of the Land by Ben Borson and Derrick Brent (which will continue in our next issue of *New Matter*), The Licensing Corner by James Roberts III, Case Comments by Lowell Anderson, and now with this issue we begin the Federal Circuit Report by Glenn Von Tersch. For those interested in contributing a regular column in the future, please contact our Production Editor, Anthony Craig who oversees approval of the columns and assists the managers with editing and printing their columns in *New Matter*.

As always, thank you to our entire *New Matter* editing staff and the authors. Also, thank you to members of the IP Executive committee and State Bar staff who provided guidance and support for this issue. On behalf of all the *New Matter* staff, we hope you are enlightened and informed by this issue of *New Matter*. ◂◂

INTERNAT'L IP DEVELOPMENTS
continued from page 25

31. Jones Day, "Amendment to Trademark Act of Japan—New Marks, Including Color Marks and Sound Marks, Can Be Protected," September 2014, available at http://www.jonesday.com/amendment-to-trademark-act-of-japan---new-marks-including-color-marks-and-sound-marks-can-be-protected-09-04-2014/.
32. Aoshima, Emi, "Revision of the Japanese Trademark Act and Design Act in 2014," *Yuasa and Hara Intellectual Property News* Vol. 40 June 2014, available at http://www.yuasa-hara.co.jp/english/news/pdf/ipnews040.pdf.
33. The Copyright and Rights in Performances (Quotation and Parody) Regulations 2014, Draft Legislation, available at http://www.legislation.gov.uk/ukdsi/2014/9780111116029.
34. Johan Deckmyn and *Vrijheidsfonds VZW v Helen Vandersteen* and others, C-201/13, court documents available at http://curia.europa.eu/juris/liste.jsf?language=en&td=ALL&num=C-201/13.
35. C-201/13 at paragraph 33. *See also*, Laurent, Philippe, "The concept of parody and the legitimate interests of parodists and copyright holders," *Kluwer Copyright Blog*, September 8, 2014, available at http://kluwercopyrightblog.com/2014/09/08/the-concept-of-parody-and-the-legitimate-interests-of-parodists-and-copyright-holders/.
36. Coulson Harney, "The Re-Introduction of Capital Gains Tax in Kenya," *CH Newsletter*, available at http://communications.bgafricagroup.com/public/messages/view-online/tlLN0skdL-SipWt20/NGGl6oH0SG9RllHN.
37. http://www.eregistration.copyright.gov.ng/ last visited October 6, 2014.
38. For an in-depth comparison and walk through of the system, see Schultz, Aurelia J., "Nigeria Welcomes Online Copyright Registration," *Afro-IP*, September 4, 2014, available at http://afro-ip.blogspot.com/2014/09/nigeria-welcomes-online-copyright.html.
39. Orphan Works Guides are available at https://www.gov.uk/government/publications/orphan-works-diligent-search-guidance-for-applicants.
40. Copyright Exceptions Guides are available at https://www.gov.uk/government/publications/changes-to-copyright-law.