## Tell me about your proudest professional achievement. It can also be a personal or school project.

I’m proud that I was able to contribute to all large and small successful projects that I worked on over the years. But probably one project I would be proud of because it was the most challenging one that I worked with. I worked in KPMG with the SOLR search engine. Among other things SOLR ingested and indexed a lot of not structured documents like MS Word, text, and PDF available for search. The user was able to search those documents by certain terms. At one point I was presented with an assignment to parse thousands of legal contracts in PDF and extract from them few pieces of information such as the name of the company that KPMG has contracted with, the address of this company, and dates of beginning and the end of the contract. First, I search on the internet if any open source or commercial solutions were available for such a task but found nothing. I started reading few randomly selected contracts and found that there is no standard or even common pattern for specifying the information I needed. The most challenging part is that I have to start from scratch without any available solution. My first approach was to hopefully use the neural network or some other machine learning (ML) software to extract data from contracts. However, to use the ML system, first I need to train it. I had to tag particular parts of the document that can be used for training. When I keep analyzing documents, I found a certain pattern that I can use for tagging. Using a custom build program using natural language processing libraries I was able to simplify my targeted search splitting into the sentences converting them into lower case removing stop words. After that, I used stepwise refinement search applying regular expressions extracting names of contra parties, another set of patterns extracting addresses and dates. Eventually, the program was designed as a decision tree of regular expression patterns from most common to more specifically used to extract all terms that I need. After converting PDF file to text format and parsing those documents I was able to process 80% of documents that was acceptable to users. 20% of documents after converting from PDF to text files were humanly unreadable because they contained non-readable characters inside the words, some words were just unreadable, etc. However, an 80% success rate was acceptable by users because it dramatically reduce the time to search those contracts.

## 2. Tell me about something you have read recently that you would recommend and why. (Can be a Github Repo, Article, Blog, Book, etc)

It’s not recent but I would recommend the following books:

Modern Java in Action: <https://smile.amazon.com/Modern-Java-Action-functional-programming/dp/1617293563/ref=sr_1_3?dchild=1&keywords=modern+java+in+action&qid=1634662828&qsid=139-2122694-1714111&sr=8-3&sres=1617293563%2CB08288LT6T%2C1617291471%2C1617294543%2C0134685997%2C1617296023%2C1617294691%2CB079MD426Z%2C1449370772%2C1617291994%2C149197317X%2C1617290769%2C1617292737%2C1937785467%2CB008VRFNOC%2C1484242777&srpt=ABIS_BOOK>

# Effective Java 3rd Edition: <https://smile.amazon.com/Effective-Java-Joshua-Bloch/dp/0134685997/ref=pd_bxgy_img_1/139-2122694-1714111?pd_rd_w=IX7DX&pf_rd_p=c64372fa-c41c-422e-990d-9e034f73989b&pf_rd_r=9VWBCW42ZPF7ZHH4PCMB&pd_rd_r=ece688ae-aabd-418b-be7e-93573f47932e&pd_rd_wg=EaibR&pd_rd_i=0134685997&psc=1>

## 3. How would you explain to your grandmother what Availity does?

Gramma. If you go to the availity.com site from your phone or computer you can check your coverage, benefits, name of the medical plan including Medicare, coverage date, member id (that you can tell your doctor billing office). It allows doctor office and hospitals access your information from your insurance companies and assist them to pay your medical claims.