Google Drive security considerations in an academic and research space

August 23, 2021

By [Mark Krenz](mailto:mkrenz@iu.edu) and Ishan Abhinit



**About the authors**

Mark Krenz is chief security analyst at the Center for Applied Cybersecurity Research (CACR). He is focused on cybersecurity operations, research and education. He has more than two decades of experience in system and network administration and has spent the last decade focused on cybersecurity. He serves as the CISO of the ResearchSOC and also the Software Assurance Marketplace (SWAMP).

Ishan Abhinit has been a senior security analyst at the Center for Applied Cybersecurity Research (CACR) for the past two and a half years. He has a master’s degree in Cybersecurity from Northeastern University, Boston. Before moving to the US, he had worked with IBM India Pvt. Ltd and Infosys Ltd in security domains.

**About ResearchSOC**

The research community is large, highly collaborative, uses diverse infrastructure (like telescopes, control networks, sensors, and high performance computing, to name a few), and enjoys a fair degree of autonomy. These factors challenge standard cybersecurity frameworks and require expertise in scientific research and cybersecurity to develop effective cybersecurity programs.

ResearchSOC helps make scientific computing resilient to cyberattacks and capable of supporting trustworthy, productive research. We do this by providing the operational cybersecurity services, training, and information sharing necessary to a community as unique and variable as research and education (R&E).

The Research Security Operations Center (ResearchSOC) is an NSF-funded collaborative security response center that addresses the unique cybersecurity concerns of the research community.

ResearchSOC provides a combination of operational services, including a security operations center, vulnerability identification, decoy computers (honeypots), threat intelligence tailored for research, a community of practice, and virtual security staff.

In addition to over 30 higher education partners, STINGAR has been deployed in 3 labs under the NSF-funded ResearchSOC.

The ResearchSOC is supported by the National Science Foundation under Grant 1840034. The views expressed do not necessarily reflect the views of the National Science Foundation or any other organization. More information at researchsoc.iu.edu.

Cloud storage systems are used by millions worldwide for data storage. There are many cloud storage systems available. Popular options include Google Drive, Google Shared Drive, Microsoft OneDrive, and Box. In this paper we will provide a brief introduction to Google Drive and Google Shared Drive and review the issues of concern to security professionals which we have encountered during use.

The academic and research communities often share and collaborate beyond organizational boundaries, making these security concerns especially relevant. There are many concerns that create significant security risks to the academic and research communities.

This paper is a product of several years using Google Drive within a variety of contexts and projects, most notably ResearchSOC[[1]](#footnote-0), Trusted CI[[2]](#footnote-1), SGCI[[3]](#footnote-2) and SWAMP[[4]](#footnote-3). These projects often involved collaborating on documents with personnel from multiple institutions, a common use case in the academic sector. The authors of this paper are or have been involved with these projects. During use, we observed several security risks and developed methods for mitigating these risks.

# Introduction to Google Drive

Google Drive, including Google Shared Drive and GSuite for Enterprise / Education, is a cloud application that can be accessed via <https://drive.google.com/>. The service allows users to create, upload, download and organize folders and documents. It is also tightly integrated into the Google Docs suite of tools, which includes Google Docs, Google Sheets, Google Slides, Google Drawings, Google Forms, and other associated web applications.

When accessing Google Drive through a web browser, the user is presented with a file system access interface similar to that which would be seen on an operating system's file system browser.

## Object storage

Google Drive allows multiple users to collaborate on documents stored in its system and the ability to control access to those documents.

Google Drive looks and in many ways behaves like a filesystem hierarchy on an operating system such as Windows or Unix, however it has some unique characteristics and is classified more as an object storage system. Instead of referring to files and directories, documentation usually refers to documents and folders, with folders being just documents of a certain type. One important difference is that a document can exist in multiple folders simultaneously.

## Conventions of this paper

Unless otherwise stated, the term 'document' may refer to a document such as a Google Docs, Sheets, Forms, Slides, or an uploaded file. It can also refer to a folder on Google Drive.

The term 'context menu' refers to the menu that typically appears when the mouse button or other pointing device with default settings has it's right mouse button pressed. This could also be the left mouse button for left handed users.

When referring to a menu item within the Google Drive interface, the right arrow symbol '➞' will be used to indicate the separation between the sequence of menu items that should be clicked on in the order specified.

## Historical changes to Google Drive

Due to the live nature of cloud applications, it is not possible for the general public to reference previous versions of the Google Drive application to check features. There have been changes to what is allowed within the object storage system. Google also periodically makes changes to their user interface and the wording of features, which may affect the accuracy of our descriptions in relation to the current version.

Previously, a user could link a folder to their top level "My Drive" folder and it would appear as if it was any other folder that had been created. Google later changed this behavior so that linking a folder to "My Drive" or any other folder creates a shortcut. This behavior is analogous to the difference between a hard link and a symbolic link on a Unix operating system. However, files created prior to the switch may have retained their format as an additional location for the file instead of being a shortcut. One way to determine that a file that is not a shortcut has been placed in multiple folders is to select the file and select details from it's context menu. If the file is in multiple folders then each folder will be listed in the value for 'Location'.

Google previously limited Google Drive folder downloads to 2GB of documents at a time. This created issues when trying to backup large folders. Google has since increased this limit to an undefined higher value that fluctuates depending on the current load on Google's services.[[5]](#footnote-4)

In September of 2021 Google will make a change to the way documents are shared by link. By default, documents that are shared by link will require that the viewer make a request to view the document. This behavior will be controllable by the owner of the document, but existing documents will have this security feature enabled.

## Google Drive service overview

The Google Drive cloud service can be subcategorized into three categories that allow for different patterns of access and control depending on the use case. With identities with a gmail.com email account, your service category is always normal Google Drive. For identities using an email from a different domain, the category of service that one receives when they login to Google Drive is determined by whether the organization that manages that domain has registered for enterprise level service.

### Standard Google Drive

This is the category of service that you get for free when you sign up for a Gmail account or for a Google Drive account using a domain that has not registered for enterprise service. The category has a basic amount of controls and the lowest class of storage quota. Currently each user is provided with a 15GB storage quota.

Documents and folders can have their sharing permissions changed depending on the use case. By default a document is restricted to just be viewable by the owner, with the exception of any terms and conditions provided by Google in their policies for their service. This class of permissions is called 'Restricted'. If the user wishes, they may share a document directly with other users by going to File ➞ Share and entering their email address in the 'Share with people and groups' section. For each user specified, a specific level of access may be granted to that user that, by default settings, gives them privileges as explained in the table on the following page.

When a user creates a document, the document will be owned by that user. The user has the ability to change permissions on that document including adding other users specifically or allowing "Shared by link" access and sharing the link to the document.

By default, new files inherit the permissions of the folder in which they are created.

Documents inside more restricted folders can be more permissive. For example you can have a top level folder that is restricted and you can have a sub folder which can be open to the world. You can open up access wider to the documents and folders as you go lower in the hierarchy.

| **Google Drive sharing privilege abilities by access level** | | | | |
| --- | --- | --- | --- | --- |
| **Ability** | **Privilege Level** | | | |
| **Owner** | **Editor** | **Commenter** | **Viewer** |
| View the document | ✔ | ✔ | ✔ | ✔ |
| Copy the document | ✔ | ✔ | ✔ | ✔ |
| Print the document | ✔ | ✔ | ✔ | ✔ |
| Download the document | ✔ | ✔ | ✔ | ✔ |
| View document details (owner, created, last modified) | ✔ | ✔ | ✔ | ✔A |
| See other profiles viewing document | ✔ | ✔ | ✔A | ✔A |
| View and submit a message in document chat | ✔ | ✔ | ✔A | ✔A |
| Add comments to a document | ✔ | ✔ | ✔ |  |
| Make suggestions to a document | ✔ | ✔ | ✔ |  |
| Rename the document | ✔ | ✔ |  |  |
| Move the document to another folder | ✔ | ✔ |  |  |
| Remove the document | ✔ | ✔ |  |  |
| Copy the document | ✔ | ✔ |  |  |
| Modify sharing settings | ✔ | ✔ |  |  |
| View names & email accounts in sharing settings | ✔ | ✔ |  |  |
| Provide a user with temporary access | ✔ | ✔ |  |  |
| View document version history | ✔ | ✔ |  |  |
| View document view history | ✔ |  |  |  |
| Change ownership of file | ✔ |  |  |  |
| Restrict the ability of editors to change sharing settings. | ✔ |  |  |  |
| Restrict the ability to download, copy, and print | ✔ |  |  |  |
| Receive notifications for suggestions and comments | ✔ |  |  |  |

A - Anonymous users who are not authenticated to a Google account do not have this ability or it has been restricted in some way. When viewing profiles, only the name is shown and not the email. Authenticated users who have restricted the sharing of information about their account may also not have some of these abilities.

### GSuite for Enterprise / Education

Google also offers a more advanced version of Google Drive called GSuite for Enterprise and also is sometimes called GSuite for Education. These services are compatible and interact with the standard Google Drive but offer the ability for an organization to have it's own administrator and more control over user identities within its domain.

The institutions' GSuite administrator, which will be referred to as the local admin, can then turn on and off certain features and has the ability to limit installation of add-ons and API access.

The local admin can also perform maintenance operations such as recovering deleted files, moving files between users within the same domain, managing permissions, and more. The local admin however only has power over their local domain name, such as example.edu, so they cannot perform functions such as moving documents between accounts on different domain names.

Also, just because a Google Drive account is tied to an identity on an institutions' domain doesn't mean that that institution has subscribed to GSuite for Education.

### Google Shared Drive (formerly known as Team Drive)

In Google Shared Drive the organization is the owner of the files rather than the individual. Unlike files in standard Google Drive, files in Google Shared Drive belong to the team instead of an individual. Even if members leave, the files remain owned by the team. More information about Google Shared Drive is provided in the solutions section 'Google Shared Drive (Deep Dive)'.

# Security concerns

## Sharing problems

Google Drive offers users the ability to control access to documents with multiple degrees of access. These categories currently include "Restricted", "Anyone with link", "Anyone at <INSTITUTION>" where <INSTITUTION> varies according to the local institution name , and "Public".

### Sharing with link

Google Drive offers the ability to share a document with anyone who has the URL. This permissions category as described in the interface is called "Anyone with the link". The URL is available through the address bar in the web browser and also by selecting 'Get Link' from the document's context menu. In an earlier version of Google Drive, the "Get Link" feature was called "Get Shareable Link" and clicking on this would automatically change the permissions category to "Anyone with the link". This meant that a file with stricter permissions could be accidentally opened up to a broader audience through this feature. The authors of this document and their colleagues found that this was frequently the case.

You can open up the permissions so that anybody with the link can open the file. The problem here is that people who leave the project and have access to the file may still have the link. So, if the file is still openable by the link, they could access the file and could possibly change it. They can also change the permissions to it making it accessible to more people.

### Linked from other documents

This happens a lot where we have one GDoc linking to another GDoc. Websites posts also link back to a GDoc. Documents that contain links may be granted a broader level of access or may become public without the realization that they contain links to documents that are intended to be more restrictive.

### Email could be compromised

Email addresses get compromised all the time. And if you have a document that has been shared with that compromised email, then that document can be accessed by other people.

### Sharing with Public

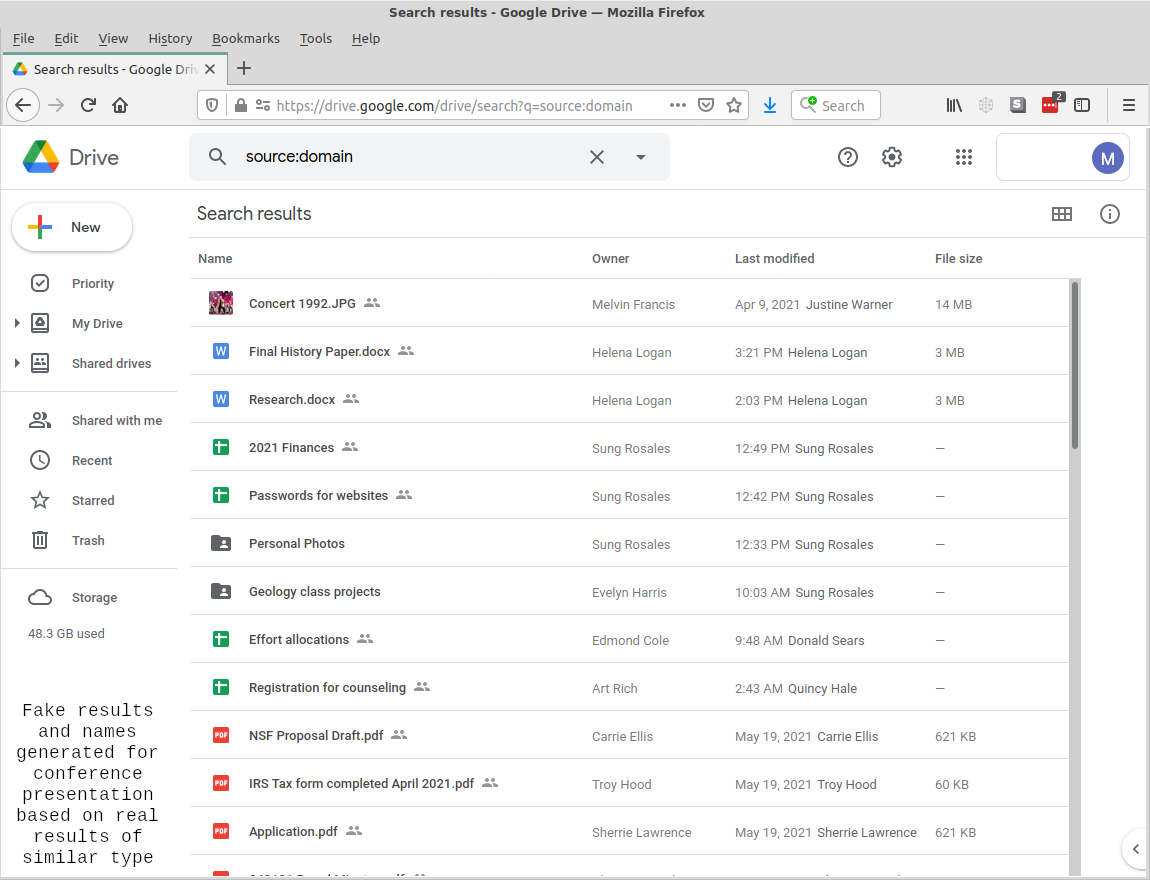
The "Public" or "Anyone" permissions category is no longer available through the Google Drive interface, however older files that were set to this may have retained this permission category. Documents with this category of permissions can be found through a Google search. For example, you can search for "site:docs.google.com password managers" on Google.com and you will find a document that has been shared with "Anyone". The key difference between this category and "Anyone with the link" is that the user doesn't need to know the link and thus the document can be discovered.

### Sharing with domain / Anyone at <Institution>

Another way you can share a document is by sharing it with the domain. Within an organization that has a GSuite for Education / Enterprise account, you can set permissions on files and folders to be viewable by anyone on your domain. Unlike 'Sharing with link' where the viewer must know the URL of the document, this level of permissions allows other users within the domain to discover documents they don't know about. Google provides this ability through the publicly disclosed[[6]](#footnote-5) search keyword called 'source:domain'. By searching Google drive using this keyword, a user can find all documents at an institution that have been shared in this way.

In our research it became clear that according to the names of many documents, the users most likely did not intend to share the documents in this way. Our theory as to how this happens involves the history of the interface of Google Drive. As mentioned in the section "Sharing with link" in this paper, documents would accidentally be opened up to the public if they had the link to the document. We found that after institutions started to adopt GSuite for Education that the accidental "sharing with the public" became accidentally "sharing with the institution". When this started to happen to folders, whole document hierarchies began to be exposed.

The screenshot below, while consisting of fake data, is based on real results from a search for 'source:domain' at a large educational institution and is meant to provide a sense of the type of exposures that could be possible at an educational institution.



*A faked screenshot showing the potential results of searching for documents shared with a domain.*

During our research we found indications that the public was already becoming aware of these types of exposures and that technical administrators were ignoring or downplaying the risk.[[7]](#footnote-6) A recent incident involved the exposure of sensitive Google Drive data over 3000 student and 100 employee accounts at New York City schools[[8]](#footnote-7). Based on the description of the incident, it is possible that the documents were exposed through this method.

The diagram below is a representation of a folder hierarchy in Google Drive for reference in the following two sections. These scenarios are representations of situations that we've encountered while onboarding and offboarding users.

* 🗁 Company (A)
  + 🗁 Management (B)
    - 🖹 "Policies" (B)
    - 🗠 "Budget" (B)
  + 🖻 "Newsletter" (A)
  + 🖺 "Document" (C)

### Revoking access

Revoking from the top of the tree is not 100% effective all the time. We have found for unexplained reasons that some users, when removed from the top of a folder hierarchy, still retain access to some folders despite being removed from the top-level folder.

The diagram above represents a folder hierarchy on Google Drive. Let's suppose we have 3 different sets of permission patterns that we've applied to files (A), (B), and (C). Each of these sets has a different group of users in the set and those sets may overlap somewhat, but they are not identical.

Next, we add a new user at the top of the hierarchy to the 'Company' folder and Google helps you by granting that user the same permissions recursively to all the other folders. The user now has access to all the documents in this hierarchy.

Later the user leaves and is offboarded from the organization. Management removes the user from the top level 'Company' folder and assumes that the user has been removed from all the sub folders. However the user has somehow retained access to all the documents that have the permissions pattern (B).

As expected, the user also retains access to documents they own. If new folders are created within folders they own, that can lead to access re-propagation of their permissions.

### Granting access gives too much access

User set (A) at an organization is given access to the top level folder 'Company' and documents below that in the hierarchy. Management at the organization restricts access to everyone but management in folder 'Management' and below to just access pattern (B). Later, a user is granted access to the top level folder 'Company' and Google applies access down the hierarchy, including giving access to the new user in the 'Management' sub-hierarchy. This gives the new user more access than is warranted.

### User accounts with privacy enabled

A user may choose to enable privacy settings within their Google account. While this is beneficial to the user, it limits the ability of document owners and editors to see who accessed a file and who made changes.

### Multiple institutions

There are a lot of times when universities collaborate with one another on a project and sometimes you might need to transfer the file owned by a person at university A to someone at university B. Gdrive does not provide a straightforward way to do so i.e. you can not transfer the file’s ownership between different domains. Other issues could be: some universities might not have GSuite for Education. Different universities could have different security policies. For example: one university could have two-factor authentication, might demand that the project have one as well, whereas another university might not require two-factor authentication.

### Offboarding when someone leaves

When someone leaves they retain access and they still know the URL. The issue is that they can access the file and make changes. If the user continues to own folders within the hierarchy, then new documents created under those folders may end up with the user gaining editor level access privileges.

### Ownership transfer

In a project that is spread across multiple institutions, offboarding can be especially challenging. Ownership of documents belonging to the organization must be transferred to a remaining colleague. While a user can transfer ownership within their own domain, Google Drive policy prevents users from transferring ownership to users on other domain names. For example if we have one user on the domain 'example.com' and 3 on the domain 'example.edu'. The user on 'example.com' owns 10 files belonging to the project. When this user leaves the project, they will not be able to directly transfer files to any of the other users on the project. Files can be copied by a receiving user, but the meta data such as version history and document ID will be lost in doing so.

### When a user deletes their account

Another issue arises when they delete their account. When someone deletes their account who had files owned through that account, the access goes away immediately for everyone who had access to those files and people get kicked out of the documents. The subfolders and files that were nested end up becoming orphaned and this creates an issue with file hierarchy that was set up.

## Ransomware threat

Ransomware has been a big threat in the last few years. Here, Google Drive Sync can be used as an attack vector. Google Drive Sync allows you to mount the Google Drive as if it were a drive on your local system.

Ransomware could infect this and encrypt Google Drive files if someone’s system gets infected with a ransomware who was using Google Drive Sync. Suddenly, you will find some of your files are encrypted because of one individual system being hit by ransomware. Here, you may be saved by doc versions by rolling the encrypted document back to the previous version. But this mostly depends on how the ransomware behaves. If the ransomware deletes the file entirely and then creates a new encrypted file instead, then this method of rolling back to the previous version will not work because all the metadata will be lost.

## Live document chat

The Google Docs suite of tools has a feature that allows authenticated users to talk to each other in real time within the document similar to how an instant messaging application works. The feature may not be enabled on your account and requires certain conditions in order to work. For users to see the chat icon and use the functionality, another user has to be viewing the document at the same time. Each of the users must be listed with their email identities with a specific permission level in the sharing settings of the document. Chat history is also not saved.  
  
A potential security issue with the chat feature is that it provides no way to moderate messages or report messages for spam or abuse. If a document happened to have it's sharing set to 'Anyone with the link can edit' then a Google drive user could add a target's email address to the sharing settings without notification and harass them when they view the document. Additionally, an authenticated user will see chat messages pop up even if they don't have the chat window open.

# Solutions

Below we will cover some of the solutions to the issues that we have highlighted above.

## Restricting permissions changes

Google has given the owner of a document the ability to restrict editors from making permissions changes to documents such as adding or removing users. This setting can be accessed by accessing the sharing settings of a document and clicking on the 'gear' icon in the upper right of the pop up window and uncheck 'Editors can change permissions and share'.

## Limiting access by time

It is possible to limit the amount of time that a specific user has access to a document. For instance, if you want to allow students to have viewer access to a file until the end of a semester, you can grant each student individual access and set a deadline of the last day of the semester. This is accomplished by first granting access to the user at the level you want. Then accessing the sharing settings again for that user and clicking on the access level drop down you should see an option 'Give Temporary Access', which allows you to set an expiration time for the permission.

## Auditing Permissions - GSuite Administrator

When an organization such as a university signs a contract with Google for use of Google Drive on their domain, they have the ability to designate someone as a GSuite Administrator. This person has special access to Google Drive and can perform many operations that normal users may not. There are still some restrictions as to what can be done. For instance, files cannot be transferred between accounts on different domain names. Also, files moved to the trash cannot be recovered after a certain number of days has passed. Those who are GSuite Administrators may find the open source tool set called *gam*[[9]](#footnote-8) useful. A GSuite Administrator can use these tools to determine the files a user has access to, remove, or grant permissions.

## Auditing Permissions - Google Docs Add-on

The Google Docs suite of tools may have add-ons that can be used by people who are not GSuite administrators, nor have API access. One of these add-ons is 'Drive Permissions Auditor' by 'Digital Inspiration'[[10]](#footnote-9), which is an add-on for 'Google Sheets'. It doesn’t require the user to apply for special API access, all you need is add-on permissions. The free version of this add-on is limited to scanning only 100 documents. As a general principle you need to be cautious about any Google Docs add-ons since you are giving a 3rd party developer permission to access your data and this brings additional security risks[[11]](#footnote-10). For example, the 'Drive Permissions Auditor' add-on required the following permissions as explained by Google when authorizing access to the add-on:

* See, edit, create, and delete all of your Google Drive files
* See, edit, create, and delete all your Google Sheets spreadsheets
* Connect to an external service
* Allow this application to run when you are not present
* View and manage data associated with the applications
* Display and run third-party web content in prompts and sidebars inside Google applications.

The add-on makes the statement under 'What is your privacy policy?' that 'Your data is neither uploaded anywhere, stored nor shared with anyone.' While this may be their policy, you are still granting them the potential for this level of access and a vulnerability in their code could provide an attacker with any of the privileges listed above that you have agreed to.

## Auditing Permissions - Cloudperm

Google also offers API access to Google Drive. An application suite called Cloudperm[[12]](#footnote-11), which was written by Mark Krenz and others, is available for performing several auditing and permissions management operations against Google Drive. This suite currently allows a user with API access to list all the files and permissions under a Google Drive folder, revoke permissions on documents, and move files. It also will give a warning when a document has been shared by link. As of this writing functionality that will assist in migrating files between accounts on different domains using the method described in this paper is in development.

These solutions for auditing permissions are especially notable considering that since April of 2021, Google has restricted or broken the ability for users to search documents trees for access rights using their own user interface 'Shared with' or the 'to:' keyword. Google has not responded to the author's repeated requests for help on this problem.[[13]](#footnote-12)

## Setting policy

Another thing one can do is set up policies i.e. setting up requirements for access management, setting up a strong password, using 2-factor authentication, etc. You can also set up a policy to restrict add-ons except the ones made by Google since they already have access to our data, so using their add-ons should be fine. Another setting restriction on mounting and syncing softwares like Google Drive Sync and File Stream as they can be used by ransomwares as explained earlier.

## Using document name tags

In an organization where there are a variety of sharing use cases, for instance sharing documents with outside partners by link, it may not be clear what the intended level of sharing is for a document. This presents problems with trying to audit the permissions of documents, as it is not known whether a document is properly shared or not. To help alleviate this problem within our organization, we have implemented a Document tagging policy that requires the use of text based tags within the name of a document on Google Drive. For example, a spreadsheet document called 'Global Temperature Readings' that contains scientific data meant to be shared with the public might have it's sharing settings set to "Anyone with link may view". To indicate to the security staff that this is the intended level of sharing, the document owner could add the text '[SHARED-LINK]' at the end of the document name, thus the full name becomes 'Global Temperature Readings [SHARED-LINK]'.

In order to reduce the clutter of tags on every document, we have allowed for 'inheritance' of tags and the use of an '[INTERNAL]' tag at the top level folder that propagates its meaning down the hierarchy. This means that mostly documents that are not Internal to the organization need to have a tag. Below are examples of the tags that we've developed for our use and their meaning.

* No tag - If the document has the same sharing intent as its parent folder, then no additional tag is required. The document will be considered to "inherit" the tag of it's parent folder.
* [INTERNAL] - Use this tag when the document is only for sharing with organizational staff. Only current organizational staff members should be listed by email address in the sharing settings. Additionally, a document with this tag should have its link sharing setting under the 'Get Link' section of the sharing settings set to 'Restricted'.
* [SHARED-EMAIL] - Use this tag when the intention is to share the document with a specific set of users outside of organizational staff, who are identified by email address within the sharing permissions of the document. A document with this tag should have its link sharing setting under the 'Get Link' section of the sharing settings set to 'Restricted'.
* [SHARED-DOMAIN] - Use this tag when the intention is to share the document with anyone at an institutional domain. For example, "Anyone at <Institution>" or "Anyone at example.edu". A document with this tag should have its link sharing setting under the 'Get Link' section of the sharing settings set to the institutional domain for which it is intended to be shared. Unless the intention is really to allow anyone at an institution to discover and use a document, you are recommended not to use this tag and level of sharing.
* [SHARED-LINK] - Use this tag when the intention is to share the document with anyone who has the link to the file. A document with this tag should have its link sharing setting under the 'Get Link' section of the sharing settings set to 'Anyone on the internet with this link'.
* [SHARED-OWNER] - Use this tag when a file has an owner that is outside of the organization and this is intentional.
* [PUBLIC] - Use this tag when the intention is to make the document completely public and discoverable through a Google search. Google no longer offers this feature in sharing settings, although there could be some older documents that still have this level of permission.
* [SKIPBACKUP] - This tag can be useful to indicate to backup systems that you wish to skip a file, perhaps due to sensitivity or size issues.

Note that these are only by our own naming convention, it is not a feature of Google Drive.

## Google to restrict access to 'Shared by link' files

In September of 2021 Google is scheduled to change the way 'Shared by link' access works. Previously anyone publicly with or without a Google Drive account could access a document simply by knowing the URL for the document. They will change this access model so that by default someone visiting the document who is not listed in the permissions explicitly and who has not visited the document before will need to request access to the file. This method of requesting access is expected to be similar to the method used for request access now when a user does not have access. The owner of the document has the option of turning off this restriction on a document by document basis if they do not wish to handle requests.

## Backing up Google Drive documents

Some of the security risks around Google Drive include the possibility that documents will be accidentally deleted, moved to an unknown destination, or subject to a ransomware attack. Because of this, users can benefit from periodic backup of their Google Drive documents. Currently there are some significant limitations to any of the solutions listed below. First of all, document IDs will not be preserved when documents are restored to Google Drive from backups. This could mean that links between documents or links from the Internet or from bookmarks will be broken. Another drawback is that version history on files will not be preserved in the downloaded file. Also, Google Drive permissions will not necessarily be preserved on documents, so if you had some special permissions, those will be lost. Fortunately, permission auditing solutions like Cloudperm mentioned in this paper can help restore permissions to their original state.

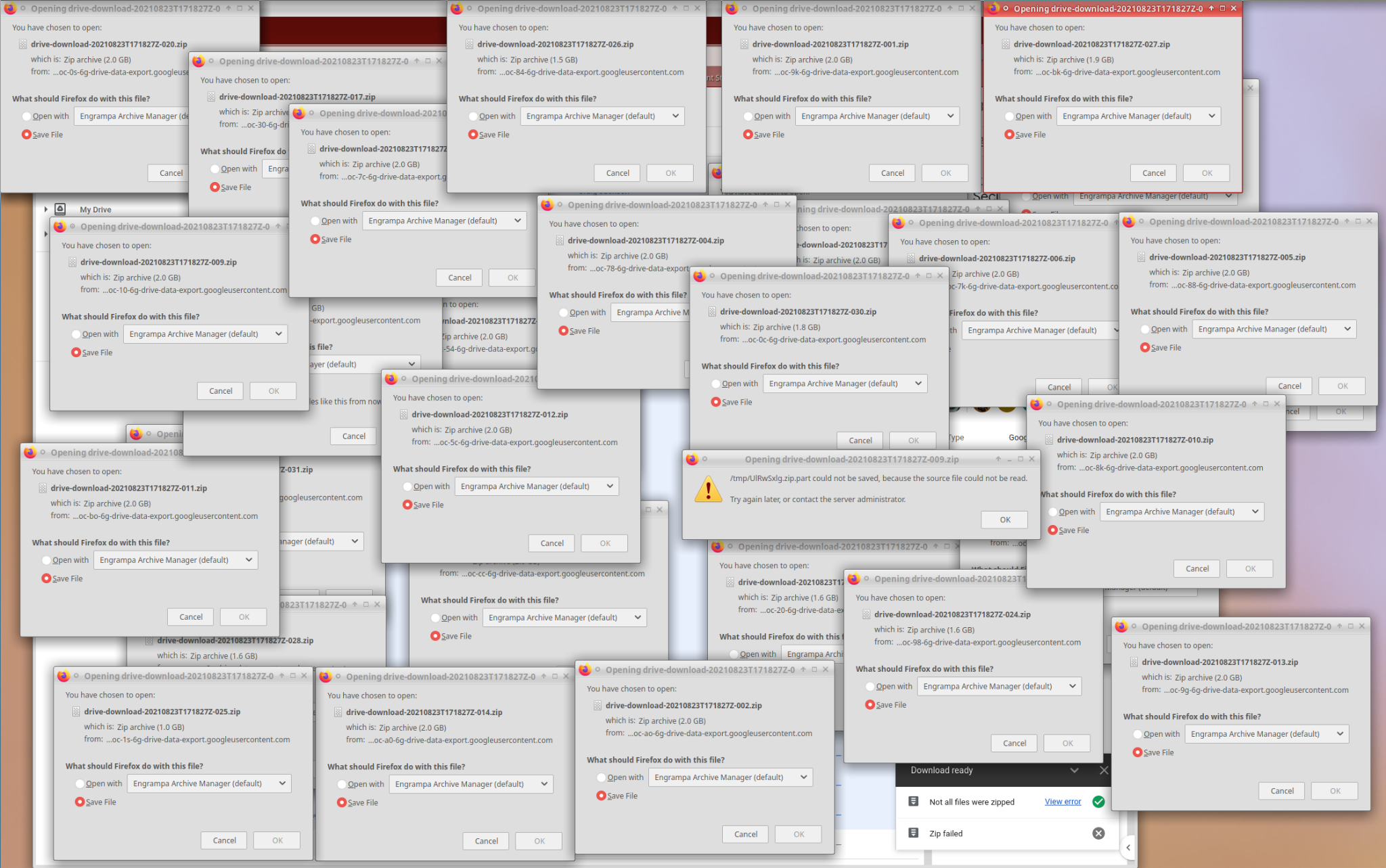
Historical note: Previously comments in documents were not preserved when downloading a document from Google drive through any method, however that feature has been added in recent versions.

### Using the Google Drive UI

Google provides a simple method for backing up documents through the Google Drive interface. This is accomplished by opening the context menu for a document and selecting 'Download', or by selecting multiple documents or folders and selecting 'Download'. While this is a simple method to employ, it does have some limitations and drawbacks.

The first drawback is that this method can take a long time and give you little feedback on it's progress, especially if you have many files or large files. Performing a download of a gigabyte of data may take several minutes just for Google to prepare the file for download, let alone the time to download the file itself. This means a large document tree (ie. 10GB or more) may take several hours before the user is prompted to begin the download process. During this preparation time it's possible to navigate to other folders, but you must leave the browser window or tab open that is performing the backup.

The authors have also run into issues downloading more than 2GB of documents and received errors when trying to do so. This meant that in order to get a complete backup of a folder tree using this method one would have to select different combinations of subfolders and documents to stay under the 2GB limit. Recently Google has changed it's interface to allow the downloading of more than 2GB of documents at once, however it splits up the resulting zip file into 2GB chunks, resulting in a potentially large number of download pop-up windows and potential user confusion about the completion of the integrity of the download. This is prone to error because Google Drive provides no obvious way through the standard interface to determine how many bytes of space are used under a folder.



*A screenshot of all the pop-up download windows asking to download 2GB sections of a large download.*

One advantage through the Google Drive UI is that Google will scan smaller downloaded files for viruses. Larger files are skipped. Virus detection can also sometimes block a download from completion.

### Google Sync and Google File Stream

Windows and Mac users have the ability to use software provided by Google that will allow them to treat their Google Drive space as a network drive on their local computer. This is an interface that could allow the backup of Google Drive documents en masse. However, the threat of ransomware affecting Google Drive through this method should encourage you to look for a better solution. At least consider creating a dedicated host that is only for the purpose of using Google File Stream and doing backups that have tight security controls.

### FUSE filesystem

On Linux systems, a kernel module called FUSE allows novel connections to be made between various services and the filesystem interface. There is a FUSE module called *google-drive-ocamlfuse[[14]](#footnote-13)* which allows a Google Drive folder to be mounted to the Linux filesystem. This uses an API key to access a Google Drive hierarchy. This has the benefit of allowing the user to use standard Linux tools such as 'ls', 'cd', and 'tar' to navigate the folder hierarchy and perform backups. When documents are copied through this interface, they are converted into their related offline forms. For example Google Doc documents are converted into Open Document Text Format files. These files can be opened by popular word processors. They can also be reuploaded to Google Drive where they can be used through the Google Docs suite of applications. Files can also be uploaded to Google Drive by placing them inside the google-drive-ocamlfuse mountpoint.

Using this method it is possible to download several gigabytes worth of files in an unattended way. For example, you can use the tar command to backup an entire hierarchy using the following commands:

* google-ocaml-fuse /mountpoint
* tar pcvf google-drive.tar /mountpoint

You can expand upon this method to create encrypted backups without an intermediate file by pipelining the output from the *tar* command directly into the *gpg* command as shown in this command:

* tar pcv /mountpoint | gpg --encrypt -r email@example.com --output \ google-drive-encrypted.tar.enc

Backups of large files such as videos through this method can be slow and you may consider skipping the download of those files using tar options such as --exclude. As shown in the section on document tags in this document, you can tag files that you wish your backup process to skip by putting a tag such as [SKIPBACKUP] in a filename and then modifying your tar command by adding the option --exclude='\*\[SKIPBACKUP\]\*' before the directory you wish to backup.

### Rclone

Another option that uses an API key is the program *rclone[[15]](#footnote-14).* The *rclone* program is a more mature and more fully featured cloud storage access solution that also provides the ability to access over 40 different cloud storage systems in addition to Google Drive.

## Cross domain file ownership transfer

Transfering the ownership of files between domains in Google Drive is a big issue if your project is a multi-institutional project. According to Google you can only transfer the ownership of files within your institution's domain. That means if a user at example.edu owns files, they can’t transfer the ownership of these files to someone else at some other domain like gmail or example2.edu. One solution is someone manually creates a copy of files I want to transfer to them and then delete my older files.

There is another way to transfer these files to someone else at some other organization.[[16]](#footnote-15) This solution requires creating a ‘shared drive.’ Then you add the current owner and the ‘to be’ owner of the files as a member of the shared drive. Even if they are members of different domains, you can add them to the same Shared Drive. And that is how the magic tunnel works.

Then, the person leaving or the original owner transfers their files to the shared drive. When they transfer their files to the shared drive, the drive becomes the owner of those files.Next, the new owner would transfer the files from the shared drive back to where they originally belonged and they own it.

There are few caveats with respect to this method:

* The folders can’t be transferred.
* File permissions are not preserved.
* File hierarchy is not preserved and will be lost.

## Google Shared Drive (deep dive)

Google Shared Drive, which was formerly known as Team Drive. is an organizational structure that lives parallel to My Drive, but uses a different document ownership model. In shared drive files are owned by an organization rather than a single person. Unlike files in My Drive, files in shared drive belong to the team instead of an individual. So, even if the members leave, the files stay exactly where they are, so your team can continue to share information and get the work done.

Some important points to note about Google Shared Drive:

1. Files can be moved from My Drive to shared drive or from one shared drive to another.
   1. Moving files and folders to a shared drive changes ownership from the user to the shared drive.
   2. When you move a file to a shared drive, only members of the shared drive and people the file is directly shared with can access the file.
   3. If the original owner of a file is in your organization but not a member of the shared drive, they lose ownership but can still access the file as its creator.
   4. Files and folders moved to a shared drive don’t remain in other Drive locations, such as My Drive, but are still accessible from the user's Shared with me and Recent views.
2. Collaboration with users from other domains (external users):
   1. The shared drive appears in the external user's Google Drive.
   2. Any work an external user contributes (for example, edits to, creating, or uploading a file) is transferred to and owned by the domain that created the shared drive.
   3. You and other admins can’t migrate folders or files owned by external users even if the external user is a member of the destination shared drive.
3. When a member leaves, the files stay in the shared drive and can be used by the other members since the content is owned by the team.
4. As a G Suite administrator, you can restore files that were deleted from a shared drive. If the entire shared drive was deleted, you can also restore the shared drive and its contents.
5. Shared Drive Limitation:
   1. Shared Drive can contain a maximum of 400,000 files and folders.
   2. A given file within a shared drive can be directly shared with a maximum of 100 groups.
   3. A folder in a shared drive can support up to 20 levels of nested folders.
6. Moving the files from the shared drive to My Drive and vice-versa doesn’t change the URL or the ‘Share Link’ of the files. Moving the files from one shared drive to another doesn’t change the URL of the files, either.
7. Access Permissions: There are two classes of permissions in shared drives
   1. Users can be granted access to the shared drive itself where they can access all the files and folders within the Shared Drive.
   2. Users can be granted access to the specific files and folders in shared drive on the need to know basis.

#### Access levels

Google Shared Drive also provides you the ability to assign different roles to individual users. The various roles are like manager, content manager, contributor, commenter and viewer. Different levels of access for users along with permissions (actions) are provided below from the Google Support website[[17]](#footnote-16):

| **Permission** | **Manager** | **Content manager** | **Contributor** | **Commenter** | **Viewer** |
| --- | --- | --- | --- | --- | --- |
| Can view files and folders | ✔ | ✔ | ✔ | ✔ | ✔ |
| Can comment on files | ✔ | ✔ | ✔ | ✔ |  |
| Can edit files | ✔ | ✔ | ✔ |  |  |
| Can create and add files, can create folders | ✔ | ✔ | ✔ |  |  |
| Can add people and groups to specific files | ✔ | ✔ | ✔ |  |  |
| Can restore files from the Trash (up to 30 days) | ✔ | ✔ | ✔ |  |  |
| Can move files from My Drive to a shared drive | ✔ | ✔ | ✔ |  |  |
| Can move files and folders to Trash | ✔ | ✔ |  |  |  |
| Can move files and folders within a shared drive | ✔ | ✔ |  |  |  |
| Can add people and groups to specific folders in shared drive | ✔ |  |  |  |  |
| Can move files from one shared drive to another shared drive | ✔ |  |  |  |  |
| Can add or remove members of a shared drive | ✔ |  |  |  |  |
| Can change member access levels | ✔ |  |  |  |  |
| Can permanently delete files in the Trash | ✔ |  |  |  |  |

## 

## Using Google Groups for IAM

Google Groups[[18]](#footnote-17) are assigned an e-mail address that can be referenced in sharing rights for a document. This provides the ability to give access rights to an entire group and its members at the viewer, commentor or editor level. It is also possible to increase the rights of an individual who is a member of a group by adding them individually to the list of shared privileges on a document. For example, if user A is part of group B, but group B only has "viewer" privileges, then user A's privileges can be increased to commentor or editor by adding them to the document's share settings with a different level of access.

It does not allow the reduction of privileges of a user. For example if user A is part of group B, and group B has editor access, you cannot reduce the privilege level of user A to commentor or viewer or remove their access altogether.

The authors of this paper have not yet thoroughly tested this method of providing access.

# Conclusion

There are many security concerns to consider when using a cloud storage system such as Google Drive. The academic and research communities are especially vulnerable to many of the concerns that have been presented in this paper. It is recommended that the security team for these organizations become aware of and mitigating these risks.

The authors of this paper would like to hear from you if you have suggestions for improving this document. Please contact the authors at [mkrenz@iu.edu](mailto:mkrenz@iu.edu) or [iabhinit@iu.edu](mailto:iabhinit@iu.edu).

1. https://researchsoc.iu.edu/ [↑](#footnote-ref-0)
2. https://www.trustedci.org/ [↑](#footnote-ref-1)
3. https://sciencegateways.org/ [↑](#footnote-ref-2)
4. https://continuousassurance.org/ [↑](#footnote-ref-3)
5. https://support.google.com/drive/thread/2035857/download-quota-exceeded?hl=en [↑](#footnote-ref-4)
6. <https://support.google.com/a/thread/5253719/how-can-i-find-all-shared-files-in-my-drive-that-can-be-found-by-anyone-in-my-organization?hl=en> [↑](#footnote-ref-5)
7. <https://www.reddit.com/r/k12sysadmin/comments/jb474i/student_you_guys_should_go_through_sourcedomain/> [↑](#footnote-ref-6)
8. <https://support.google.com/a/thread/5253719/how-can-i-find-all-shared-files-in-my-drive-that-can-be-found-by-anyone-in-my-organization?hl=en> [↑](#footnote-ref-7)
9. <https://github.com/jay0lee/GAM> [↑](#footnote-ref-8)
10. <https://digitalinspiration.com/> [↑](#footnote-ref-9)
11. <https://spinbackup.com/blog/security-risks-google-add-ons/> [↑](#footnote-ref-10)
12. <https://github.com/deltaray/cloudperm/> [↑](#footnote-ref-11)
13. <https://support.google.com/drive/thread/105396997/the-shared-with-search-functionality-is-not-working?hl=en> [↑](#footnote-ref-12)
14. https://github.com/astrada/google-drive-ocamlfuse [↑](#footnote-ref-13)
15. <https://rclone.org/> [↑](#footnote-ref-14)
16. <https://www.tabgeeks.com/tabgeeks-blog/how-to-transfer-ownership-between-domains-in-google-drive> [↑](#footnote-ref-15)
17. <https://support.google.com/a/users/answer/9310249#1.2> [↑](#footnote-ref-16)
18. <https://groups.google.com/> [↑](#footnote-ref-17)