

Artificial Intelligence vs. Human in the Legal Profession

AI vs. Legal Profession

Yonatan Aumann, Professor in the Department of Computer Science, Bar Ilan University

May 10th 2018

LR 2018, Page 74 to 76 (overall 3 Pages)

Much has been written about AI overtaking humans at diverse professions, ranging from [health](#) to [education](#). At first, the legal profession may seem more immune to this trend, as it is notorious for using complex language, incomprehensible to all but legal experts. Recently, however, it is becoming clear that AI is poised to make headway in this profession as well. Several [studies](#) have considered the potential areas in which AI technology can be beneficial and how this could [transform the role of lawyers](#) in delivering legal services. 1

While it is apparent that AI will be able to perform some legal tasks, it is universally assumed that the quality of the work will be inferior to that of humans, especially when compared to professional lawyers. Training, experience, professionalism, and tradition have convinced lawyers - and many of their clients — that manual labor ensures the best outcomes. But is it really so? Do trained lawyers necessarily outperform machines? 2

1. The Design

A startup company for which I consult (LawGeex), tested this presumption in a [new study](#). The study was designed in collaboration with top legal academics and professionals, including: [Erika J.S. Buell](#) of Duke,¹ [Gillian K. Hadfield](#) of USC,² and [Bruce Mann](#) of Morrison Foerster, LLP.³ [Christopher Ray](#), an independent lawyer and consultant,⁴ administered the test. The basic study design was as follows. 3

¹ Director of the program in Law & Entrepreneurship at Duke Law

² The Richard L. and Antoinette Kirtland professor of law and economics at the University of Southern California

³ Former senior partner at Morrison Foerster, LLP. Recognized with a Lifetime Achievement Award as one of the top corporate lawyers in America; has handled more than 300 IPOs and over 200 M&A transactions

⁴ Graduate with distinction from Suffolk University Law School, licensed to practice law in Massachusetts and New Hampshire

The study was designed to compare the performance of the AI engine with that of humans at this core task of "issue spotting." Focusing on this well-defined task allowed the design of a rigorous, well defined study. [Erika J.S. Buell](#), Director of the Program in Law & Entrepreneurship at Duke Law, and one of the study's leaders, said "the issue spotting test is a solid way to compare the performance of a lawyer with that of an AI review." 4

2. The Structure

The study focused on NDAs (non-disclosure agreements), which are one of the most common legal agreements used in business. Five commercial NDAs, with a total of 153 paragraphs, were chosen for review. The five NDAs (taken from the "Enron Dataset") were real agreements from companies including Enron, InterGen, Pacific Gas and Electric Company, and Cargill. The agreements were given to the AI engine and to a group of human lawyers for review. The participants comprised 20 US-trained lawyers, with extensive legal and contract review expertise. The lawyers were paid for their participation as if the review was for a real client. Both the AI engine and the lawyers were tasked with spotting key legal concepts in each of the five NDAs. It was not enough to merely skim the agreements; as the test simulated real lawyers' work, a deeper analysis was required. 5

For example, it was not enough to identify which information was considered confidential but also what information is excluded from confidential treatment. If either the AI or a lawyer missed a relevant exception, points were deducted for accuracy. Similarly, participants lost points if they falsely identified an exception. To achieve the maximum score, each participant had to identify all and only the right topics in the right text paragraphs. The results of both the AI and the human lawyers were reviewed by the study's administrator, [Christopher Ray](#), who determined the correctness of each attribution. Accuracy measures were then computed for each of the participants. 6

3. The Results

The performance of the human lawyers varied. The top human lawyer exhibited an impressive 94.94% accuracy, but there were several that scored under 80%. The average human score was 84.84% and the median was 87.56%. The AI scored 94.55%, placing it well above the average and the median human score, and essentially at par with the top human performers (the advantage of AI over human is statistically significant – using by the Mann-Whitney's U test). On average, the lawyers completed all five NDAs in 92 minutes. The slowest lawyer took 156 minutes while the fastest finished in 51 minutes. The AI engine completed the review in 26 seconds (most of which was consumed by loading the data and model). 7

4. Conclusions

This study is, to the best of my knowledge, the first to methodologically compare AI and human performance in an actual legal-profession activity, in which the lawyer participants were highly experienced in the exact field of law they were tested on. The results clearly indicate that AI can be competitive, and in fact even better, than experienced human lawyers. In addition, AI sets objective standards that meet the highest levels of professionalism, instead of having individual standards like we have today in the legal profession. Using AI, one is sure that the highest professional standards are utilized in routine tasks.

8

I find it interesting that while all lawyers in the study were very experienced - with years of experience in precisely the domain of interest - none scored perfectly, and the average and median were well under 90%. Somehow, the legal profession accepts this level of inaccuracy on a regular basis but is still wary of incorporating AI for fear of errors. The study questions the validity of such fears. Once adopted, AI could free up lawyers to focus on the more challenging and less repetitive aspects of the legal work. [Grant Gulovsen](#), one of the lawyers participating in the study said, "Participating in this experiment really opened my eyes to how ridiculous it is for attorneys to spend their time (as well as their clients' money) creating or reviewing documents like NDAs which are so fundamentally similar to one another."

9

Adopting AI will also promote consistency in legal processes, which is a common problem in large companies. Lucy Bassli, former assistant general counsel at Microsoft, says that [she has experienced cases where](#) five paralegals, all reviewing contracts, are doing so in five very different ways. The AI, in contrast, remains consistent. The engine applies the same contract rules — pre-approved by a legal decision-maker — in each and every review.

10

5. Implications for the future of law

This study represents an important milestone in legal technology, but robots will not replace lawyers any time soon. Lawyers must and will continue to play a vital role in strategic legal work, and there are many aspects of the profession that will not be automated in the foreseeable future. The use of technology is not meant to displace lawyers, but rather to make their work more efficient – and more interesting. Personally, I am confident that AI will continue to make strides in the legal realm, and that it will ultimately be widely adopted in the profession. It was encouraging to see how the lawyers who participated in the study gained more clarity into the simple power of AI to do common legal tasks, which have been carried out manually for decades. Ultimately, lawyers embracing AI — for its accuracy, consistency, speed, and cost effectiveness — will help secure the legal professionals' relevancy, allowing them to remain competitive going forward into the next decades.

11