

**C L I F F O R D**

**C H A N C E**



**ARTIFICIAL  
INTELLIGENCE AND  
THE FUTURE FOR  
LEGAL SERVICES**



**— THOUGHT LEADERSHIP**

NOVEMBER 2017



## ARTIFICIAL INTELLIGENCE AND THE FUTURE FOR LEGAL SERVICES

Artificial intelligence (AI) is growing at a phenomenal speed (think, for example, of driverless cars, or, simply, interaction with the personal assistant on your smartphone) and is now set to transform the legal industry by mining documents, reviewing and creating contracts, raising red flags and performing due diligence. We are enthusiastic early adopters of AI and other advanced technology tools to enable us to deliver a better service to our clients. In this paper, we look at the current state of the market and explore how things will develop in the coming months and years.

### NATURAL LANGUAGE PROCESSING

This technology is able to comprehend language in its natural form (whether that is a legal contract or a spoken question). Early attempts to program a computer to understand language involved a series of rules. While this is fine for some basic concepts, it becomes complicated as you cater for exceptions to the rules. Increasing computer power means that instead of trying to codify thought processes as rules, the latest machine learning tools use statistical pattern recognition techniques to create their own rules (known as predictive algorithms) from large volumes of examples. For example, if you show such a system a collection of documents and their translations into another language, the system can determine the statistical patterns between the documents and work out how to translate from one language to the other, without having to understand what the individual words mean or the underlying rules of grammar.

#### Ediscovery tools

Ediscovery software tools have been available for some time to help legal teams with document management and review. The current generation of ediscovery software includes machine learning functionality that enables technology-assisted review (also known as predictive coding). This takes place in the context of litigation or investigations and involves the analysis of large collections of electronically stored information, to determine which documents are responsive to a particular issue.

The system is exposed to a training set (a sample of documents that has been reviewed by an experienced lawyer or subject matter expert) and it develops a preliminary algorithm based on the expert's decision about the relevance of the documents. This algorithm is then applied to further documents and, through an iterative process where the system's coding decisions are subject to human review, the system is further trained until its results reach a statistically acceptable level of accuracy. The final algorithm is then applied to the entire population of documents to identify, or prioritise, relevant documents.

Studies have shown that this approach is more accurate and cost-effective than the traditional approach (which involves having paralegals carry out the initial review of all the documents). This technology, combined with other ediscovery techniques (for example, de-duplication and tools that group email threads together), can substantially decrease the number of documents for review or, alternatively, help find the most relevant documents in the fastest way possible.

While predictive coding tools do not acquire "knowledge" that can be transferred between matters, they are still useful in situations (including litigation, regulatory investigations and antitrust filings), where significant numbers of documents need to be assessed.

#### Contract review tools

A second wave of machine learning tools has recently emerged. These have two key differences from the ediscovery tools: they do acquire "knowledge" that can be transferred between matters and they operate at the clause level within the

#### Case study – a matter before the US Department of Justice

We used technology-assisted review to comply with a Request for Additional Information and Documentary Materials (also known as a "second request") in a matter before the Department of Justice in the United States. Our process and methodology was approved by the Department of Justice. After initially collecting over 11.5 million documents, through technology-assisted and other ediscovery tools we were able to cut down the total number of documents for review to approximately 1 million. This saved the client significant costs in terms of hosting the documents, the number of contract attorneys required and the time it took to conduct the review.

document, rather than just at the document level. As such, they are able to identify certain types of clause and extract information from documents.

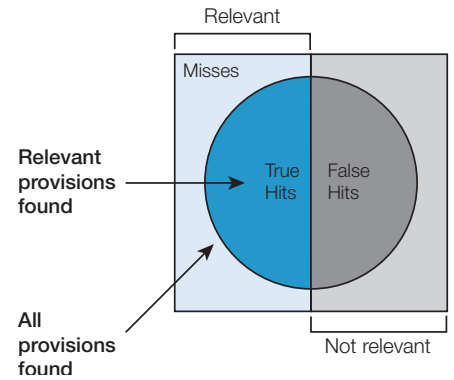
We have been working with leading AI company, Kira Systems, to develop and use one of these tools called Kira. We chose this system after more than a year of due diligence and piloting, because, while it already has some “knowledge” out of the box, it can be “trained” by our lawyers to learn from our specialist expertise, enabling it to deliver more effectively against our clients’ specific requirements.

Tools like these can make two different types of mistake: false negatives – that is to say, missing something that is in fact there and false positives, indicating that they have found something that is not actually correct. False negatives are measured using the “recall” statistic (the proportion of all relevant provisions actually found) and our testing has concluded that tools like these can

achieve a recall rate of about 90%. False positives are measured using the “precision” statistic (the proportion of found provisions that are relevant).

There is generally a compromise between recall and precision, so these systems can, sometimes, generate a number of false positives that need to be reviewed and excluded. It is worth noting that human reviewers tend to be “tuned” in favour of high precision (in other words, if a human reviewer says that a clause is, for example, a “change of control” clause, then it most probably is, but there is a risk that the rushed or tired human reviewer misses such a provision when reading through).

Given the 90% recall figure, there clearly is a trade-off to be made between accuracy (in other words, the risk of missing something important) and efficiency. Depending on a client’s particular risk appetite for a transaction (or for that subgroup of documents on that transaction), we envisage a spectrum of options:



Approach	Approximate efficiency saving from ‘artificial intelligence’	Risks
Conduct a ‘human review’ of the documents (in the normal way) and use the tool to cross-check findings	0%	Reduced due to cross-checking
Use the tool as a first pass and then have lawyers carry out some degree of quality control	20-50% depending on the seniority of the lawyers involved and the extent of quality control	Not all documents may be reviewed by a lawyer so some relevant provisions may be missed (although accuracy should be higher than 90%)
Rely on the tool to locate relevant clauses and only apply human review to ‘tidy up’ the false positives	90%	Accuracy is only 90% so some relevant provisions will be missed

At the moment, we are tending to take the first approach outlined above (using the tool and human reviewers in parallel). There is academic evidence that the combination of human reviewers and systems like these is more accurate than either operating alone. While this means that we are not yet fully exploiting the potential savings of the artificial intelligence engine, there are some immediate benefits including: optical character recognition, searching across documents, tagging, quickly identifying duplicate documents,

workflow automation and using automatic language detection.

A logical next step will be to use tools like these to analyse which sections of the data room should be prioritised for review. This approach might bring efficiency savings if it helps identify any “deal breakers” earlier in the process.

As we (and our clients) gain more experience of these tools, we would expect a more nuanced approach to





### Case study: Kira

We have recently used Kira to facilitate the due diligence process for an acquisition involving more than 4,000 documents, across several jurisdictions. Although the legal team checked the findings of Kira manually, they concluded that using the system saved them about 20% of the time it would otherwise have taken.

develop: on the most important documents, you apply AI with a human check (to ensure the highest accuracy) and on the less important documents, you just apply the AI software (and accept a slightly lower accuracy). This would mirror the approach we have seen taken with predictive coding on ediscovery matters, but requires a change in mindset not to think of due diligence as “looking for a needle in a haystack” but rather as a risk-based analytical process.

Tools like these are at their most useful when applied to relatively large data sets. Future use cases within the context of an individual transaction may therefore involve expanding the population of documents that are subject to review, to include those that might not otherwise have been reviewed. Some of the use cases we plan to explore are:

- using a tool to examine all the documents on a transaction in situations where it might previously have been cost-effective only to review a sample of them (for example, individual leases on the acquisition of a property portfolio). In this use case, we would gain a 90% understanding of 100% of the documents, rather than a 100% understanding of, say, 30% of the documents
- using a tool to examine the purchaser's own contracts (in addition to those of the target) to look for pricing and other synergies.

The other way to increase the population of documents is to apply tools like these across several transactions. This allows us to train the system to recognise new provisions that may be particularly relevant to our clients. It is important that any such training is carried out by senior lawyers with relevant expertise (and not outsourced to paralegals, etc).

### Legal research tools

A slightly different application of natural language processing is to create an intelligent legal research tool, which can accept queries in ordinary language, identify the crux of the question and present the answer (or the most relevant answers) back to the enquirer. This will save time and will make future results even more accurate.

## AUTOMATION

Automation refers to technologies that use rules to carry out tasks. Most of these systems are based on decision trees; a type of flowchart that poses a series of questions, the answers to which determine which branch is followed, until there are no more questions and a conclusion (or decision) has been reached. The decision tree can be created by a lawyer or derived algorithmically by a computer based on training data. These systems tend to be used for either giving advice or drafting documents.

### Advice systems

Early examples of these tools include our Cross Border Acquisitions Guide and Cross Border Financing Guide, which offer a concise and practical overview of the legal and market developments in jurisdictions around the world. Users can quickly and easily create a tailored report relevant to their transactions. They can use the guides' comparative tables to assess the inconsistencies between the laws of the countries involved and find fast answers to specific queries. This enables clients to undertake an initial assessment of the feasibility of a proposed transaction before instructing external counsel to undertake a detailed legal analysis, or to sense-check legal advice already obtained. As the system can be used and accessed by the client directly, it can significantly cut costs.

Recently we entered into partnership with Neota Logic, a company that uses decision tree technology that can accept large amounts of data and provide weighted options for lawyers to take next steps. We have applied it to the impact of regulatory rule changes on financial institutions. For example, we have built a MiFID2 and MiFIR Client Documentation Toolkit, which provides an analysis of the direct requirements and indirect implications of the level 1 and 2 MiFID2 regulations and UK FCA implementing rules on client-facing documentation. The toolkit is capable of being filtered by client type, business type, activity type and theme.

### Drafting systems

Automated drafting tools are becoming increasingly sophisticated. With Clifford Chance Dr@ft, we offer an automated

document assembly system, which is based on the Contract Express software. It allows clients to generate quickly and independently tailor-made and house-styled documents within the secure Clifford Chance Dr@ft private cloud. This builds on our internal use of Contract Express to offer our lawyers an efficient approach to document drafting that removes much of the routine work, giving our lawyers additional time to focus on the more complex aspects. Our use of document automation is extensive, with a high proportion of our most used templates automated and all offices in the Clifford Chance network taking advantage of the technology.

Rather than working from model documents or precedents, users simply answer a single online questionnaire, from which the system generates one or more documents. Each document generated is an output of multiple possible text variations (which themselves are dependent on the answers to the questions), rendering a tailor-made result. Clifford Chance Dr@ft is suitable for many types of document, including: loan documents; litigation documents; sales contracts; service contracts; supply contracts; share purchase agreements; corporate housekeeping documentation; and HR documents as well as other documents drafted on the basis of models or precedents. These can be either models/precedents that are already available, or models/precedents that are drafted by us.

Clifford Chance Dr@ft also offers workflow functionality (internally at the client or in conjunction with Clifford Chance lawyers) for the review and approval of generated documents. Clifford Chance Dr@ft adds value to our clients' documentation process by:

- saving time: approximately 50% per document or, in the case of a suite of documents, up to 85% (depending on the type of document or combination of documents). Information entered for the purpose of a document can be re-used at a later stage when creating one or more other related documents (for example, an amendment agreement)
- improving quality and consistency: related generated documents are

based on the same "mother template" and corresponding questionnaire, whereby legal and factual checks and balances are built in using warnings and/or preventing irrelevant options from being offered. In this way, they provide a high level of consistency, resulting in a high-quality first draft

- decreasing the burden on the legal and/or compliance departments and therefore a reduction of internal costs. Legal content and knowledge are embedded in the questionnaire and templates themselves, thereby reducing the costs of review and backup assistance of these departments
- improving efficiency in updates: an automated template can encompass many variations that previously required separate model documents. As a result, updates in texts that are the same for all those separate documents need only to be made once.

## WHAT'S THE FUTURE?

This is a fast-moving area and it is difficult to predict with certainty how the technologies will evolve over the next few months, let alone over the next few years. However, these systems break the link between the cost of providing the service (once built, the marginal cost is relatively low) and the price charged to clients. These systems can, therefore, help provide a charging mechanism that is more closely aligned to the value received by the client than the current billable hour model that is often used for legal services.

### Natural language processing

We can see a third wave of contract review tools on the horizon that can not only extract information from contracts but analyse them in context (for example, not just identify a "change of control" clause but determine if it will be triggered by the factual circumstances of the transaction). These sorts of systems might, for example, also be able to keep company policies up-to-date to reflect changes in laws and regulation.

This may be facilitated by the emergence of deep learning, a type of machine learning based on large neural network architectures (designed to mimic the structure of the human brain). These algorithms are obtaining striking results across disparate areas such as machine





translation, image recognition and drug discovery. After being out of favour for decades, the availability of vast amounts of data and computational resources have enabled the resurgence of neural networks. This highlights an important hurdle for the application of deep learning to the legal sector, which is characterised by an absence of large-scale training datasets. However, given the fast pace at which the field is advancing, it is likely that new techniques may be developed to mitigate the need for such large training datasets. Indeed, one of the major research topics in deep learning at the moment is transfer learning; a set of techniques that aim to leverage knowledge between tasks and domains. Transfer learning may provide techniques for leveraging general language understanding capabilities acquired from large-scale datasets such as Wikipedia, to reduce dramatically the data requirements for developing tools to understand legal documents.

#### Automation tools

Automation tools are likely to become more sophisticated and we will likely see a transition towards decision trees that are created by machine learning from training data, rather than being codified by humans. The principal challenge here is that it is difficult to verify the accuracy of a decision tree that has been derived algorithmically and so a conversation about efficiency and accuracy (similar to that relevant to contract analysis tools) will likely need to happen in this area.

The other area of automation that is likely to see growth in the near term is “robotic process automation”. This term describes software that can be programmed to operate other pieces of software through the user interface in the same way that a person would (rather than requiring a deep technical integration). Many people will already be familiar with macros in spreadsheets and these tools are, essentially, “macros on steroids” that can operate across several systems. While there are obvious back office applications, many routine legal tasks, such as downloading information from online registries and/or making online submissions and filings, may benefit from this approach.

#### Conversational interfaces and bots

Another next step is to combine natural language processing tools with automation tools to create a system that can understand questions, extract relevant information, follow a decision tree to determine what needs to be done and, potentially, produce a first draft of the necessary documents to effect the change. There is a lot of focus in this area around the creation of bots. Initial applications may be in the B2C space, to mitigate access to justice problems, but B2B applications for access to simple self help are likely to follow. We may also see things like negotiation chatbots for simple documents.

#### Conclusions

Many of the current systems essentially automate the way in which a legal task is currently carried out but they do not transform the legal approach. For example, the contract review tools described above are a solution to a current problem, but may not be necessary in the long term. In essence, they convert unstructured data (legal documents), into structured data. Many of those legal documents would have started life in a more structured form (for example, a term sheet) or, at the very least, it would have been better if the structured summary had been prepared at the same time as the contract was drafted. In the medium term, we would expect to see an increasing use of contract management systems with summary information stored as metadata (or even with the contracts themselves stored as data on blockchain platforms) which would negate the need for these contract review tools.

While it is difficult to know exactly how these systems will evolve, what is certain is that lawyers will need to become more technology-savvy so that they can advise their clients on which tools to use in which circumstances. There is also likely to be a growing role for technology experts (not necessarily lawyers) to provide advice on matters as the law firm of the future takes shape.

#### Case study – PGGM

We have developed a template for an automated asset management agreement which Dutch pension fund service provider, PGGM, is using through Clifford Chance Dr@ft. PGGM also makes use of the workflow functionality we have built into the system, enabling PGGM users to review efficiently first drafts generated by their colleagues and submitting an internally approved draft to Clifford Chance funds lawyers for final review. Where traditionally, the whole drafting process of the relevant agreements was done by external lawyers, PGGM now creates the first draft itself (which it can comfortably do with all legal checks and balances built into the system) and then submits that draft to Clifford Chance lawyers for review, thus considerably reducing costs and improving efficiency of the documentation process.

# CONTACTS

## Our artificial intelligence experts:

**Bas Boris Visser**  
**Partner**  
T: +31 20711 9624  
E: basboris.visser@cliffordchance.com

**Paul Greenwood**  
**Chief Information Officer**  
T: +44 207006 1034  
E: paul.greenwood@cliffordchance.com

**Oliver Campbell**  
**Global Head of Client Service Solutions**  
T: +44 207006 2062  
E: oliver.campbell@cliffordchance.com

**Anthony Vigneron**  
**Director – Legal Technology Solutions**  
T: +44 207006 1453  
E: anthony.vigneron@cliffordchance.com

**Felix Sanchez-Garcia**  
**Senior Manager – Data Science and Innovation**  
T: +44 207006 3337  
E: felix.sanchezgarcia@cliffordchance.com

## Our ediscovery expert:

**Clare Chalkley**  
**Litigation Support Manager**  
T: +44 207006 8721  
E: clare.chalkley@cliffordchance.com

## Our contract review tool expert:

**Tessa Bartley**  
**Senior Knowledge Operations Specialist**  
T: +44 207006 4104  
E: tessa.bartley@cliffordchance.com

## Our document automation experts:

**Thijs Lommen**  
**Senior Advisor**  
T: +31 20711 9614  
E: thijs.lommen@cliffordchance.com

**Judith Edixhoven**  
**Associate**  
T: +31 20711 9298  
E: judith.edixhoven@cliffordchance.com

**Gail Swaffield**  
**Director – Knowledge Management Systems**  
T: +44 207006 3694  
E: gail.swaffield@cliffordchance.com

All product names, trademarks and registered trademarks are the property of their respective owners. All company, product and service names used in this document are for identification purposes only. Use of these names, trademarks and brands does not imply endorsement.



# CLIFFORD CHANCE

This publication does not necessarily deal with every important topic nor cover every aspect of the topics with which it deals. It is not designed to provide legal or other advice.

[www.cliffordchance.com](http://www.cliffordchance.com)

Clifford Chance, 10 Upper Bank Street,  
London, E14 5JJ

© Clifford Chance November 2017

Clifford Chance LLP is a limited liability partnership registered in England and Wales under number OC323571  
Registered office: 10 Upper Bank Street,  
London, E14 5JJ

We use the word 'partner' to refer to a member of Clifford Chance LLP, or an employee or consultant with equivalent standing and qualifications.

If you do not wish to receive further information from Clifford Chance about events or legal developments which we believe may be of interest to you, please either send an email to [nomorecontact@cliffordchance.com](mailto:nomorecontact@cliffordchance.com) or contact our database administrator by post at Clifford Chance LLP, 10 Upper Bank Street, Canary Wharf, London E14 5JJ.

Abu Dhabi • Amsterdam • Bangkok  
Barcelona • Beijing • Brussels  
Bucharest • Casablanca • Dubai  
Düsseldorf • Frankfurt • Hong Kong  
Istanbul • London • Luxembourg  
Madrid • Milan • Moscow • Munich  
New York • Paris • Perth • Prague  
Rome • São Paulo • Seoul • Shanghai  
Singapore • Sydney • Tokyo • Warsaw  
Washington, D.C.

Clifford Chance has a co-operation agreement with Abuhimed Alsheikh Alhagbani Law Firm in Riyadh.

Clifford Chance has a best friends relationship with Redcliffe Partners in Ukraine.