

# Project Description

Online File Manager is a very useful service. It can be a standalone service like Dropbox and Box.com, or it can be part of a web application like in CSNS and Canvas. In this semester we will develop an online file manager application and use it to practice the things we learned in the class. The requirements of the application are as follows.

## Users

Users must log in to use the online file manager. A user can only access the files he or she uploaded, unless the file is explicitly marked as public by its owner, or the file is shared with the user by its owner. Folders

A user can create, rename, copy, move, and delete folders. Folders can be created either at the top level (i.e. no parent folder), or under an existing folder. Copying/moving/deleting a folder will copy/move/delete all the folders and files in the folder.

## Files

A user can upload, rename, copy, move, and delete files. Files can be uploaded either at the top level (i.e. no parent folder), or under an existing folder. If a user uploads a file to a folder that already has a file with the same name, the system automatically creates a new version for the file. The system keeps all versions of each file (in case a user wants to revert back to a previous version). Copying a file only copies the latest version of the file; moving/deleting a file will move/delete all versions of the file.

## Folder and File Sharing

The owner of a file or folder can mark the file or folder as public so it can be accessed by any user (including user who are not logged in). Note that marking a folder public means anyone can see the names of the files and folders in this folder, but it does not automatically make those files and folders public.

The owner of a file can share the file with other users. These users must be users of the system and they must log in to access the file. For simplicity we will not allow sharing folders.

## Design and Implementation

There are two types of data in the system:

- The information about users, folders (e.g. name, hierarchical structure), and files (e.g. name, type, size)
- The actual files (i.e. the content of the files)

The first type of data must be kept in a database. There are two important reasons for this: efficiency and flexibility. Disk operations are quite expensive, so by keeping folder and file information in database rather than creating those folders and files physically on disk, it's much more efficient to perform certain operations such as renaming/moving folders and files. Also by decoupling file information and file storage, it allows us lots of flexibility on how the files can be stored. For example, we could store the files on multiple disks or multiple servers or in multiple cloud storage services, and to the user it still seems like their folders and files are in one place. The second type of data, the actual files, can be stored either in database or outside database (i.e. as regular files on disk). There are pros and cons for each approach. For an application like online file manager, I would recommend storing files outside database as it'll be faster, and much cheaper if you host the application in a cloud.