



## Collecting Cloud Attachments with Aid4Mail

---

### *User Guide*

Fookes Software Ltd  
Charmey, Switzerland  
[www.aid4mail.com](http://www.aid4mail.com)

## Collecting Cloud Attachments with Aid4Mail

Aid4Mail 6 introduces a powerful feature that allows you to collect modern cloud attachments from platforms like **Google Drive** and **Microsoft OneDrive** (Personal, Business, and SharePoint). This includes the ability to gather document revisions, enhancing your eDiscovery capabilities. Google refers to these as “Linked Drive Files,” and Microsoft uses the term “Shared Drive Links.”

Here’s how to manage these attachments using Aid4Mail:

## Configuring Cloud Attachment Collection

On the **Settings** tab under the **Source** section, you’ll find a new option labeled “*Collect cloud attachments and their metadata.*” When enabled, Aid4Mail checks each processed email for hyperlinks pointing to files stored in the cloud. The behavior of this feature depends on your selected settings:

### Microsoft Accounts

- **Account Selection:** This drop-down lists email accounts accessible via the Microsoft Graph API. If the desired account is missing, click **Add**. Note that this option is hidden when you’ve already selected the Graph API as your source format.
- **Credentials:** Personal accounts like Hotmail and Outlook.com only access files on OneDrive Personal, while work accounts (Microsoft 365, Exchange Online) access both OneDrive Business and SharePoint.

Aid4Mail uses the selected account’s credentials to download files from OneDrive and SharePoint. If no account is selected, Aid4Mail will only access publicly available files.

### Google Accounts

- **Account Selection:** Similar to Microsoft accounts, this drop-down lists Google accounts accessible via the Google API. Click **Add** if you don’t see the desired account. This field is hidden when Gmail is selected as the source format.

Aid4Mail uses the selected Google account credentials to download hyperlinked files from Google Drive. If no account is selected, it will only access public Google Drive files.

### File Size Limits

Use the “*Collect files up to*” drop-down to control the size of the cloud attachments Aid4Mail collects:

- **Only metadata:** Collects only file metadata without downloading the actual files.
- **No limit:** Collects all files, regardless of size.
- **Size limits:** Choose a limit between 10 MB and 500 MB.

In all cases, Aid4Mail collects metadata and stores it in a CSV file at the location specified in the **Cloud attachment location** setting under the **Target** section. For more information, see the **CSV File with Metadata** section.

## Exporting Google Docs

Google Docs, Sheets, Slides, and Forms are exported in their corresponding Microsoft formats (Word, Excel, PowerPoint) or CSV. Keep in mind that Google's export size limit is 10 MB—files exceeding this limit won't download.

## Document Revision Matching

When the **Match document revision** option is enabled, Aid4Mail retrieves the version of a cloud document that was current at the time the email was sent. This is useful if you need to review the exact version that was shared. When Aid4Mail downloads a revision, it appends the document version number in square brackets to the end of the file name—for example, **MyFile [59].docx**.

However, document revision matching is a double-edged sword. This is because cloud providers typically send email notifications when a document is updated. For example, Google and Microsoft both send a notification when a document is initially shared. Afterward, Google sends notifications for every batch of document edits or comments. Microsoft only notifies for comments, not for edits alone.

This means that heavily edited, or frequently commented, documents can potentially generate hundreds of revisions and notifications. When Aid4Mail processes the notification emails with **Match document revision** turned on, it treats each document revision as the current one for its associated email. This results in Aid4Mail downloading all of them.

While this approach is certainly comprehensive, it has some serious side effects:

- It can significantly increase processing time.
- It may quickly lead to bandwidth throttling if the cloud provider's quotas are exceeded.
- You end up with a multitude of versions of the same document.

To mitigate this problem, you could filter out automated revision notifications from your search. For example, to eliminate Google Docs notifications, you would use the following filter:

```
NOT From:comments-noreply@docs.google.com
```

However, this is risky because it assumes the original document link was sent in a regular email. This may not be the case. If the original document link was only communicated through one of these automated notifications, the above filter would prevent any collection of the shared document at all.

For this reason, we recommend leaving **Match document revision** unchecked and not filtering out the notifications. Then, Aid4Mail will only download one version of a shared document—the latest available at the time of collection. You can still access details of earlier document edits and comments from the revision notifications.

By using the latest version of the document and email body searches of the notification emails, you can usually identify any relevant text and eliminate the need to collect multiple versions of the same file.

## Recommendations for Efficient Cloud Collection

To optimize your cloud attachment collection while minimizing bandwidth and time costs, follow these best practices:

- **Filter early:** Use pre-acquisition and local filters to limit email collection. Narrow the scope to relevant time periods, contacts, and exclude irrelevant folders and labels.
- **Minimize revisions:** Only download document revisions when necessary.
- **Limit file sizes:** Modern attachments can be enormous (up to 5 TB in Google Drive). Aid4Mail offers file size limits to prevent downloading unnecessarily large files.
- **Use incremental processing:** If you encounter interruptions, Aid4Mail allows you to resume collection without starting over.
- **Export first with Google Vault** (when possible): This will avoid Google's bandwidth limitations and subsequent throttling, enabling Aid4Mail to process the files locally and significantly faster. Make sure to include linked Drive files in your export. In this case, you can turn off the *"Collect cloud attachments and their metadata"* option in Aid4Mail unless you also need to collect Microsoft cloud attachments.

## CSV File with Metadata

Aid4Mail saves cloud attachment metadata in a CSV file named **FileMetadata.csv**, containing the following columns:

- **Author, Collaborators, Viewers:** Email addresses of those with access to the file.
- **DateCreated, DateModified:** Timestamps (UTC) of when the file was created and last modified.
- **DocId, FileName, FileSize, MimeType, Title:** Properties of the cloud attachment.
- **DownloadLink:** URL to download the file, available when accessing versioned documents or files that don't require credentials.
- **Hyperlink:** Link from the email that points to the cloud document.
- **Version:** Document revision or version ID; used if the **Match document revision** option is enabled.
- **Status:** Indicates the file's accessibility, with values such as:
  - **File found:** File accessible and metadata collected.
  - **File inaccessible:** Metadata couldn't be retrieved.
  - **File not found:** 404 HTTP error.
  - **File saved:** Successfully downloaded and saved.
  - **File unavailable:** 429 HTTP error—temporarily unavailable.
  - **Forbidden:** 403 HTTP error—insufficient permissions.
  - **HTTP error:** Errors above 299, excluding 401, 403, 404, and 429.
  - **Unauthorized:** 401 HTTP error—invalid credentials.
  - **Save error:** File found, but an error occurred while saving.
  - **Unzip error:** Corrupted ZIP container when accessing local attachments from Google Vault.