

## Bulk Import PDFs

- Overview
- Extract PDFs from your prior system
- PDF directory and naming conventions
- Import PDFs into SameGoal
- Reviewing an import
- Common Issues

### Overview

During implementation, PDFs from your prior system may be imported into SameGoal, **giving staff access to historic documents authored in your previous system even after converting to SameGoal.**

Imported PDFs will not bootstrap program participation, deadlines or reports. In order to bootstrap your SameGoal environment in year one, use the Data Import tool in addition to the PDF Import tool.

The PDF Import tool associates an unlimited number of documents per student going back arbitrarily many years. Documents are associated by attaching them to existing student records in SameGoal. Because this tool does not create student records, either run a demographics integration or use the Bulk Loader to create student records prior to import.

## Extract PDFs from your prior system

**In order to import PDFs into SameGoal, you must first extract the PDFs from your prior system.** For most special programs vendors, you'll need to submit a request to let the vendor know you are planning to transition from their system and request your data be made available as PDFs. Many vendors provide this service at no cost or for a marginal fee.

Each special programs vendor has a slightly different process, but SameGoal has helped districts transition from a wide variety of systems. Check with your implementation coordinator regarding best practices districts have used previously for transitioning from your particular vendor.

## PDF directory and naming conventions

How a PDF directory is processed is driven by the directory directly under the domain directory (eg "students", "students-nonstudentid", or no directory at all).

- **Option 1 (Best, typically highest association rate)** - Place PDFs in a per-student directory in which the directory is named based upon the student ID number of the student.
  - All PDFs in a given directory are associated with a student record based on student ID match only.
  - No best effort document text extraction is used.
  - e.g.: **<domain>/students/<studentID>/\*.pdf**
- **Option 2** - Place PDFs in a per-student directory that has a name other than the student ID number (eg BobSmith). All documents in the directory are assumed to be for a single student.
  - Best effort document text extraction is performed (looking for student identifiers that uniquely identify the student record in SameGoal).
  - Clustering is performed (if such a unique student is found, **all** documents in the directory will be associated with that student).
  - e.g.: **<domain>/students-nonstudentid/<per\_student\_directory>/\*.pdf**
- **Option 3 (typically lowest association rate)** - Place PDFs in any directory structure with any filename.
  - Best effort document text extraction is performed only (looking for student identifiers that uniquely identify the student record in SameGoal).
  - Documents that do not contain text-based unique student identifiers (eg scanned image-only signature pages) will not be attached to a student record (whereas they might have been in options 1 or 2).
  - e.g.: **<domain>/\*.pdf**


**Tip Required:** To ensure that the directory being uploaded into an environment is for the desired district, the top directory must be the district's SameGoal domain (e.g. **applegrove.org/...**) or the import will fail. When initiating a new import, the district's domain will be listed at the top in the blue bar.

## Import PDFs into SameGoal

The **PDF Import Tool** bulk imports PDFs in 2 stages. First, files are uploaded into SameGoal by an administrator, then they are automatically associated with the corresponding student.

### Steps *To bulk import PDFs:*

1. Log in to SameGoal as an administrator.
2. Visit **Settings > PDF Import**.
3. Click the **New PDF Import** button.
4. Select the folder on your local computer that contains PDFs from your prior system below.
  - **Required:** The directory must be named the district's domain.
  - Once a folder is selected, the tool will automatically begin uploading the PDFs to SameGoal servers (**Stage 1: Upload**).
  - A progress table will show the upload status of each document.
  - While PDFs are being uploaded, **do not interrupt the process** (eg close the tab, close the browser, hit the back button, log out of SameGoal, etc). Doing so will interrupt this stage and require you to restart.
5. A summary of the number of uploaded documents for processing and any upload errors will appear on screen after Stage 1 is complete.
6. Once PDFs have been uploaded to SameGoal servers, they are automatically inspected and associated with the corresponding student in SameGoal when possible (**Stage 2: Associate**). Because this phase happens server-side, you may close the tab, browser, work on something else in SameGoal or even log out while PDFs are being associated.
7. Once Stage 2 is complete, you will receive a notification email with a link to review results. Follow this link to see a summary report.

 **Tip** This tool may be re-run multiple times for any folder. PDFs that have already been associated will not be re-associated.

## Reviewing an import

After the PDF Import tool is run, review the report to ensure all of the desired PDFs were associated to the correct student record. If you find PDFs that should have been associated but were not, address any issues and re-run the PDF Import tool. This can be done as many times as necessary to get the desired results. You can also manually attach PDFs if you wish.

**Steps** *To view the results of a PDF import:*

1. Log in to SameGoal as an administrative user.
2. Go to **Settings** > **PDF Import**.
3. Select the import summary you'd like to view.
  - Summaries are sorted by the date and time the import tool was run.
4. Use the **Result** and **Note** columns to review. You may filter on these columns in-app if you're on the Plus or Pro edition of SameGoal. Otherwise, you may download the report and filter the results locally.
  - **Result:** This column shows whether or not the document was associated. Potential results include:
    - **Associated:** The PDF has been associated to the student listed in the Student column.
    - **Previously Associated / Duplicate:** The PDF was already associated to a student in a previous import.
    - **Not Associated:** The PDF was not associated and is less than 3 years old.
    - **Not Associated (older than 3 fiscal years):** The PDF was not associated and is more than 3 years old. The document's age is not the reason it did not associate. However, this note is included since if a PDF is older than 3 fiscal years, it is likely less relevant to current document and might not need to be reimported.
  - **Note:** If the document was not associated, this column lists the reason why. (See the common errors below.)

**Tip** SameGoal does not store PDFs that were not associated with a student record. It is the responsibility of the district to save and store these documents.

## Common issues

The most common reasons a document is not associated to a student record include:

- **Unsupported file type:** Only PDFs are imported into SameGoal. No other file types (eg: doc, docx, jpeg, etc) can be associated. If you wish to import a different file-type, you must first convert it to a PDF prior to uploading it to the SameGoal PDF Import tool.
- **File is too large:** SameGoal only associates PDFs that are allowed by your current tier.
  - SameGoal Standard - 50MB
  - SameGoal Plus - 100MB
  - SameGoal Pro - 150MB
- **Student not in SameGoal:** The student a PDF should be associated with is not currently in SameGoal (e.g. graduated, transferred, etc. prior to implementation).
- **Missing student demographics:** Each PDF must meet at least one of the criteria below:
  - Be contained in a per-student directory that is named with the student ID number (see option 1 above).
  - Contain both the (a) student ID number and (b) student last name.
  - Be contained in a per-student directory that includes at least one other document that contains the student ID number and student last name (see option 2 above). Without this information, it is not possible to associate the PDF with a particular student.
- **PDF is image only:** When a PDF does not contain text, the SameGoal PDF Import tool cannot directly match the document to a student record. Therefore, if you have many image only PDF documents, it is highly recommended to use a per-student directory naming convention for the import.