Oracle copied Amazon's API—was that copyright infringement?

Enlarge / An Oracle building in the San Francisco Bay Area.
Smith Collection/Gado/Getty Images

reader comments

177 with 63 posters participating

Share this story

Charles Duan is the Director of Technology and Innovation Policy at the R Street Institute, a nonprofit think tank based in Washington, DC . He has authored several amicus curiae briefs in the litigation between Oracle and Google, as well as the article Internet of Infringing Things: The Effect of Computer Interfaces on Technology Standards. The R Street Institute has received financial support from Google. The opinions expressed here do not necessarily represent those of Ars Technica.

Early this year, the Supreme Court will hear an important case that will determine the legal status of application programming interfaces under copyright law. If the high court sides with Oracle in its multibillion-dollar lawsuit against Google's Android platform, it could stifle competition and entrench dominant technology firms—possibly including Google itself.

Oracle has accused Google of infringing copyright law by copying the API of the Java programming language. An API is essentially a language for instructing a computer on what to do. It includes a vocabulary of named commands tied to

grammatical structures for how those commands are to be used. To cause Java software to perform pre-defined tasks, such as calculating a sine function or encrypting a message, a programmer must use those named commands and grammatical structures with precision, much in the same way that a Waffle House diner invokes exact code words like “scattered, smothered, chunked, and diced” to get a hash brown order correct.

Google’s Android was designed to be compatible with the Java API so that programmers already familiar with Java could easily bring their software and knowledge over to the new mobile device platform. To do this, Android had to exactly copy the relevant Java API commands and grammatical structures. Oracle’s argument is that this “reimplementation” of the Java API is on par with writing an unauthorized “Harry Potter” novel and thus an infringement of Oracle’s copyrights in the command names and structures of the Java API.

But the Java API is not the only API, and Android is not the only reimplementation. APIs are found all over modern technology, and reimplementation is a key part of ensuring competition in the computer industry and preventing large-firm lock-in.

Did Oracle infringe copyright by copying Amazon’s S3 API?

David Ryder/Getty Images

Consider Amazon's popular data storage platform, S3. To allow programmers to store and retrieve files on S3, Amazon built a comprehensive, detailed API for interacting with the service. To get a list of stored files, for example, one sends the command name *GET* with the folder name as a grammatical object, along with cryptically tagged information such as *encoding-type*, *continuation-token*, and *x-amz-date*. Software must use these exact, cryptic terms and a bevy of others to work with Amazon S3.

Unsurprisingly, competitors have sprung up to Amazon's market-leading cloud services. To convince programmers to switch away from Amazon's offerings, those competitors reimplement S3's API. In doing so, the competitors must mimic the command names, parameter tags, "*x-amz*" phrasing, grammatical structure, and overall organization of the S3 API—in other words, exactly the kind of thing Oracle argues is protected by copyright.

To be sure, a competitor may use a different programming

language than Amazon did so the internal software code might not look like verbatim copying. But implementing an API in another computer language is simply an act of translation, and translating a copyrighted work into another language is specifically known to be a copyright infringement.

Among the companies offering a copy of Amazon's S3 API is Oracle itself. In order to be compatible with S3, Oracle's "Amazon S3 Compatibility API" copies numerous elements of Amazon's API, down to the *x-amz* tags.

Did Oracle infringe Amazon's copyright here? Ars Technica contacted Oracle to ask them if they had a license to copy Amazon's S3 API. An Oracle spokeswoman said that the S3 API was licensed under an Apache 2.0 license. She pointed us to the Amazon SDK for Java, which does indeed come with an Apache 2.0 license.

However, the Amazon SDK is code that uses the S3 API, not code that implements it—the difference between a customer who orders hash browns and the Waffle House cook who interprets the orders. Code that uses an API will be organized completely differently from code that implements one; it may not even contain the whole API. And Oracle has for years argued that using an API is unrelated to reimplementation and not an infringement of copyright (or else every app developer using Java would infringe). Oracle can't simultaneously argue that API-using code does not embody copyrighted material from an API, and yet API-using code embodies all copyrights in the API necessary to give Oracle the right to reimplement S3.

Even if the Apache license does apply, Oracle doesn't appear to comply with the terms of the license. Section 4 of the Apache license requires notices and attribution statements attached to derivative works. Yet I find no attribution to Amazon or relevant mention of an Apache license on Oracle's documentation or anywhere on the site.

API copyrights could create a legal minefield

Oracle is by no means alone in reimplementing an API like Amazon S3. A handful of cloud storage systems do the same. And there are abundant other APIs and reimplementations across the technology industry. Technical standards like Wi-Fi and Internet protocols—the glue that keeps the information economy connected—all involve APIs, and every computer and device that uses any modern communication technology necessarily reimplements them. Oracle's copyright theory could turn almost anything you do with a computer—from reading an online document to watching a video—into a legal minefield.

To avoid these far-reaching consequences, Oracle and the appeals court that endorsed Oracle's arguments have tried to limit copyright infringement to partial reimplementations of an API that are "incompatible" with the original. But partial reimplementations are commonplace and often expected. Even Oracle's S3 compatibility API notes its numerous "differences" and incompatibilities with Amazon's API.

The danger of Oracle's copyright lawsuit is that it could prevent smaller technology companies from building compatible versions of dominant software platforms. Without such compatibility, software programmers familiar with a dominant firm's API will effectively be locked into that firm's offerings. Sure, the competitors can ask for permission to use an API, but giving companies gatekeeper power over competition through copyright law cannot bode well for a robustly competitive software marketplace.

Update: Oracle sent Ars Technica the following statement.

> *The Apache license is a permissive open source license and any use-case-specific argument (e.g., "SDK vs. non-SDK") doesn't change the availability or nature of the license. It*

allows downstream users to copy and modify code, and use the code in their own projects, with few restrictions. That is why Apache-licensed code is widely used in commercial products. AWS understands this and obviously consciously chose to use a more permissive vs. more restrictive license (such as the GPLv.2).

https://arstechnica.com/?p=1635339