# James Skelley

Technology Law - IP Prosecution - Transactions

MAIN

**UPDATES** 

LAWMUX KNOWLEDGEBASE

**SOFTWARE** 

FAO

CONTACT

**SUBSCRIBE** 

## **Expediting Portfolio Review**

② February 18, 2020 SKELJ Uncategorized Unca

It's a federal holiday and the USPTO / my clients are on vacation, so you know what that means . . . more non-billable-software-showing-off-nonsense!

If you follow me on GitHub you know I have a **thing** for improving portfolio review efficiency. One tedious artifact that's bugged me for years is having to print a patent and match reference numbers in the text with those in the image. An automated tool perfectly matching reference numbers isn't feasible, but the process by hand can take ~45 minutes, so if I could build something whose errors took less time than that to correct, then I decided I'd call it a win.

Well, over the previous X-mas holiday I finally decided to do something about it and, having time today, I migrated a demo version onto my website.

Enter RefBot. RefBot takes a patent PDF and text then spits out a GUI with automated images <-> text matches. You can find an online demo here (n.b., as with all my browser tools, they're built for FireFox - so I take no responsibility if you're using something else! Actually, I take no responsibility no matter what you're using).

Now, I run this thing ragged when I'm using it in the office because I have a local copy of the patent database. For an online demo, uploading >100GB of PDFs isn't feasible, though. Pulling PATFT is feasible, but I'm always nervous about imposing on the USPTO's servers so . . . we're not going to do that either. Instead, the demo has a drop down in the top left corner letting you select from a handful of unprocessed assets on the server that you can play with.

#### **RECENT UPDATES**

DMCA Study

**Skelley Snapshots** 

CPC Based Deal Search (i.e., not a COVID post)

#### **NAVIGATION**

Main

**Updates** 

LawMux Knowledgebase

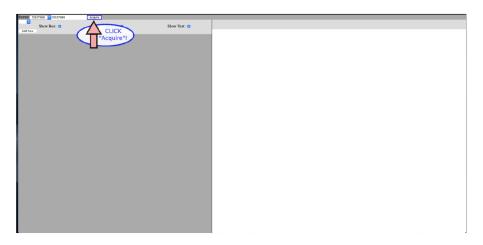
Software

FAQ

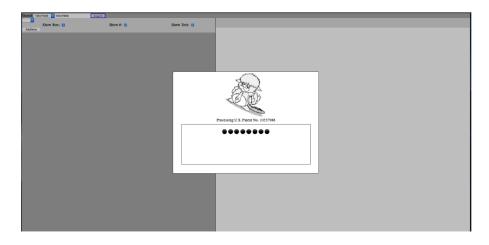
Contact

Subscribe

So if you follow that link you'll see a page like this:



On the right side is where the text references will appear and on the left is where the annotated image will appear. By default, the search field is loaded with U.S. Pat. No. 10537988 (if you select from the drop-down to the left of the search box, you can play with some of the other assets I randomly selected). Per the flashing invitation, if you click on the "Acquire" button . . .

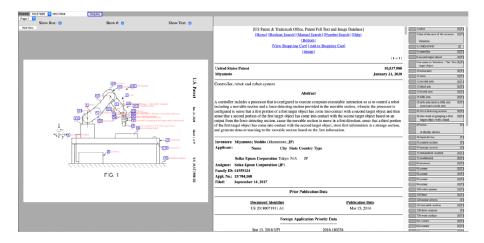


You'll be greeted with a dancing LawMux as the patent processes. Now, my local copy doesn't have these dots, instead presenting processing diagnostics (long story, but these are really useful when processing large portfolios). In my deep and profound ignorance, however, I still haven't figured out all the nuances of Flask + Passenger + Ajax on my website, so . . . you just get dots.

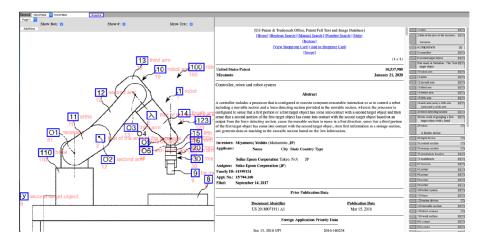
### - "\$@% it Jim, I'm a lawyer not a programmer!"

It can take almost a full minute, so be patient.

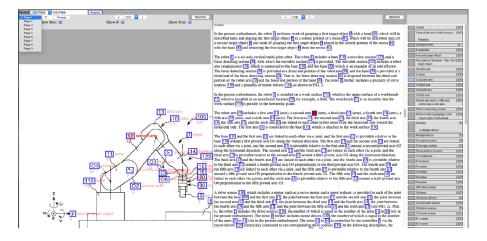
Once finished, you will see . . .



as promised, the annotated images are on the left and annotated text on the right. Again, it's not perfect – text recognition is ~85-90% (it's ~100% when all the assets are by the same author, but there's such a diversity in fonts, etc., that general recognition takes a hit). You can drag the image by clicking and dragging your mouse within it. If you have a mouse wheel you can zoom in and out by scrolling the wheel.

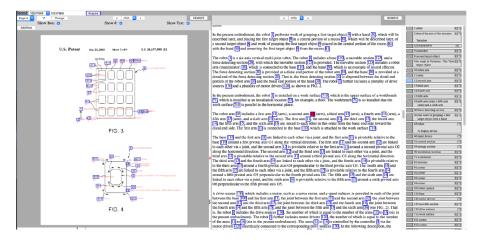


Though not shown in the above image, if you scroll down in the text, you'll see additional blue reference boxes. Clicking on a reference box either in the image or in the text will focus that reference, as shown in the below image where "12 arm" was selected. You can select other pages from the patent with the drop down in the upper left corner also as indicated in the below image:



Once a reference is selected you can iterate through its instances in the image and in the text with the left and right arrows in the top of the GUI (the "1/22" indicates this is the first instance of the 22 instances).

By selecting the asset drop down above the page drop down, you can look at other assets, e.g.:



My version of the tool also includes an editor, so that I can quickly tweak and make corrections to the machine-generated annotations. In the most extreme cases it now takes ~5-10 minutes – quite an improvement over the previous 45 minute approach.

Since using PATFT isn't feasible, I don't think I could provide the full tool online for free, since I'll need to find some (cost incurring) method for hosting the PTO dataset. Still, if you're interested, feel free to reach out and I'll see if there's a feasible way to set things up: james@jstechlaw.com.

**PREVIOUS ARTICLE** 

**NEXT ARTICLE** 

WP Template Copyright © 2016 MH Themes   Content and Supplemental Software Copyright © 2015, 2016 James Skelley No content on this website is intended, or should be construed, as legal advice. Rather, the content provided here merely serves to notify a wide audience of various issues' general character. If you have a legal question contact an attorney.