API Privacy: A Look at G Suite Marketplace Permissions and Policies

Irwin Reyes    Michael Lack

twosix LABS
The story actually begins with Spotify. A popular streaming service.

At the end of 2019, they released Spotify Wrapped: Statistics about your listening habits over the course of the year.
I thought this was really cool, and wanted more. Found a reddit post about other Spotify stats services. Nice!
Of course I decided to install these services. By “install,” I mean consent to whatever terms and permissions this nag screen showed.
Before long, I realized I just granted a whole bunch of these things programmatic access to my Spotify account via the Spotify application programming interface (API).

But this is small potatoes. Who cares if external services can read my Spotify data? What’s the worst that could happen?
Well, this was the same mechanism that resulted in Mr. Zuckerberg being hauled in front of Congress in 2018.

A researcher had written a quiz app that obtained programmatic access to people’s Facebook data. The scope of that access was far broader than people expected (it collected data from quiz app users *and* their friends), and was later used in targeted political advertising. This was the Cambridge Analytica scandal.

Facebook eventually placed additional restrictions and safeguards on apps using the Facebook API. Limits on what personal data third-party apps can access, and authorizations expiring if the user hasn’t used an app for a while.
This got me thinking: Have other major online services placed similar safeguards on their APIs?
Looked at company Google account. We use Zoom a lot internally, and my account automatically had Zoom integration authorized.

Some interesting permissions.
Google offers a large number of APIs, that give third-party software programmatic access to private user data.

This is just a small sample.
Google uses OAuth to allow users to authorize third-party software to access their private data held on Google servers.

If I want to authorize an app, I have to be logged into Google, find where the app is being offered, then install it onto my account.
The app installation process will notify me of the app’s API permissions and ask for my permission.

This screen is generated by Google, and is standard for all apps that ask for authorization for user data. Same prompt for web apps, mobile apps, Slack addons, etc.
There’s a whole ecosystem of software products that request access to Google user data. Totally legitimate purposes, like email apps, Slack bots, document editors that save to Google Drive, etc.

Got me thinking about research questions though.
There’s a whole ecosystem of software products that request access to Google user data. Totally legitimate purposes, like email apps, Slack bots, document editors that save to Google Drive, etc.

Got me thinking about research questions though.
The G Suite Marketplace is a convenient centralized repository of web apps that integrate with Google APIs.

Mostly business and productivity focused.
Scraped the G Suite marketplace.

Corpus of 987 installable apps. Some apps were only installable on proper G Suite accounts with admins (i.e., organizations)
<table>
<thead>
<tr>
<th>Permission</th>
<th>Frequency (from N=987 apps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display and run third-party web content in prompts and sidebars</td>
<td>50%</td>
</tr>
<tr>
<td>Connect to an external service</td>
<td>49%</td>
</tr>
<tr>
<td>See, edit, create, and delete your spreadsheets in Google Drive</td>
<td>27%</td>
</tr>
<tr>
<td>Allow this application to run when you are not present</td>
<td>25%</td>
</tr>
<tr>
<td>See, edit, create, and delete all of your Google Drive files</td>
<td>21%</td>
</tr>
</tbody>
</table>

Most common permissions requested
Most common permissions requested for apps that can connect to an external service.

481 apps can connect to external services

21% have full access to Google Drive data

17% can read email while running as an add-on

3% have full access to Contacts data
There’s no mandatory disclosure for what external services are used or for what purpose.

Developers can voluntarily disclose this in the app description or the privacy policy, but that’s optional. User has to go out of the way to find out, and not even reliable.
Given these risks, what’s the platform doing?

Some apps threw this warning at installation. Something about app verification, but can still be installed.
Apps require verification if they use “sensitive” or “restricted” scopes for API authorizations. The most stringent review is for the Gmail and Drive scopes.

Apps under review have a limit on new users, but subject to Google’s discretion. What does this discretion look like?
277 unverified apps encountered

144 unverified apps could still be installed

initial data collection on January 2, 2020

re-examined 144 unverified apps 16 days later
<table>
<thead>
<tr>
<th>Unverified App</th>
<th>User growth (Jan. 2 to 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemistry Question Generator</td>
<td>+9,398</td>
</tr>
<tr>
<td>YouCanBook.me</td>
<td>+7,158</td>
</tr>
<tr>
<td>siteMaestro</td>
<td>+3,632</td>
</tr>
<tr>
<td>Hippo Video</td>
<td>+2,332</td>
</tr>
<tr>
<td>ezShared Contacts</td>
<td>+1,135</td>
</tr>
</tbody>
</table>
Close off with thoughts on future research directions
This initial investigation only looked at web apps offered on the G Suite Marketplace.

There are many other classes of software that integrate with these APIs: mobile apps, desktop apps, etc.

Worth it to look at a broader set of software that use these APIs.
I asked a friend to look at his Google account for API authorizations. Found lots.

How prevalent are these apps? What permissions do they use? How long have they been authorized? Are users surprised?
How about controls and disclosure?

We’ve learned a lot about how users navigate permissions on mobile devices. Install-time permissions are out, and run-time permissions are in.

Users can understand context and infer when and why an app needs to exercise a certain privilege. Can we do the same for web API authorizations?
THANKS!

apiprivacy.irwinreyes.com

@irwinreyescom

irwin.reyes@twosixlabs.com