

(LEGAL) LANGUAGE IN LEGALTECH

Miklós ZORKÓCZY*

LL.M., lecturer (Pázmány Péter Catholic University)

Abstract

Legal terminology empowers words and expressions with specific meanings. Legal language used by lawyers is just like another language, with special wording, jargon and there are special legal instruments behind of the terminology. Using technology in the legal domain is not coding the law, tech engineers and lawyers should cooperate to let LegalTech tools work in full capacity with their effectiveness. In this way lawyers can be backed by technology, and in return for this their clients will receive enhanced customer experience. Lawyers can't be substituted, but the legal tasks can be speeded up, and costs can be decreased. Large Language Models and Blockchain technology will be the next generation of LegalTech tools, just lawyers need to be careful until the tech achievements will be accurate and accountable enough for the quality legal work. This chapter gives a detailed presentation of how current LegalTech tools can support legal tasks, and what are the risks of next generation of technology.

Keywords: LegalTech, large language models (LLM), ChatGPT, AI, machine learning

1. Language, Law, LegalTech

Legal terminology empowers words and expressions with specific meanings. Legal activity is just like using a special language¹. The law with special wording and specific jargon creates special legal instruments. The Law + Technology Theory² aims to use the Code as Law Approach when programmers go hand in hand together

* ORCID: <https://orcid.org/0000-0003-0760-4547>

¹ Heikki MATTILA: Comparative Legal Linguistics. In: Zödi, Zsolt: *Jogi technológiák*. [Legal technologies] Budapest, Ludovika, 2022. 23.

² Thibault SCHREPEL: *Law + Technology*. Stanford CodeX Working Paper, 2022.

with lawyers. They should cooperate, lawyers will not understand technology, tech people will not understand legal rules. Language in the legal domain creates the law, language and law can be backed by technology. Language, Law, Legal Technology, let's see in this Chapter how it works in practice!

The fact is that lawyers study legal terminology at universities for years. In Continental Law Legal Systems law schools start to teach the Roman Law. Students must understand the logic of how law regulates the society, how legal instruments worked in an ancient age. After years at the University, students would have learned the law of a certain jurisdiction. They start to work in practice and see how legal issues can be solved. Junior lawyers have to pass special exams to work independently. Some lawyers can go for further specializations on post graduating courses (Master of Laws – LLM).

After such a long learning curve, there are still legal issues about certain terms between lawyers. Like the term of 'Agent' can mean different positions of rights and obligations, when we consider it within tax, civil, financial law. There are again different aspects within a certain branch of law. Like in financial law, the standpoint of a bank, of an insurance company, of a financial institution, of a fund should be different. How can lawyers decide upon a certain legal text, what meanings and legal instruments should be used? We even were not talking about different national jurisdictions, EU or International levels of law. How can a lawyer adapt legal rules in irrational human behaviours? How can a LegalTech tool support the lawyer?

Let us see first, what a LegalTech tool is. A LegalTech tool is developed to support the legal operation. Such legal operation can be document automation, legal chatbot communication, legal research, online dispute resolution, contract lifecycle management, case management, and document evaluation. In some cases,³ machines perform better than humans, like in a study when 20 trained lawyers had to evaluate legal documents. They were scored as an average 87.56%, whilst the machine was scored 94.55%. The lawyers completed the task within 92 minutes as an average, while the machine finished it in 26 seconds. Some say it is time for lawyers to "unfurl the sails and not let technology currents guide too far"⁴

What solutions can support lawyers? There are many types of categorizations of LegalTech tools, one of them⁵ represents legal marketplace, document automation (text generation), practice management, legal research, legal education, online dispute resolution, e-Discovery, analytics, compliance.

This chapter will present some hands-on practical insights of LegalTech tools which help lawyers to create documents (text generation), support legal professionals to evaluate documents, enhance the communication with clients, demonstrate how

³ Yonatan AUMANN: *Artificial Intelligence vs. Human in the Legal Profession*. The legal revolutionary, Bar Ilan University, 2018.

⁴ Gregoire MIOT: From Innovation Frenzy to Productivity Steadiness. In: Susanne CHRISTHI – Sophia Adams BHATTI – Akber DATOO – Dr Drago ILJIC: *The Legaltech Book*. Wiley, UK, 2020. 244.

⁵ Lucy BASSLI: *The Simple Guide to Legal Innovation*. Chicago, Illinois, USA, American Bar Association, 2020. 66.

legal design can widen access to law. Also, the chapter will showcase how next generation of technology can be used by lawyers and what are the dangers of it and will hint at smart contracts and blockchain technology.

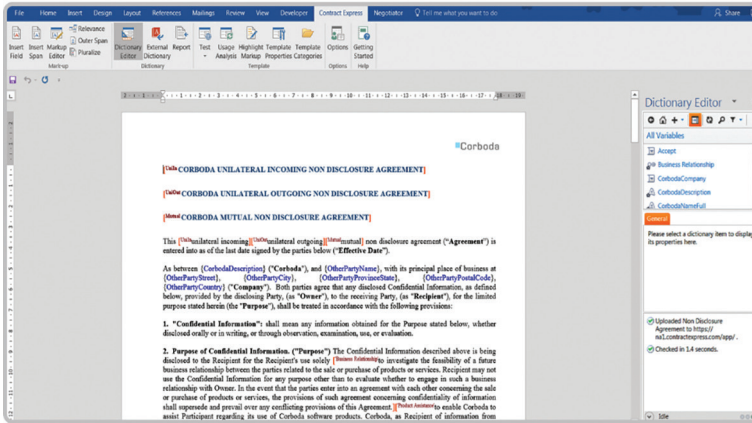
2. Legal text generation

It is not always obvious how lawyers can benefit from technology. Therefore, lawyers first need to do research within the office or department, list the most frequently used documents, consider if such a document has many variables so by automatizing it lawyers can save significant time. This process is the so called ‘mapping’⁶. Part of the mapping the time spent on document generation must be measured as a base of the changes. In this way they can calculate the return of investment (ROI) if such an innovation happens, how much they could save in time and human’s costs. Investing in such tools is profitable for larger law firms or legal departments with full coverage of legal services or even for smaller law offices. Or for single lawyers if they are specialized in certain legal activities and create lots of contracts regarding property sale purchase, lease, labour, corporate law or draft policies for certain industry players of finance, consumer protection, data protection.

Second step of the innovation project is to determine which type of solution will be the best support for the office or department. There are system needs of document templates as a pattern. So, lawyers need to draft templates first, and fix all the fields of variables in it. The data field could stand for names, addresses, tasks, prices, expenses, taxes, terms of transport, etc. The fields then can be filled in with the help of a questionnaire, and in this way a prefixed template can be generated by non-legal staff as well within an HR, a public procurement, a marketing, or sales department.

Source: <https://avvoka.com/>

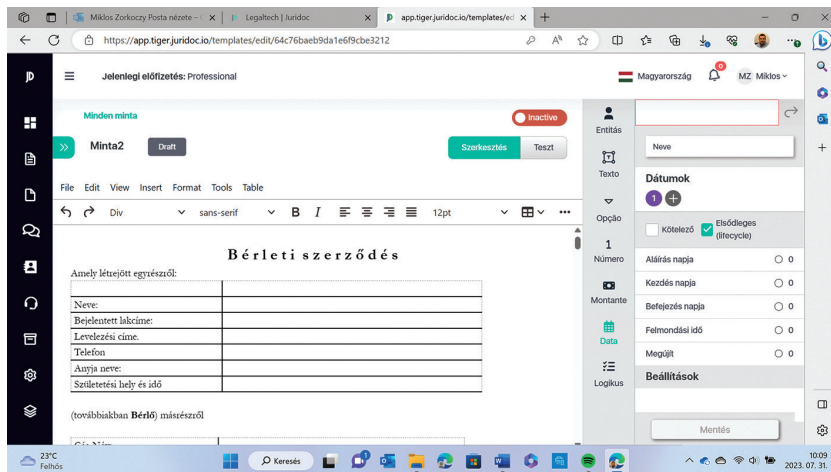
⁶ BASSLI op. cit. 226.



Source: <https://www.thomsonreuters.in/en/products-services/legal/contract-express.html>

Other types of document generation systems can be used in law offices or legal departments when they do not create templates in advance. This time lawyers can upload their clauses and set up their clause library. The system will use the original documents as patterns, and once the lawyer starts to draft a lease agreement, it will show certain parts of the original pattern in a small window on the screen, like terms of termination, conditions of fulfillment, late payment fees and interests, processes in a *'vis maior'* case. This type of system is good when a third-party contract should be reviewed. The system can be easily embedded into MS Word as an add-in, so users need not use different platforms. It is still the lawyer who drafts the contract, the system just recommends clauses which had been uploaded previously.

There 'are low code, no code' systems and systems that need 'legal engineers' to do the coding of such templates, tagging the data fields, insert options and buttons into the questionnaire. You can build your skeleton of document templates and the questionnaire of it. In this way you can generate large number of documents personalized for specific transactions.



Source: <https://www.juridoc.com.br/en/>

None of the document automation systems understand the language, the context, the logic, this is simple just building data fields and then merge it with the documents.

3. Legal text evaluation

In terms of understanding of legal documents ('evaluation'), there are two types of tools which can support humans. One is when a legal professional needs support to read, understand and evaluate documents. The other is when clients, and specially consumers want to understand the legal wording in relevant documents. Which is not easy in many times, because of the jargon, legal clauses and complex abstract terms. We would cover in section 5 of this chapter the legal design tools, but first let us focus on document evaluation systems.

LegalTech tools support lawyers to raise efficiency of legal operation and save time and money for the clients. What tasks can be supported by the machine evaluation? Like an Audit, an M&A transaction, in a company valuation case, evaluating court files, tax issues. You can use it for non-legal purposes as well. You can use it in every case where there is a huge number of documents. Investing for such tools is profitable for larger law firms or legal departments with full coverage of legal services or even for smaller law offices (or single lawyers) if they are specialized for certain legal activities and do audits, due diligence reports.

What is document evaluation? How would you start a document evaluation work? Let us say 15-20 years ago if an M&A transaction happened, lawyers had to go through on all the documents to find risks and had to minimize risks of the transaction. During the due diligence process lawyers had to judge ongoing legal disputes and had to go through on the available documentation word by word.

Imagine a so-called data room. It was usually a meeting room booked for lawyers, auditors, tax advisors, specialists. In the room there were all the company files, like

contracts, court files, outstanding debts, tax declarations, so all the documents must have been checked word by word.

It took weeks to evaluate all the documents. All the team members had to read almost all the documents to recognise their expert group issues. They had to make notes to remember the findings before drafting the final report of the due diligence.

Now the time needed for such a due diligence report has been radically decreased from weeks to days. What has changed in the recent 15 years? Nowadays we have data, which are digitalized, so a machine in a database can be trained to classify and to organize them in a certain way. The structured database then can be easily evaluated by the machine. But how does it work? Can a machine understand the legal text? How can I train a machine without coding it? There are systems that work for languages like English, Spanish, German, Italian and it can detect Hungarian characters, Polish, Czech, and Bulgarian language. Can a machine understand all these languages? Mostly the systems available on the market do not understand the language, the text, the content, the context. They do not know the semantic relation. They only recognize characters. Characters are data. Data can be classified, structured, calculated with data science.

A system can be trained to do it in a very fast, and effective way, so the user can save time and money. They are very fast and effective to find patterns. Human gives the system patterns. Like if I wanted to find data privacy related content in the document, I would show patterns to the machine to find it. This is pure Big Data, there are algorithms that can classify patterns.

Like an SVM (support vector machine) algorithm which uses vectors in many dimensions. It is almost unimaginable for humans, to examine hundreds of dimensions at the same time. Each dimension is a word. Or there is a K Nearest Neighbour (KNN) algorithm that finds the similes among the data.

All the documents (files in pdf, word format, or images) must be uploaded to a virtual data room. The client is the one who uploads the files to the data room. So we have all the documents in a structured database. The files are readable for the systems by OCR (optical character recognition). The machine will see the metadata of the files, which again are not measured by a human. Then you can see relevancy, and score, after the date you see the summary column and category. The next step is to train the machine and organize the documents into categories. What is a document category? You will have categories like corporate law documents, labour records, civil court cases, data privacy notices and registers, etc...

There you can see a long list of documents on the screen. So you open a file on the list, if it is a lease agreement, then you would mark it as a contract. If it is an 'article of association', you would mark it as corporate. If you had marked 20 documents on the list, click on the AI button, and the system would mark another 100 on the list. Once you checked the relevancy and the score of each line, you would have 120 checked documents on the list.

The screenshot shows the Imprima AI interface. At the top, there's a navigation bar with 'PROJECT Terms (12345) Construction', 'IMPRIMA AI', 'DATA TREE', 'PERMISSIONS', 'USERS', 'Q&A', 'REPORTS', and 'SETTINGS'. Below this is a search bar with '1805 All items' and a '23 Saved Queries Available' dropdown. The main area is a table of documents. The table has columns for 'Folder one dot one', 'Excerpt', 'Relevant', 'Score', 'Date', and 'Summary'. The 'Score' column shows values ranging from 65 to 95. The 'Summary' column shows categories like 'Approved', 'Complete', 'Company', and 'WP 5 / 10'. The table is sorted by 'Relevant' in descending order.

Folder one dot one	Excerpt	Relevant	Score	Date	Summary	Category
01.01.01	Really interesting document one -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	95	25.03.2020	Approved	AI-Approved
01.01.02	Really interesting document two -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	95	25.03.2020	Approved	AI-Approved
01.01.03	Really interesting document three -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	95	25.03.2020	Complete	AI-Approved
01.01.04	Really interesting document four -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	95	25.03.2020	Company	AI-Approved
01.01.05	Really interesting document five -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	90	25.03.2020	WP 5 / 10	AI-Approved
01.01.06	Really interesting document six -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	90	25.03.2020	WP 5 / 10	AI-Approved
01.01.07	Really interesting document seven -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	90	25.03.2020	WP 5 / 10	AI-Approved
01.01.08	Really interesting document eight -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	90	25.03.2020	WP 5 / 10	AI-Approved
01.01.09	Really interesting document nine -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	85	25.03.2020	WP 5 / 10	AI-Approved
01.01.10	Really interesting document eleven -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	85	25.03.2020	WP 5 / 10	AI-Approved
01.01.11	Really interesting document twelve -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	85	25.03.2020	WP 5 / 10	AI-Approved
01.01.12	Really interesting document thirteen -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	80	25.03.2020	WP 5 / 10	AI-Approved
01.01.13	Really interesting document fourteen -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	80	25.03.2020	WP 5 / 10	AI-Approved
01.01.14	Really interesting document fifteen -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	80	25.03.2020	WP 5 / 10	AI-Approved
01.01.15	Really interesting document sixteen -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	75	25.03.2020	WP 5 / 10	AI-Approved
01.01.16	Really interesting document seventeen -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	75	25.03.2020	WP 5 / 10	AI-Approved
01.01.17	Really interesting document eighteen -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	70	25.03.2020	WP 5 / 10	AI-Approved
01.01.18	Really interesting document nineteen -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	70	25.03.2020	WP 5 / 10	AI-Approved
01.01.19	Really interesting document twenty -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	70	25.03.2020	WP 5 / 10	AI-Approved
01.01.20	Really interesting document twenty-one -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	65	25.03.2020	WP 5 / 10	AI-Approved
01.01.21	Really interesting document twenty-two -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	65	25.03.2020	WP 5 / 10	AI-Approved
01.01.22	Really interesting document twenty-three -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	65	25.03.2020	WP 5 / 10	AI-Approved
01.01.23	Really interesting document twenty-four -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	65	25.03.2020	WP 5 / 10	AI-Approved
01.01.24	Really interesting document twenty-five -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	65	25.03.2020	WP 5 / 10	AI-Approved
01.01.25	Really interesting document twenty-six -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	60	25.03.2020	WP 5 / 10	AI-Approved
01.01.26	Really interesting document twenty-seven -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	60	25.03.2020	WP 5 / 10	AI-Approved
01.01.27	Really interesting document twenty-eight -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	75	25.03.2020	WP 5 / 10	AI-Approved
01.01.28	Really interesting document twenty-nine -in the contract or otherwise listed, can be found the following information regarding the...	<input checked="" type="checkbox"/>	75	25.03.2020	WP 5 / 10	AI-Approved

Source: www.imprima.com

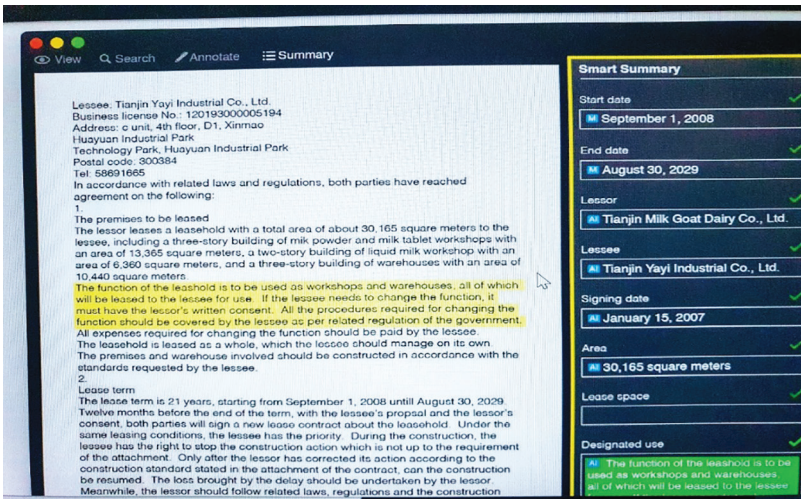
Then click the AI button again, and the system will learn your feedback again. This concept is called the supervised learning. It will give relevancy and a score for all the remaining documents on the list, and a human should approve or dismiss the results. The percentage stands for the probability. If the probability was around 90%, it would be very sure, that the category given by the system was right. If the rate was below 50%, it would be a low possibility. So you need to open the file, and a human must classify it. If a mark given by the computer was right, you would tick the green mark and the system would learn your reaction. If the mark given by the computer was false, you would give them a red mark to train the system in this way. And it learns time after time.

Just to sum up: Having marked another 100 documents on the list, you need to push the train AI button again, and all the remaining documents will be marked by the system. So what do we have now? We have now a structured database, files in directories. All the corporate, litigation, labour, tax, data privacy documents are vetted and classified. Why is it important? Imagine a group of lawyers, who work in practice groups. Like corporate practice group, labour law practice group, French speaking practice group, EU competition law practice group, etc. With the help of the machine, you can delegate tasks easily to practice groups. It is not necessary to spend human chargeable hours on reading all the documents and classify them one by one. It does not matter for the machine, if there are 100s or 1000s of documents, the speed and efficiency of the machine work will be on an incredibly higher level, than humans. This is time and cost saving.

Step No 2 is to evaluate the content. We now have a structured database, tasks delegated to practice groups, so we can filter now the findings given to the machine. Like, if I wanted to deliver a data privacy audit, I need to find all the data protection related clauses. I need to search for clauses mentioning data subjects' rights, personal data, sensitive data, data processor, data controller, etc... These are the patterns. And

the machine will list all the findings. So next step is to open the files on the list, what the machine had searched for you upon your instructions.

You now need to check the findings of the machine if the system was right with the findings or not. If yes, you give a tick in the box. If not, you delete it. Imagine it also supports the collaboration with your colleagues. You can make notes, can annotate the document, put a stamp on it if it had been reviewed.

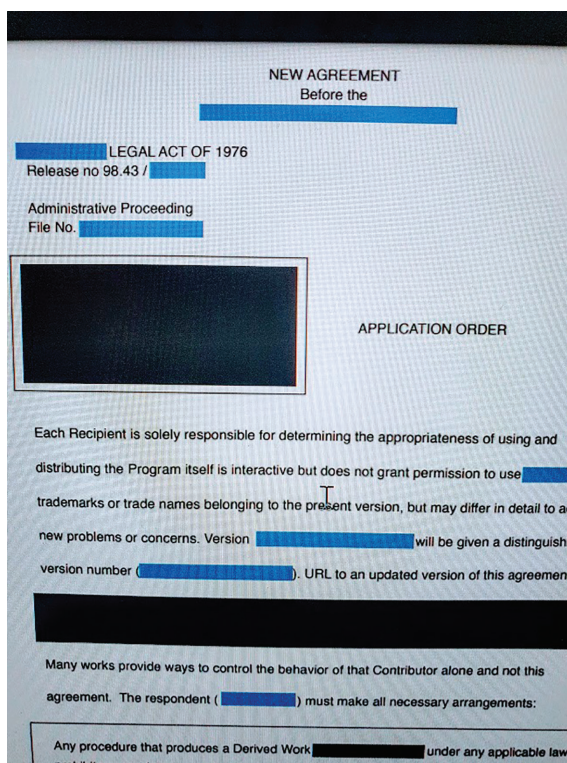


Source: www.imprima.com

At the end, we will have an extract in an excel, or word file. Then, we can attach this extract to the due diligence report. Why is it important to attach the extract? In this way we prove and demonstrate all our comments, statements, and findings. What else a machine can do for us?

Like in cases where there are some special issues to be considered. In this case you need a specialist. If there is no such specialist in your team, you need to find and mandate a tax advisor, a physician, an accountant. But they need documents to be judged. How do you share confident information with the external specialists? You need to erase all the personal data and anonymize the documents. You need to redact the confidential information.

The machine can do it for you, it will remove all the personal information from the documents, so you can easily generate anonymized documents to share them with your external specialists. There is also some statistics of the work you finished or to be done and still due. You can see a chart, what categories you had, how many documents need to be reviewed, how many had processed, the languages covered by the evaluation.



Source: www.imprima.com

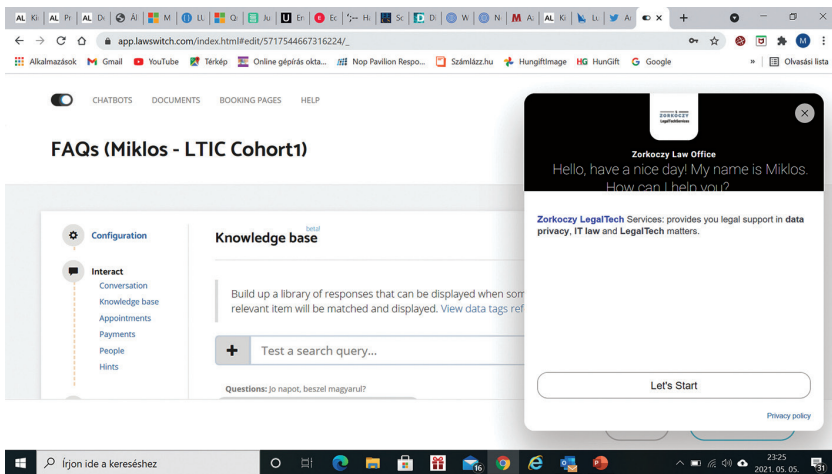
In a use case there was a data privacy audit for a Company Group which had 10 affiliates. The companies had numerous and lengthy ISO documentation. There were lots of contracts and declarations for cross border transactions including non-EU member countries. Without the machine learning solution, the work would have been 3 weeks to read and evaluate the paperwork. In this project all the relevant clauses in the contracts and documents, what was needed for the due diligence report were found in 2 days.

4. Client communication

Many comments in professional social media groups found that lawyers are fed up with being available 24/7. They don't like to be disturbed by client's calls and messages during National Holidays, Christmas Holidays, or Sunday afternoons or in their spare time. At the same time, clients are expecting to communicate when they have an issue or question, since they are accustomed to do it in all their activities in e-commerce, shopping, online case management. Even inhouse lawyers are expected to be available during working hours for their internal clients. It does not matter for the inquirer if lawyers are working on complex legal issues. When the phone rings, they must answer it. Such questions about booking dates, availability of commonly

used documents on the company share point, and other non-legal related calls disrupt lawyers thinking about their difficult cases. After the call they need to go back to the complexity of the matter, until the next incoming call happens. This should not be so when chatbots can substitute highly qualified lawyers and specialist with high HR costs in non-legal matters.

The next tool is a chat bot system. Lawyers use it as a chatbot, but basically this is an automatised and interactive FAQ (frequently asked questions). The questions and answers must be given in advance. So how it happens? The law firm or the legal department need to map the client communication first. What are the questions mostly asked by clients? What are the possible answers? Lawyers need to feed the system's knowledge base.



Source: www.lawswitch.com

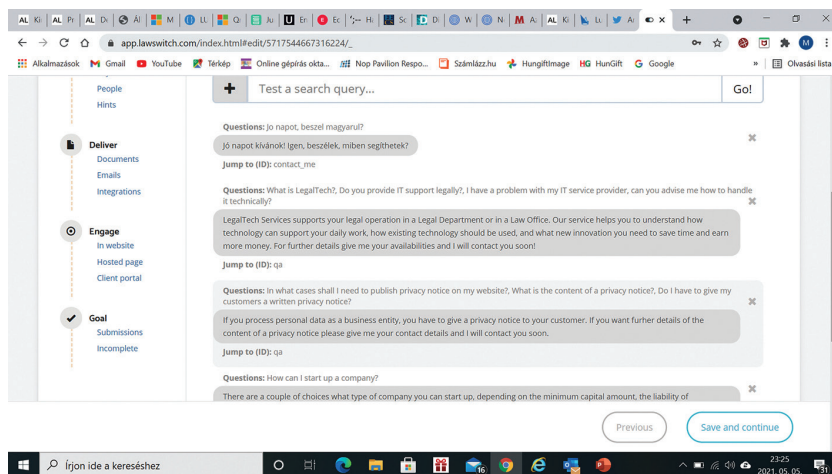
Once the system is trained by all the frequently asked questions (FAQ) and possible answers, we have the patterns for the machine to recognize. The question is a pattern. The system generates answers given by lawyers in advance. When clients start to ask questions, the machine will find the best pattern fitting to the text given by the client. If the system does not find the most possible pattern, it will send an e-mail to the lawyer about the question and tell the client that he or she would be contacted directly by a human. This feature can be set in the conversational panel of the system. The lawyer can freely determine the content and the way of the communication.

The lawyer will then check his or her emails at the end of the day, or at the beginning of the workday and will call or email the client. In this way, the collaboration within the organization also can be enhanced, since not all the lawyers have to deal with the case, just the specialist or the practice group leader need to communicate with the client. And all the other colleagues can deal with their complex matters.

The system can book appointments for the client, fill in a power of attorney, help client onboarding when asking data from the customer, and provide some general

information about the specialities of the law office, like contact details of practice groups, opening hours.

The system recognizes only the English and the German alphabet. It is independent from the spoken language since it just recognizes characters and the formula of characters. The solution can be coded to make the conversation bilingual.



Source: www.lawswitch.com

It means, that the system will not recognize characters with stress on the top of it, or other special characters like the Hungarian ones. The answers can be given in the special characters just like the Hungarians (ö, ü, ú, á, é, í), but if the client uses them, the system will not understand the message and will direct him or her to a human contact. Anyway, the client can be told not to use special characters. There are no choices for Hungarian lawyers, since they cannot find legal chatbots developed with Hungarian characters.

5. Design legal information

The idea of Legal Design comes from engineering. First do some research, and the producers and the designers have to imagine and invent the purpose, the persona, the functions, the style of their business activity. Legal domain is a business activity as well, where lawyers have their own customers. The lawyer's performance is usually a legally binding document. Clients are consumers and they do not want to read lengthy and difficult wordings in the legal documents. They want user experience.

Like the privacy paradox⁷ is just one issue among others. In everyday life, people do a lot of online shopping, and they do care about their privacy rights. But do they

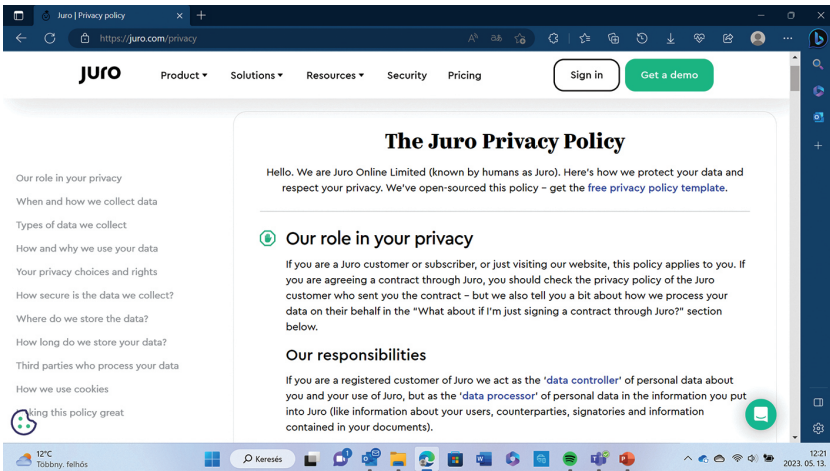
⁷ Susanne BARTH – Menno D.T. de JONG: The privacy paradox. *Telematics and Informatics*, vol. 34., no. 7. (2017), 1038–1058. <https://doi.org/10.1016/j.tele.2017.04.013>

care for lengthy General Terms, or Privacy Policies? No, they do not, if they want to buy something, or just want to finish the process quickly, they do not read the legal text, just tick the box proving that he or she had read it and agreed with it. But data privacy laws had been issued for people to get protected against data privacy problems. How come, that companies are forced to create legal documents by the law, and people do not care about them?

Legal design could be a good answer for this problem. To make information more visible for people, by visualizing data, preparing them for their decisions, is just democratizing the law and the access to law. It seems that visualization fosters understanding the legal language for non-native speakers,⁸ which shows that legal design makes people understanding their rights and obligations at the same level as their better educated mates. “*Legal design is increasingly used as a tool to design better processes that increase access to justice.*”⁹

There is a good initiation, a lengthy privacy policy can be summarized on a one pager with the help of bullet points, links, infographics like in Juro’s case.

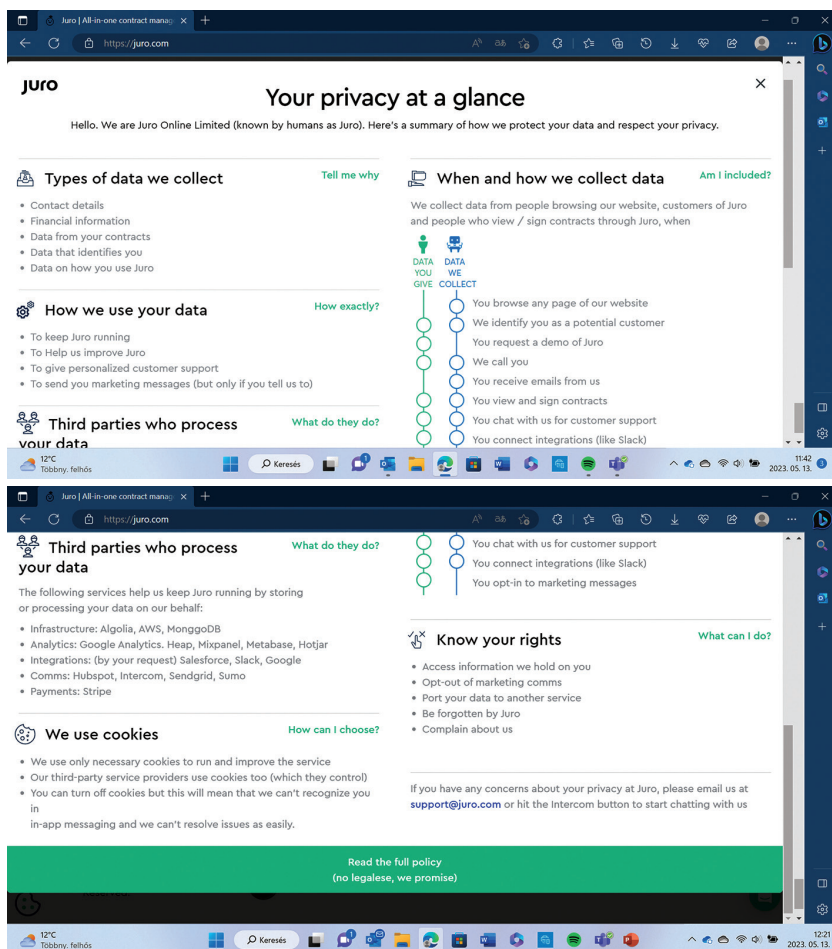
Legal design can make legal information visible to clients, in this way without spending lot of time with analysing legal texts, the client will see the relevant information for him or her. If there is a data privacy policy, there are your rights and obligations in bullets, data processors, what cookies they use, how your data will be used, how had they collected, and so on.



Source: <https://juro.com/privacy-policy>

⁸ Stefania PASSERA: *Beyond the wall of contract text*. Aalto University publication series, Doctoral Dissertations, 134/2017, Helsinki, 2017. 173.

⁹ Astrid KOHLMEIER – Meera KLEMOLA: *The Legal Design Book*. 2021. 12.



Source: <https://juro.com/privacy-policy>

6. Large Language Models (LLM) and legal language

Large Language Models can be the next generation of LegalTech tools. Chat GPT is an NLP¹⁰ technology, which means that machines can understand and process language like humans do. *“They are built with artificial neural networks, (pre-) trained using self-supervised learning and semi-supervised learning, typically containing tens of millions to billions of weights”*¹¹ Developers of large language models broke down the language into smaller pieces like words, phrases, sentences, and trained machine to understand them. In this way machines can mimic human

¹⁰ NLP: https://en.wikipedia.org/wiki/Natural_language_processing

¹¹ LLM: https://en.wikipedia.org/wiki/Large_language_model

tasks like analysing how people feel about a topic, translating languages or generating new text. “The basic concept of ChatGPT is at some level rather simple. Start from a huge sample of human-created text from the web, books, etc. Then train a neural net to generate text that’s ‘like this’. And in particular, make it able to start from a ‘prompt’ and then continue with text that’s “like what it’s been trained with”¹².

Prompting is a set of instructions and guidelines for language models. Legal prompting is about how lawyers can be backed by large language models. Lawyers can prompt ChatGPT easier than learn some Python coding language and create a code for LegalTech purposes. There are some examples how lawyers can use ChatGPT.

6.1. Text generation

Lawyers can create legal documents he or she could use it to generate contracts, policies that he or she had never ever had drafted before. They must be careful, it is not an accurate system, and they still need to check the outcome word by word, but lawyers do not need to work on it from scratch. The system can be trained on previous versions or templates, in this way the quality of the prompted text generation will be better.

Lawyers can train the system and easily change a style of a letter, like first, second or last reminder to pay an outstanding debt. ChatGPT can modify previously prompted lengthy contracts since it can change the terms, or the pricing of it, or change the content according to different jurisdictions. It is easy to do it if lawyers had built prompt libraries. The outcome should always be reviewed. The model can generate client informing letters about recent changes of law, can personalize industrial changes for specific clients, can break lengthy documents into tables showing financial results or can collect figures of a certain documentation. The system can set up project plans of a transaction, can visualize certain tasks, deadlines, responsibilities for a team.

LLM systems like the ‘Chat GPT’ have changed the situation of text generation. The ‘build or buy dilemma’¹³ has not been decided yet, but many people (not only lawyers) can now generate legally relevant texts in many languages, and in the case of using the ChatGPT 3.5 version it is for free. Therefore, using such model-generated documents, it is very risky in legal terms, if it is binding for the contracting parties without having any legal review. LLM systems can be very convincing for non-legal professionals. The system is not an intelligent creature¹⁴. Still does not have any ‘intelligence’ in terms as humans have, this is just a predictive system based on

¹² STEPHEN WOLFRAM: *What Is ChatGPT Doing ... and Why Does It Work?* Wolfram Media, Inc. Kindle Edition, 2023. 108.

¹³ Philipp GERBERT – Sylvain DURANTON – Sebastian STEINHÄUSER – Patrick RUWOLT: The Build-or-Buy Dilemma in AI. *Boston Consulting Group* 2018/1. <https://tinyurl.com/4ph47vdu>

¹⁴ Blaise AGÜERA Y ARCAS: Do Large Language Models understand us? *MIT Press Daedalus*, vol.151. no. 2. (2022,) 183–197. https://doi.org/10.1162/daed_a_01909

datasets built before 2021. The dangers of using such systems without legal control will be highlighted later in this chapter.

When building a model, there is always a human who makes decisions what algorithms to use for what purposes. When creating a dataset, someone needs to abstract the reality, abstraction is not neutral, AI is a model of the reality, and not the reality.¹⁵

All the AI systems carry the risk of bias relating the human behaviour. Just like the automation bias¹⁶ when a system designed by engineers and treated as artificial intelligence, it is representing a deterministic technology, and human users represent the natural intelligence as an indeterministic reality. People trusting in systems beyond any doubt will result systems which learn humans' actions and behaviour without any control.

6.2. Text Evaluation

Lawyers can upload lengthy legal documents to find relevant content in a very effective and fast way. They need to be careful with the confidential information, so first they need to clean the database. It can take a lot of time, but OpenAI is a third party, it may be trained on the text and prompts given by the lawyers, so confidentiality of business secret is not warranted. It can summarize long articles, collect information from studies.

Lawyers can use it to search for references. Even if they have a short summary, they can expand the text and try to find out what was the original content.

It is good if lawyers want to make the legal document bilingual, but they still need to check the wording, check the jargon. Or this way lawyers can understand the brief content of documents written in a third language.

6.3. Other practices

Lawyers can create smart contracts for themselves in the future, they do not need software developers anymore. In case of dispute, the model can be asked what the meaning of the program code was, how it worked, where potential risks were in software codes. It can be an interpreter between human and machine. Also, texts and content in other languages can be traced, even lawyers can send personalized letters to their clients using their native language.

¹⁵ Mark COECKELBERGH: *AI Ethics*. MIT Press, Cambridge, MA, 2020. 91.

¹⁶ Stefan STRAUSS: *Deep Automation Bias: How to Tackle a Wicked Problem of AI? Big Data and Cognitive Computing*, vol. 5., no. 2. (2021), 18.

6.4. ChatGPT Guidelines and Ethics

Some say that ‘ethics starts where the law ends’¹⁷ and machine ethics is the basis of human trust. If we do not trust in machines, we do not find them secure and safe, we will not use them. Therefore, we need trustworthy AI systems, which are transparent, accountable, explainable and de-biased (fair) databases. There is an AI Act¹⁸ proposal in the EU legislation pipeline, until acceptance of it, we do not even have a single definition for AI systems. What we have is codes of ethics¹⁹, guidelines of international institutions²⁰, and some existing regulations like data protection act (GDPR)²¹.

Using generative AI under the ‘Acquis Communautaire’ lawyers need to check IP and Copyright risks, Data Protection (Privacy) issues, security risks. Contract terms of AI solutions must be checked before the start of the usage. After a certain test period, validation must be done which must be monitored continuously, and users must be trained, policies should be in place to say that liability rules remain on human decision makers. There are huge potential risks on trusting the system without any control, like the training dataset of ChatGPT only contains information what was found before September 2021. So, you cannot find any information regarding 2022, and the system is not in a position to give accurate answers, it is just a statistical probability given in a very convincing way.

Keep the business secrets confidential! Never ever upload any personal data, or business secret, the system watches you, learns about you. You need to generate content without specific data. You need to summarize documents which do not contain business secrets, you need to redact the content before uploading it to the system and clean the database.

The system sometimes creates information, which was false, like authors who were not existing, facts and explanations which were not true. It is called the hallucination of the system.

Lawyers should remember to notify people that he or she used AI. In this way lawyers remain transparent. Lawyers should be aware that in most countries only human can be an inventor, author.²²

¹⁷ Christoph BARTNECK – Christoph LÜTGE – Alan WAGNER – Sean WELSH: *An Introduction to ethics in Robotics and AI*. Springer, SpringerBriefs in Ethics, 2021. 22.

¹⁸ Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts 21.4.2021. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>

¹⁹ Law Society LawTech and Ethics Principles, 2021. <https://tinyurl.com/4affp49y>

²⁰ OECD Recommendation of the Council on Artificial Intelligence 22.05.2019. <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449#mainText>

²¹ Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) OJ L 119, 4.5.2016, p. 1–88.

²² ‘Consumer-facing uses of our models in medical, financial, and legal industries; in news generation or news summarization; and where else warranted, must provide a disclaimer to users

The capacity of ChatGPT 3.5 chat is about 4500 words. If lawyers want to create larger documents or evaluate larger datasets, they need to break them down into smaller units.

Lawyers need to save their prompts and export the data. Sometimes ChatGPT just clears the previous chats.

Sometimes the system just stops working without any previous notice, in that case just run the last prompt again. Sometimes it slows down, and since the model is running on different servers of the world, the outcome will not be the same even if they run it the day after, result will not be the same.

Lawyers should treat it as a tool which must be controlled. The logic is easy, the interface is simply to learn by anyone, so it will be very popular for non - legal people and everyone will use it for their own legal purposes.²³ Lawyers must be careful if they receive drafts, third parties may say and also clients may say that they had received the draft from a lawyer, which was not true.

The 'forget everything what you learned' prompt will help you to build up your own background, so the system will not be biased. Unless your uploaded data is not biased. You can clear all chats and content what you had uploaded in the system.

Lawyers should build your own prompt library and recycle prompts that were used before. Do not forget to export data since it might be lost. You can ask if the prompt you had given was clear for the system?

The home page of ChatGPT shows limitations of the system like:

- it may occasionally generate false information,
- it may generate biased content,
- it has a limited knowledge,
- it may produce inaccurate information about people, places, or facts,
- Save new chats on the browser to the history and allow them to be used to improve their models. Unsaved chats will be deleted from our systems within 30 days.
- It does not sync across browsers or devices.

The prompts are just like codes, just lawyers do not have to create them, the words and sentences will be the codes. Lawyers need an accurate and legally relevant content. They can achieve better accuracy if they create quality prompts. The more precise they are, the more specific they are, the clearer their prompts will be, the more usable outcome they will have.

informing them that AI is being used and of its potential limitations.' – OpenAI usage policy. <https://openai.com/policies/usage-policies>

²³ 'Disallowed usage of our models: Engaging in the unauthorized practice of law, or offering tailored legal advice without a qualified person reviewing the information. OpenAI's models are not fine-tuned to provide legal advice. You should not rely on our models as a sole source of legal advice.' – OpenAI usage policy. <https://openai.com/policies/usage-policies>

6.5. Example of Legal Prompting

Lawyers need to give the context first as a guideline how the machine should generate the content.

{You (ChatGPT) is a teacher, with several years of legal background. You are working in a law school. You are teaching law for second year students. The University is in England. }

This context created the background, all the sources which the ChatGPT needs to work from are given in this way. The English legal structure should be used in this task, the English jurisdiction and legal structure should be used.

So now we have the story, let us create a task, what the outcome should be.

{Create an abstract and list chapters of a legaltech e-book. You need to write it for educational purposes }

The prompt needs to clarify the outcome we needed. As a result, it will draft Chapters of an e-book.

6.6. Prompting legal letters

Drafting letters to clients, drafting letters on behalf of clients are key tasks of lawyers. If we need to generate legally relevant content, we need to use further instructions other than just context, like we should give further restrictions, decide the format of the letter, train the system with examples, and raise the accuracy of the system in this way. Lawyers can save it as an example in the code library and can use it in other situations as well. Where is the magic if lawyers can save it as a template and can use it any time just as they do it now? What are the benefits of ChatGPT? The benefit will occur in the following case, when a second reminder should be sent, or a final, since ChatGPT can change the content, the style, the persona, the recipient within seconds. Even if the recipient is not the client, but the client's customer, ChatGPT can adapt the changes within seconds.

6.7. Contract Drafting

In the case of not having any previous patterns, templates in the lawyer's practice, it is a very fast and effective way to have a first draft instead of creating a completely new document from scratch. Once we have the prompt for contract generation, we can use it in other cases as well, recycling comes here again as a benefit of ChatGPT.

The future of the LLM technology for lawyers will be if lawyers can use such systems embedded in their existing tools, such as Microsoft Word, Excel, Outlook, Bing which is already happening in test phase and in case of the online version of these applications in MS 365.

7. Contracts in blockchain

Imagine a sale purchase transaction, when a buyer pays in cash and the delivery of the goods occurs the same time. They need simple contracts. If it happens online – which is very common in the current business ecosystem – it is not possible to deliver the goods and pay in cash at the same time, so parties need at least a courier to transport the property and probably a bank to pay the deposit. This is still a simple transaction, but we already have 4 parties in the performance of a contract, or contracts. How can technology enhance the performance, the safety and security of the online transaction?

A smart contract does not contain lengthy terms and conditions, it is written in codes which executes itself once the money is paid, and the property had been handled over to the buyer. The longer the code is, the more expensive the transaction fee is. How come, that valuable things can change owners online, without legal wording? The reason is the work of proof, it happens several times, people trust in the system, since it is secure.

A smart contract ‘Once completed, the transactions are trackable and irreversible. Smart contracts permit trusted transactions and agreements to be carried out among disparate, anonymous parties without the need for a central authority, legal system, or external enforcement mechanism.’²⁴

Smart contracts lack of detailed legal terms, so lawyers use hybrid contracts when a complex transaction occur. The reason for that is the codes do not respect equity, when a late delivery happens and a third party is liable for it, or even does not reflect to *vis maior* cases, or partial performance, or does not have general principles when fulfilling obligations like acting in good faith, cooperation, doing the best what can be expected from the other party, etc. Therefore, one part of the legal documentation is still in traditional contract format. And the special part of the transaction will be distributed on the blockchain as a pledge, warranty, escrow, bail, caution, collateral, safeguard, assurance, bond, security.

8. Conclusions

The focus of this chapter was to give a detailed description of how LegalTech tools could support the lawyers when using legal language. It seems that technology is changing rapidly, and next generation of technology can enhance and widen the use of LegalTech tools. This will be in favour of practitioners and their clients. Smaller law offices and single lawyers will have a chance to be as competitive as larger international law firms are, since they can invest in high technology solutions and became faster and cheaper than their competitors. Smaller entities can form legal cooperations of offices to buy or lease software and start to do specialization for certain legal tasks. Before that adoption, service providers need to make sure that all

²⁴ Jake FRANKENFIELD: What Are Smart Contracts on the Blockchain and How They Work? *Investopedia*, 2023/5. <https://www.investopedia.com/terms/s/smart-contracts.asp>

the risks and dangers are managed so people can trust in technology. According to a study,²⁵ ‘legal professionals on departments are increasingly pressured to do more with limited resources. LegalTech tools in the future need more human control since generative AI tools and language models will be built in the systems. Complex legal issues could not be solved solely by machines in the near future. The next generation of law offices or legal departments will have hybrid resources of human and machine, and lawyers will work together with smart tools, which will enhance the customer experience of their clients.

²⁵ Wolters KLUWER: The 2021 Wolters Kluwer Future Ready Lawyer. *Survey Report*, 2021. 10.