Upload more file types

# Why?

Not all the files that need to be shared with the field are PDFs. Spreadsheets, images and word docs are the most common other file types that field users might need access to when out on a jobsite. They will have native apps to view those files on their devices but customers want to be able to share all of that information in one place.

It will also allow us to observe which other files are the most common and allow us to investigate potentially incorporating viewing experiences for those formats.

# Data

20-30 people a day attempt to upload a non-PDF file to Attachments as seen here: <https://analytics.amplitude.com/org/16/chart/72tjyd7/edit/rzojrom>\

# Hypothesis

Allowing all file types to be uploaded will increase the amount of files uploaded and shared via PlanGrid

# Relevant objective:

Increase average number of documents uploaded to a project

# MVP

*Minimum work required to test our hypothesis*

Enable the upload of:

## Images

* .png
* .jpeg
* .jpg
* .gif
* .bmp

## Spreadsheets

* .xlsx
* .xls
* .csv

## Word/Open Doc documents

* .rtf
* .doc
* .docx
* .txt
* .odf

## Videos

* Webm
* mpg
* MOV
* F4v
* Avi
* Mp4
* Wmv

## CAD files

As we look to improve to make a truly smart record set, we need to get access to the original design files that generated the 2D sheets our customers look at in PlanGrid. This unlocks a huge amount of possibilities to understand a lot more about our user’s projects, automate a lot of activities they currently need to do and provide better insights on top of their drawings.

Allowing those design files to be uploaded to our attachment section is a step in this direction as it allows us access to those files.

|  |  |  |
| --- | --- | --- |
| Revit | .rvt | A Revit Project file |
| DWG | .dwg | A modeling file usually used with AutoCAD |
| Drawing Exchange Format | .dxf | A modeling file that can be used by multiple drawing programs |
| DGN | .dgn | A modeling file used by Bentley Microstation |
| Design Web Format | .dwf, .dwfx | A modeling file with restricted information, mainly used for viewing |

## Google Earth files

KML and KMZ files are Google Earth’s file format for storing placemarks, network link information, and other information. If the Google Earth app is installed on the device, these can be opened to link directly to that location.

* KMZ
* KML

## 

# Designs

Each file type should have relevant icons on all platforms

*Attachment log*

* Web: <https://zpl.io/1TCr>
  + Assets: <https://www.dropbox.com/sh/s2h89lchs4xsb5b/AABBTBE3Szbynq5xjaEkrybpa?dl=0>
* iOS: <https://zpl.io/8nfvn>
  + Assets: <https://www.dropbox.com/sh/48nyujbao33xjqz/AADU0WPOmQkCVM8HORDjAZ1_a?dl=0>
* Android: <https://zpl.io/ZEXtb>
  + Asssets: <https://www.dropbox.com/sh/egvv1do8tvdst6w/AAAj3wH7vdNkTtxR-IggS_Ola?dl=0>

*Image viewer*

* Web: <https://zpl.io/Z1cwV55>
* iOS: <https://zpl.io/16jcd3>
* Android: <https://zpl.io/2rY7sF>

*Video player*

* Web: <https://zpl.io/Z1MhH3>
* iOS/Android: native players

Other files, including all open document formats, should render as pdfs do; centered, drop shadow.

## User journey

* Clicking on a non-PDF file
  + Web: will download it
  + Mobile clients: will try to open it outside of PlanGrid with the default OS option

## Errors

* Unrecognised file format - “That file format is not supported. Click here to see the supported formats”
* File too big - “That file exceeds our 1GB size limit”

## User stories

* As an admin I can upload any file type to PlanGrid via the website
* As a web user I can view any file type in the Attachments log, on any platform, and download it
* As a mobile user I can view any file type in the Attachments log, on any platform, and open it outside of PlanGrid

# How attachments work with RFIs

## Problem

RFIs can have attachments added to them. They expect those attachments to be PDFs which causes the following problems:

* On Web you can no longer export RFIs
* On iOS the app will crash if you try to open an RFI with non-PDF files attached
* On Android you can view RFIs with non-PDFs but the preview will think it’s a PDF

## Solution

### For users on up-to-date clients

* If a user exports an RFI with non-PDFs, we skip the non-PDFs and warn the user that they have to download the non-PDF attachments separately

### For user on old versions of iOS

* On iOS the app will crash if you try to open an RFI with non-PDF files attached. We will email all the users who have viewed an RFI on iOS in the last month that aren’t on the most uptodate version advising them to upgrade.

# Future enhancements

*Not part of the MVP but we can do after releasing it*

View other file types within PlanGrid. Easy ones to add are:

* Image viewing on Web
* Image viewing on iOS
* CSV viewing on iOS
* Doc viewing on iOS

# Analytics

Information we should have in looker:

* Distribution of total number of file and total file size per project or per user
* File type (extension) as filterable value

Events:

|  |  |  |
| --- | --- | --- |
| **Action** | **Event** | **Extra attributes** |
| Attachment view | attachment.view | type=<extension name> |
| Attachment download | attachment.download | type=<extension name> |
| Attachment upload new | attachment.upload.new | type=[list of all extensions included] |

# Remaining decisions/questions

*What do we need to decide on before we can get to work on this*

* Why are users trying to upload photos as attachments?
  + Reach out to users who’ve tried and ask them what they are doing
* File size limits? And how does that limitation interact with the above error message?
  + File size limit is 1GB
* How do we protect against malicious files?
  + From Mark: “On Windows, when we download the files we will also ensure that they are marked with InternetZone:3 attibute stream which means windows adds additional protection when the file is opened”
  + From Mark: “There were PLENTY viruses made inside a PDF, and we weren't protected. What we should do (on all clients that upload docs) is make sure the file uploaded doesn't have a PE header (https://msdn.microsoft.com/en-us/library/ms809762.aspx?f=255&MSPPError=-2147217396) basically checking if the first two characters are 'MZ' (but I can write up a more detailed explanation). Long term, I think we should offer some antivirus for our files, which would be a differentiator - we scan your files and keep your workplace safe”
* Security concerns
  + We have an external security consultant and will consult
* Business value in doing this
  + Is it circumventing sheets
* Bucket size
  + How do we track? Per user? Per organization?
    - Looker is only daily
    - Tested mechanism to monitor and alert abuse of uploading
  + How we stop this ballooning in costs?
    - Run a trail with a subset of users and see how much they upload
* [Mark] From eng review: we should store the white list of files in an API so that Windows can get to it when we update
* [Mark] From eng review: we should store the file size in the DB so we can better understand how it’s used, if we need to add pricing, etc. restrictions.

# Reference Material

* You can upload any file type on this environment: <https://plangrid-dev-pr-4160.herokuapp.com/> using the dev api
* All the PRs https://plangrid.quip.com/zvBMAV9EzL2H