Final Project Biography: The Coding for *S H I F T*

Jessica Vandervort, December 15, 2023

Introduction:

Welcome to my final project: entity extraction for a fee-shifting research app. I have spent the semester working with my partner – myself, in 21st Century Law – and determined what would be the most helpful design for a tool to extract data from summary process cases. I have incorporated LLM work through Google Collabs, as well as tinkered with several regex codes along the way. I am thrilled to say that the Collab *works*, allowing for data to be extracted into a table, from PDF cases, far more efficiently than the time it would take for an individual to read through those cases on their own. Welcome to behind-the-scenes of *S H I F T.*

Framing and Research: *The What and the Why*

In Massachusetts, 92% of all tenants in landlord-tenant disputes, are unrepresented in court. Of the 8% who receive some kind of representation, still fewer have lawyers actually representing them in court for the duration of their case.[1] While more attorneys enter the profession every year, there is currently no real drive for these new lawyers to begin representing this unrepresented population. The goal with a Fee-shifting research app – entitled *S H I F T* – is to streamline the research process for attorneys, allowing for more efficient representation. The hope, too, is that a tool that easily demonstrates past fee-awards, would also encourage young attorneys to take on these cases more willingly, as the potential for recovering fees may be higher than they would otherwise imagine.

Currently, lawyers who take on fee-shifting cases may petition the court for an award once judgment has been issued in their favor. While at times, the fees are factored into the plaintiff's award, more often the judge will allow a ten-day, or two-week window for attorneys to file a motion for reasonable fees. The court relies on the *lodestar* method – Reasonable fees x Reasonable hourly rate = Fee Award. Judges have great discretion in determining whether the fees requested – both the rate and the hours – are reasonable, but attorneys can help bolster their case (and, indeed, must) through presenting evidence of cases in their area, where the fee awards were of a similar amount.

Unfortunately, the current options for research remain limited, expensive, and time-consuming. Tools like Lexis+, and WestLaw, not only cost a great deal of money, but also are unable to narrowly tailor searches to the degree necessary to ensure efficiency. Because attorneys in fee-award cases must find past judgments that outline hourly rates, hours worked, *and* judge's reasonings, researching these cases often means trudging through case after case to determine if the content is relevant, and if so, helpful.

---

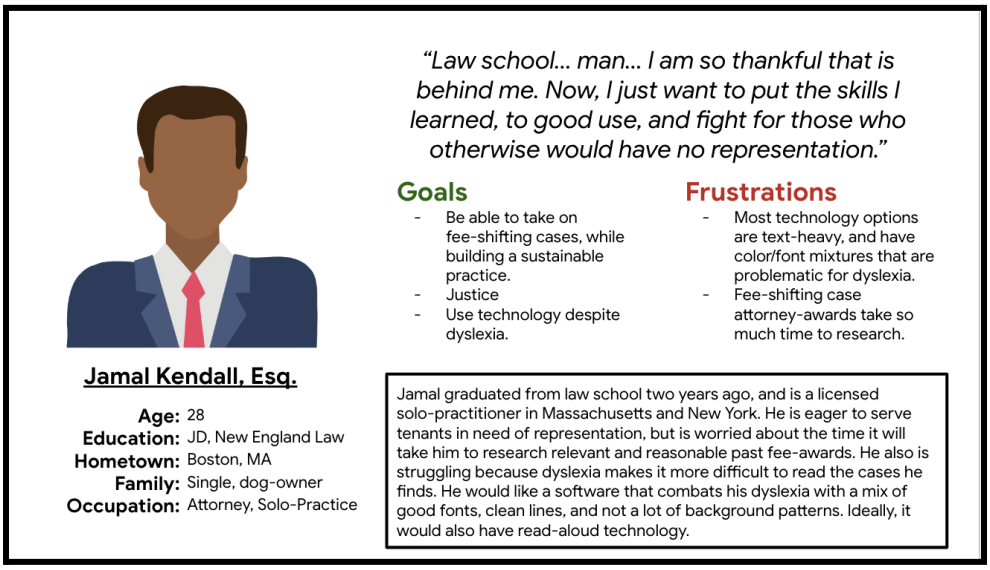
[1] Many of these individuals interact with lawyers either in pro bono clinics, on a limited representation basis, or at "Lawyer for a Day" programs at Court.

In my research, I also spoke with current housing attorneys, James Matthews, and Ed Rice. They walked me through their own process in finding cases, and also emphasized that knowing the judgments the judge's have awarded to the prevailing party in the cases, as well as the attorney's own fees, would be helpful in their use of such an app. Ideally, being able to easily access a summary of the judge's reasoning, would also be helpful.[2]

After plausible solution comes to light: An app that takes pre-collated summary process awards (and down the road, perhaps also the actual attorney motions for awards post-summary process) and allows for narrowly-tailored data extraction. This app would allow for efficient searches asking for judge and party names, the hours *and* hourly rate at which the attorney worked, the overall judgment award, and the reasoning given by the court.

This app could be used by new attorneys, paralegals, and their law student associates, at a much lower cost than Lexis, or Westlaw, with a far greater return on investment. Unlike the bigger-name apps that are juggling many different case topics and resources, the simplicity of the app's content might eventually also allow for ease in accessibility and language calibration.

Potential Personas (created for 21st Century Law)

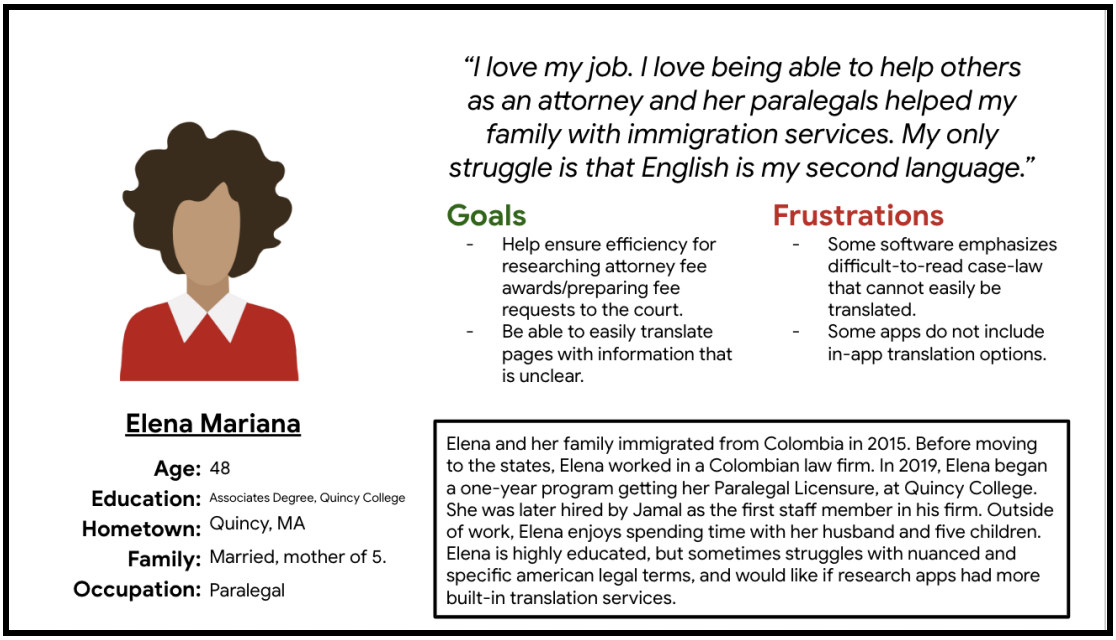


Problem Statement: Jamal is a young solo-practitioner who needs inexpensive, accessible software, because he would like to help bridge the access to justice gap for tenants, but he does not have much capital to risk on cases with no potential, and has dyslexia which makes many technology options more difficult to use.

Coding and the Law Note: My partner's wireframe design makes use of the font, Avenir, (seen in this document) which is known to help individuals with dyslexia. The wireframe design (shown below) also would incorporate high-contrast color schemes for individuals with colorblindness.

---

[2] I have incorporated both of these suggestions – the case's overall award, and the judge's reasoning, into the Collab tool.



Problem Statement: Elena is a middle-aged paralegal who needs easily translatable researching tools, because she would like to efficiently work on fee-shifting cases.

Coding and the Law Note: my partner's wireframe includes in-app translation, or links to outside translation options, as well as plain language.

But even before the creation of a full-fledged app, or the implementation of my partner's wireframe designs, the Collab already works to pull data from past cases, turning pages of text into easily discernible answers in a table. While some data such as the names of the Judges, are harder to pull from an OCR'd PDF (such as where the name of the judge is partially covered by their signature) the LLM software can save attorneys time by pulling to the front the cases that are most relevant, from which the lawyer, paralegal, or law student intern, can then examine more thoroughly. In short, *it works*. The Collab, on its own, has already begun to save my partner time in reading over cases for fee-shifting award research, even before implementing any of her app designs.

Ideation & Prototyping

For this section of my project, I was largely relying on my partner – as the ideation & prototyping were key deliverables in 21st Century Law. The partner would also be providing me with the actual data I needed to be able to extract from documents. For this reason, the reins were a bit tighter on what all I could actually ideate or prototype for the project at hand. However, I did go back and forth between my two 'personas,' considering what the best options might be given the type of data I needed to find.

I had first assumed that Regex code would be my best option. I began with some strings, and considering how to best create regular expressions for the data I needed to pull. However, I realized very

quickly that such code would be best for finding case numbers, and not much of the other data I was seeking.

While LLM's were the best option for my project, I did consider pivoting to a Document Assembly tool, which down the road I would still like to implement. This would allow for attorneys to build a motion for fees within the app, citing to the cases they have found, and inputting their hours and hourly rate. I did not work on this feature for this project, but it is something I definitely considered in the ideation process.

User Testing:

Across the semester, I worked with my 'other self' to determine best practices for use of the software. In my 21st Century hat, I walked through the Collab again and again, trying with larger samples, and longer PDFs. At different points my Coding and the Law self created alternative LLM questions that asked for various points of data such as the plaintiff's attorney name or the defendant's attorney names, or asked for 5 cases at a time, instead of 3 or 4. I also wrote in a variety of other LLM prompts that asked multiple questions in one data cell, or only two questions at a time, causing the user to need to run several different collab cells to get the data necessary.

My 21st Century user, tested each of these and in some ran into errors. In those earlier walk-throughs necessary deliverables were not always met. Sometimes the judge's name would not come through at all, or the plaintiff's attorney would be labeled as the defendant's attorney. Eventually, my user let me know that it would be more helpful to gain information surrounding the judgment, than the actual individual attorneys (facts that could be discovered after a case has been determined to be helpful). As such, this was one of the biggest changes that I made for my user, in removing the question of attorney altogether, instead looking for the party names, the judge, and the various pieces of information that my user found most helpful.

I also walked through the Collab with several others – two of my classmates, in particular – and through watching them use the software I determined that an instruction document was needed, to outline some of those specifics. Namely, if documents have been added to a specific drive for entity extraction, the drive must be refreshed before the Collab will recognize that document in it. Additionally, longer documents (greater than 9 pages) are not supported by the software, and searches can only be made every two minutes, or will return an error. These factors are all potential stumbling blocks that outside users would face, but my key stakeholder – my user in 21st Century Law – was still happy with the results – namely, that the data-pulling happens at a far more efficient rate than a Lexis+ search, or a human read-through of the case, would allow.

Refinement:

As mentioned above, I began to integrate my user feedback in as soon as I received it. After my partner (me) tested the Collab I updated the prompts and outputs accordingly, to achieve the key deliverables necessary for the Collab to reach MVP stage. After outside testing, I drafted an instruction

document, to better assist any potential other users. I also paid for an OpenAI subscription while running the project, so that I could achieve higher levels of searches with more data extraction.

Product Intro: *The Pitch*

Though the design of the app wireframe that my partner (other 'me') created, has certainly changed over time, my pitch in-class demonstrated some early mock-up concepts of what the data usage could amount to, who the stakeholders could be, and the overall 'why' of the project. It also addressed the various alternatives that attorneys currently have in researching fee-awards for fee-shifting cases.

Slidedeck:

The Problem:

Attorneys → Fee-Shifting → Reasonable Fees

My Client...

It's like looking in a mirror...

- UX Design for the Class.
- Seeking Specific Factors.
- Creating a Wireframe.

21st Century Law, Jessica

The project:

01

PROBLEM needs a SOLUTION

Attorneys seeking to prove reasonable attorneys fees need to show past cases.

02

The Project's Next Steps

Use entity extraction to pull the data from existing court cases

03

The Overarching Goal

Make fee-shifting work more feasible for the average solo-practitioner.

The Current Stumbling Blocks:

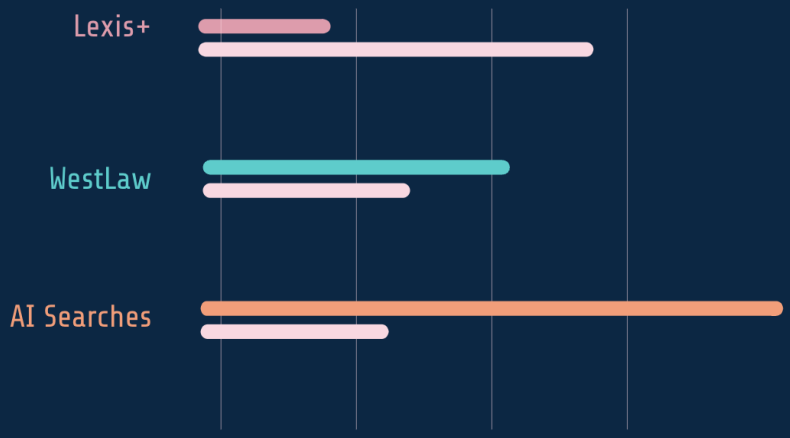
The Stakeholders:

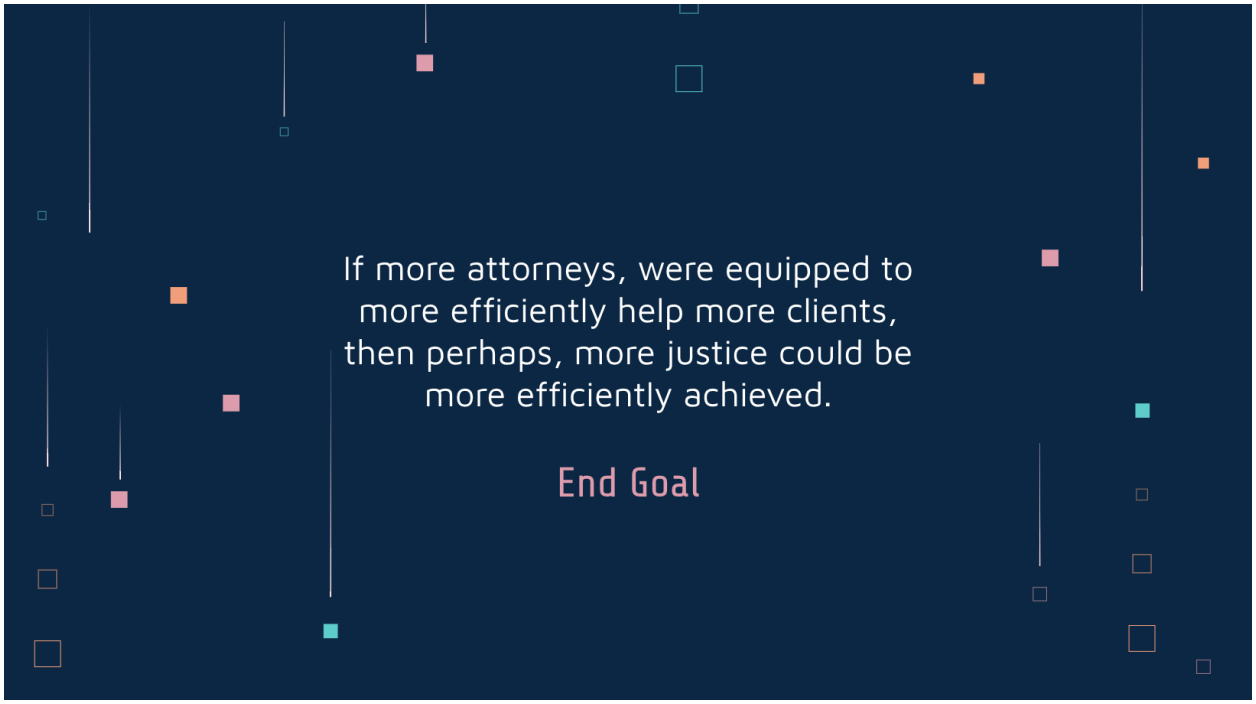
Attorneys

Staff

Court Clerks

Current Solutions:





Complexity: *The Collab*

I was definitely most concerned about this aspect of my project, as I had never used Collabs before this class, and was concerned that I would be unable to actually put them into practice. I was, therefore, very excited when I was able to have the LLM-based tools in the Collab actually put forth the data I was seeking. I developed several prompts that the Collab uses to parse through my OCR'd PDFs of summary process cases. In doing so, the LLM's datascrape the pdfs, and then provide an output of data in a CSV table, which I can then download as a spreadsheet. I also used two different Regular Expressions to parse the data as well (though I ended up only keeping one – the case number – in my final project).

Impact and Efficiency:

While I have not done the math to determine whether or not my reach has been increased by 100, I can confirm with certainty that the time extended by my partner, in searching for fee-awards, and other details in these cases, has been decreased by more than 50%. Even though parsing in the Collab cannot be performed within two minutes of each other, each run can parse through four cases at a time, returning a spreadsheet in approximately 30 seconds. This run-time dramatically reduces the amount of time necessary to determine cases that would actually be helpful for attorneys to use in their fee-award petitions. It also helps provide raw data concerning what a reasonable award, and therefore reasonable fee, might look like before an attorney even takes on a given case.

The Fit of *S H I F T*:

I am thrilled to say that my user is pleased with the end-product of this Collab. The MVP stage that it is at, allows for her to swiftly extract necessary data from OCR'd cases. The problem – namely, the amount of time it takes to research these cases – has been dramatically reduced. This impact could also be transferable to other attorneys, and is easy to implement, as the only requirement for use would be a google account.

While users would potentially have to pay a small fee (to cover the token usage in Open AI) the return on investment greatly exceeds that of other sources like Lexis or Westlaw, in actually providing helpful and accurate feedback on fee-awards and fee-shifting cases. Additionally, this could be easily integrated into the UI/UX-driven designs that my partner has wireframed in 21st Century Law.

The Documentation:

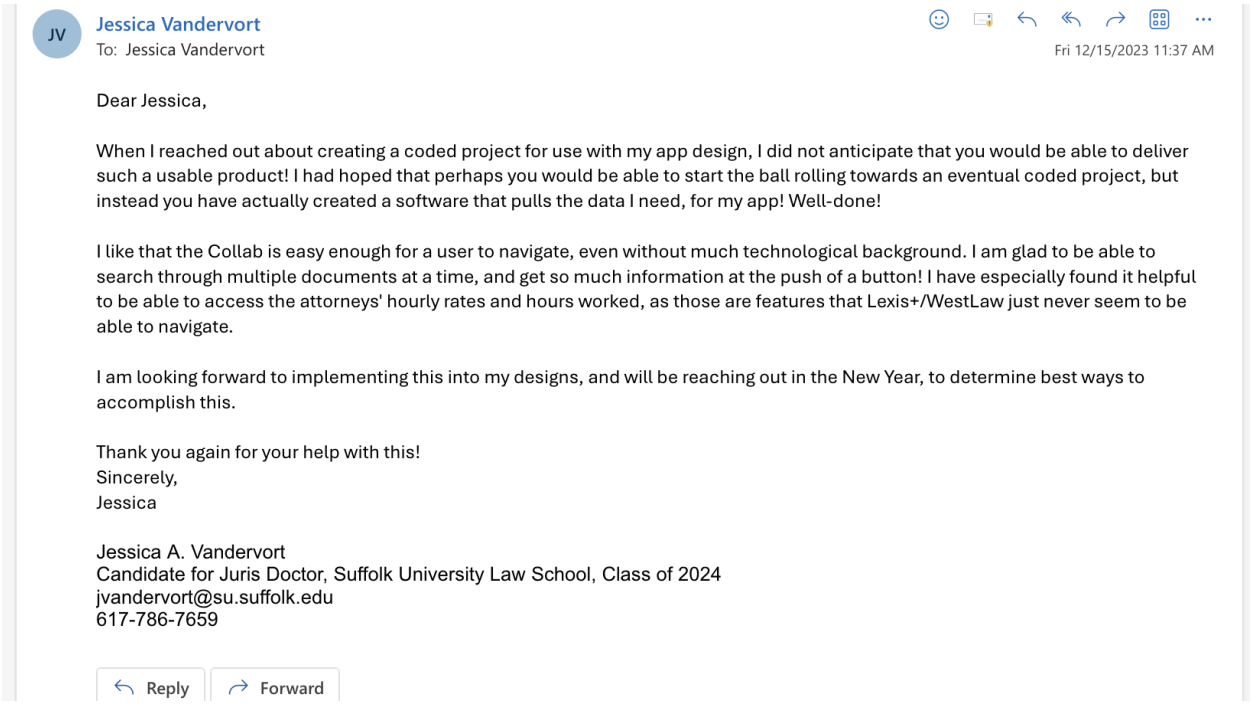
Because my key user is myself in another course, I was not as worried about the instructional documents that needed to go alongside my creation. However, I did create an instructional document which is housed in a link at the top of my Collab. It outlines the steps necessary to use the Collab, and also warns users of potential pitfalls, and ways to avoid them. That link is also here: https://docs.google.com/document/d/1nxGGDE5Hw5U8yefoi4iJRVXRZFnBGs14-DbPBr9Rhyg/edit?usp=sharing

The Real World and Sustainability:

In my real world, I have already begun to use this resource to go through cases that I was provided from a Housing Database. I have parsed through 12 different cases at this point, with more to be able to go through in the future. I have noticed that at times the judge's names or the case number are not exactly the full name or number of the case at hand. Because of this, I find that I will use this most for the *data* of the case (which remains accurate) and will then refer to cases that seem most applicable, for the details of names or full citations. This is above and beyond what I had hoped to be able to accomplish at an MVP stage. Anyone with a Google Account, and OpenAi, could use this collab with no additional technical knowledge needed. My step-by-step instructions would help them navigate the collab and allow them to download a google spreadsheet of the information they have parsed at the end.

As far as sustainability, the key factor at play is that OpenAi updates frequently. In order to sustain this project, I will need to keep up-to-date with any ChatGPT/OpenAi software updates and ensure that the most recent software is what is running the system. Additionally, there is a financial cost to this, as OpenAi tokens cost real-world money. For this reason, I have cancelled my OpenAi subscription until such time as I am using this software again, for ongoing searches.

User Email:



JV

Jessica Vandervort
To: Jessica Vandervort

Fri 12/15/2023 11:37 AM

Dear Jessica,

When I reached out about creating a coded project for use with my app design, I did not anticipate that you would be able to deliver such a usable product! I had hoped that perhaps you would be able to start the ball rolling towards an eventual coded project, but instead you have actually created a software that pulls the data I need, for my app! Well-done!

I like that the Collab is easy enough for a user to navigate, even without much technological background. I am glad to be able to search through multiple documents at a time, and get so much information at the push of a button! I have especially found it helpful to be able to access the attorneys' hourly rates and hours worked, as those are features that Lexis+/WestLaw just never seem to be able to navigate.

I am looking forward to implementing this into my designs, and will be reaching out in the New Year, to determine best ways to accomplish this.

Thank you again for your help with this!
Sincerely,
Jessica

Jessica A. Vandervort
Candidate for Juris Doctor, Suffolk University Law School, Class of 2024
jvandervort@su.suffolk.edu
617-786-7659

Reply Forward

Link to Collab:
https://colab.research.google.com/drive/1Vrtn_rXM4sUD9buoMYBRxwwlQ1kBRjw4?usp=sharing

Link to Folder of OCR'd Files Used:
https://drive.google.com/drive/folders/1kQAcTDRaudozNJZljIUbaFToVbxDZm0l?usp=sharing