

From pyramid to diamond: How AI is changing legal services

VIRAT MARKANDEYA

02/02/18

Krishna Sundareshan recalls the refrain. “You can say AI, you can say a set of fancy words,” he says, “but how is it going to make my life better?” Sundareshan is the founder of Surukam, a Chennai-based startup. In May 2017, Surukam was one of six startups selected globally to participate in the incubator lab of Mishcon de Reya, a UK law firm. What he heard was from end-clients—a hard-nosed response, to be sure—but also eager to participate, considering that practitioners of the law aren’t traditionally considered at the forefront in adopting technology.

Surukam has built a proprietary software it calls ‘CruXIQ’, which does contract review, abstraction, analytics and is typically targeted at corporates. It looks at issues like obligations, rights and restrictions in contracts. With contract review, for instance, the volume to sift through can be very high; say, how a change in law affects contracts across the board, which could be as high as 50,000 contracts of a company’s employees.

As Dan Jansen, CEO of NextLaw labs, a collaborative platform launched by the global law firm Dentons, had put it in the *ABA Journal*: “Law firms have historically had a pyramid structure that technology is evolving into a diamond.” What Jansen alludes to is how legal automation is narrowing the bottom of the pyramid, with technology chipping away at the need for human intervention.

Given the changing form of the traditional legal structure, Jansen’s concern is, to put it mildly, understandable. “If the work at the bottom of the pyramid is being automated, we want to own that technology and not be a victim of it,” he says.

In India, too, artificial intelligence (AI) applied to legal work brings up similar concerns. Sundareshan acknowledges this shift when he says, “[Law] Firms know they have to change”. Anand and Anand, Cyril Amarchand Mangaldas and Nishith Desai Associates are among those top Indian firms that have quietly adopted AI.

Moving forward

Firms have to move beyond billable hours, the legal scholar Dr Shamnad Basheer indicated in a phone call. At another level, there is room for a more commoditised digital front-facing service model even with less sophisticated technology.

Beyond law firms, the legal system is in dire need of intervention with the glacial pace of case resolution hurting the public and businesses alike. Nearly 27 million and 4 million cases are pending in the district and the high courts, respectively. A whopping Rs

8,00,000 crore is locked in tax cases alone in India, on which the government pays some Rs 5,000 crore as interest. Big tech companies have automation platforms to offer but startups building AI solutions believe platforms are going to go the Enterprise Resource Planning software way and yield low-value returns on investment. The real value lies elsewhere.

A squabble of startups

Early last year, the Indian corporate firm Cyril Amarchand Mangaldas entered into an agreement with Kira Systems, a Canadian company. Kira was one of the first legal AI startups off the blocks globally. It won a contract from Deloitte that set the ball rolling in 2016, for compliance with IFRS 16, a global accounting standard which deals with the treatment of leases across businesses. “Kira is the guy to beat in this space,” says Sundareshan.

At least half a dozen startups in legal AI have come out in India in the second half of 2017. Several are led by serial entrepreneurs, whose friends may be younger lawyers and they clearly glimpse the technology gap in the legal profession, its inefficiencies. Others have law firms and enterprises abroad as clients. A subset of the broader legal services startups, they all cater to specific tasks in the legal process.

In India, there are several low-hanging areas where technology is an easy differentiator. Basic legal research usually involves only text-based search on websites such as Manupatra and India Kanoon, which is where natural language processing-based engines can come (though there are costlier options from places like LexisNexis). Intellectual Property Rights law is a clear entry point for a legal research startup. As an indicator, trademarks registered in India grew over 30% in 2015-16 from the previous year.

There is a level of consolidation in ease-of-access waiting to happen. For example, though Indian court judgements are in the public domain, they are spread across twenty-odd websites. While lawyers pay for a commercial database, there isn't a good way to find case law for citizens. For legal startups, it is helpful too that several legal process outsourcing (LPO) providers are located in India. These are potential customers as there is an overlap between the kind of tasks LPOs do and the early applications of AI.

This is just the tip of the iceberg, says Sundareshan. Outside India, there is already an onrush to applying algorithms in delivering legal services. The US legal services market was estimated at \$290 billion, while the United Kingdom was at \$45 billion in 2016, out of a \$600 billion market. While recent figures for India are harder to get; audit firm KPMG estimated the legal services market to be \$6.8 billion in 2012.

Backstory: data deluge

Let's take a step back and look at what just happened. The legal profession has traditionally been averse to adopting technology. Susskind, the influential UK lawyer, mentions that many in the profession thought emailing legal documents would not come to pass.

But then, more abstractly, the legal profession is about documents or other material like emails that can be legally relevant as in the process of discovery. And there is a deluge now. Terabytes of data, cross-border legislation. Not to mention that we are in post-desktop publishing era.

Where there is a massive volume of data, says Milind PM, who heads technology at Nishith Desai Associates, that is where machine learning, the core technology that drives AI, comes into its own.

The successes of machine learning are well known: controlling driverless cars, beating the best human champion at Go, and at *Jeopardy!*. Could it step in and help general counsels manage legal risks and allow firms to do more?

The whole game is to find relevance, says Dipanker Bandhyopadyay, co-founder at Riverus Technology. Under the broad rubric of legal analytics, Riverus analyses public data for insights and focusses on tax tribunals and the Competition Commission at present.

The company has done preliminary work like plotting the timeline of cases disposed of at the Income Tax Appellate Tribunal, Mumbai, under similar sections related to deductions under the IT Act. In doing so, it figured cases relating to a particular tax incentive were taking especially long. Such granularity can provide inputs for clients, or for policy interventions.

Where's Watson?

For the longest time, IBM's supercomputer Watson was touted as an intelligent solution to problems of big and unstructured data. But examine a discussion between Watson executives and lawyers in October 2015 and it becomes apparent that ready solutions from Watson would take a while, at least in the legal domain.

Watson is not plug-and-play technology, says Ankit Yadav, who handles technology at Mike AI, a Delhi-based startup providing a research tool for patent law. Mike, named after the popular character in the American TV serial *Suits*, says it builds on Watson's natural language processing (NLP) tools.

NLP allows searching for conceptual relations, rather than searching text terms with the Boolean operators ("and", "or" as with Google). One can ask specific queries in spoken English such as "are pleas available for revocation of patents also available as a defence in a suit for infringement?" Or ask it to display cases where "patentees show no inventive step".

Yadav thinks the larger players like IBM are more interested in building foundational technology, which can then be developed further by others into a product. "They've provided the Lego blocks," he says.

But someone needs to create the datasets on which to train the system, create a corpus, and there is a host of other issues like implementation. "Everyone wants to solve this

sexy, cool part of AI,” complains Sundareshan, “nobody wants to do the grunt work on challenges in AI.”

Sheer volume

An IBM team a couple of years ago estimated that the data created every day was equivalent to 340 newspapers delivered to every man, woman and child on the Earth. Within 10 days, it would be 3,400 newspapers.

“Any AI product needs good training data,” agrees Mayur Mundra of AnviLegal, also in the contract space. Not only that, it needs different types of training. For AnviLegal, which started in Hyderabad, the main focus is Indian law firms, particularly for due diligence during M&A deals, though it said it was in talks with a Mexican law firm as well.

Training data poses its challenges. Riverus, which did a pilot project for a client on 250 orders passed by the Competition Commission of India, related to merger control proceedings, found that it needed more data to predict outcomes better. The data download, feature engineering and modelling required were time and attention consuming, and eventually, it decided to move on.

The Armstrong argument

Despite these challenges, legal AI startups clearly see an opening for themselves. To understand why, it might be worth turning to Kira again, in what can be called *The Armstrong Argument*.

Noah Waisberg of Kira Systems argues that though IBM may have 100x the R&D funding of all the legal startups combined, one could not assume it would succeed in legal contract review.

He sets it up by enumerating the legal problems that technology could solve, posed as questions. For example: Which of these 1 million documents are relevant to determining if anti-competitive behaviour occurred in this specific case? Which of these contracts have a change of control or exclusivity clauses? Is it illegal for individuals to have ferrets in the state of California?

Out of seven types of questions, corresponding to categories of legal AI companies, Waisberg writes that Watson was the go-to technology in only the “legal research-y” one. To assume IBM would succeed in other legal domains, he writes, would be akin to arguing that because Lance Armstrong had come off winning seven consecutive Tour De Frances, he would win the 2006 New York Marathon. (Armstrong finished 856th according to wiki).

“Machine learning is a new and glorified name for statistics,” says Harsh Gupta, who has worked with the Centre for Internet and Society and looked at topics like machine bias. The picture Waisberg presents includes areas in such a state of flux that even knowledge of what the optimal statistical techniques are isn’t clear.

Everyone and their uncle

Yet, everyone is jumping in. At an earnings call in April 2017, Vishal Sikka, the former CEO of Infosys, was asked by an analyst from Cantor Fitzgerald about its AI platform Mana, (now Nia). “It is one of a kind. No, everyone and their uncle has an AI platform these days,” a report quotes Sikka as saying, “From the times when I studied AI as a student, everyone calls their toolkits AI platforms.”

Sikka then goes on to describe how Mana was used to devise a solution pertaining to Non-Disclosure Agreements for a bank in Asia, reducing the need for 10-15 lawyers, and references Infosys’ buyout of Skytree.

When going through the well-known players in AI—IBM, Deep Mind, Nia—it becomes noticeable that the aim was towards developing generalised intelligence. First, solve for intelligence and then use it to solve everything else.

But the indication today is that machine learning solutions work best when integrated into company systems—ideally when the data is within the company’s control. As products, there is an effort to market that may not be worthwhile for the large players that chase contracts in other areas, and the startups already have moved fast in areas such as law.

Client confidentiality is another issue. Kira Systems has its solution “air-gapped” from the clients’ data, having already refined it. Platforms have become commoditised, and others—like Microsoft and Amazon—have all come in.

While talking of legal AI, Surukam’s Sundareshan believes platforms are going to go the ERP way, accounting for 5-10% of contracts but not necessarily yielding a great return on investment. “The value is not going to be in the platform,” he says, “but in the design and implementation of the solution.”

Linking diamonds makes an ecosystem

Large law firms in India are aware of the potential of the technology as well as the pitfalls of not embracing it. For instance, at Nishith Desai, technology head Milind PM says he’s evaluated technologies overseas. For the system to digest what you’re doing and act at an intelligent level, he says, it takes eight to nine months. For now, the firm is building its own internal due diligence system.

The change people like Milind already see is in the eroding hierarchy of firms. “It has become a network organisation,” says Milind. “Now you can flow around in the workflow.”

Law firms prefer in-house development, concurs Tony Joseph, the CEO of Cognitive Computer Services (CCS) based in Trivandrum.

CCS had been working for several years with an educational company called Sherston based in the UK. They were later approached by Kennedys, a £150-million law firm specialising in insurance, and have been involved in prototype development for them. The prototyping would allow insurance claims to be processed faster, escalating higher

claims to associates. Newer versions of the system will allow for semantic indexing to be able to search how a claim has gone and detect fraud.

Another reason law firms abroad and in India prefer in-house development is confidentiality of client data. We're speaking of occasions where confidentiality is so high that code words like Project Tiger are given as a means of referring to an ongoing deal, Milind says, adding that if one unsecured file on an FTP server gets disclosed, "it's the whole deal falling apart".

Away from the cloak and dagger of the high-stakes deal and large firms, there is a large pool of Indian lawyers that need basic access to technology. Bandhyopadhyay of Riverus, for instance, is keen on etching out a larger ecosystem that connects these remotely placed lawyers. (AnvilLegal, on its part, already has a slew of products targeted at mid-tier law firms, and is seeking to develop Hindi and Spanish versions, and Mike AI, too, wants to build an ecosystem.)

Unless one attempts to reach the six lakh verified lawyers out there, Bandhyopadhyay argues, it would not usher in a real change. The idea is to "reskill—provide the tools, infrastructure, and marketing—to disrupt the supply side".

Riverus eventually hopes to provide a freely available platform to lawyers and CAs. Can the sector discipline itself with technology and break its information asymmetry?

Clarification: Surukam was misspelt as Sukuram in one paragraph. This has been corrected. We regret the error.