

AI SAFETY PRINCIPLES FOR LAWYERS

Dear Sirs,

over the summer holidays, our data protection and artificial intelligence experts - including the co-authors of the guide "**AI in the work of an attorney-at-law**" ([link](#)) published a few months ago (prepared in cooperation with **the National Council of Attorneys-At-Law** and **Microsoft**) - shared tips on LinkedIn on the safe use of AI by lawyers.

Our tips were presented as part of the summer series "AI Safety Principles for Lawyers". The series took the form of weekly posts on LinkedIn and was a practical discussion of the 10 rules we developed for the safe use of AI.

As the summer holidays are now over, we have compiled all the publications/posts into one and summarised them in an accessible format. We believe that our tips will be a valuable addition to the "AI in the work of an attorney-at-law" recommendations and will encourage you to read it - you are welcome!

The authors of the series "AI Safety Principles for Lawyers":

- Agata Szeliga
 - Anna Tujakowska
 - Sylwia Macura-Targosz
 - Katarzyna Paziewska-Młodawska
 - Agnieszka Jurcewicz-Androsz
 - Sandra Słowik
-

AI Safety Principles for Lawyers

1 Use AI as a support tool,
not a replacement.

2 You remain fully responsible for the
outcomes of your work — not the AI.

3 AI lacks your experience, expertise,
and understanding of context.
You alone are responsible for
strategic decisions.

4 Always verify the information
AI provides.

5 Learn how to give clear
and effective prompts.

6 Ensure the AI tool you're using
protects data confidentiality.

7 Start by using AI for routine or
repetitive tasks, giving you more time
to focus on work that requires your
judgment and creativity.

8 Don't neglect your writing and
critical thinking skills while using AI.

9 Stay up to date on technological
developments and your
organisation's internal policies.

10 Be curious — experiment with
how AI can support your work.



**[Download a printable
bookmark of the 10 principles](#)**



#1: Use AI as a support tool, not a replacement

This is one of the fundamental principles of using AI in legal practice. While it may seem obvious, reality shows that there are still situations where lawyers use AI-generated results without properly verifying their accuracy.

One of the most well-known examples of lawyers' *blind faith* in AI's infallibility is the case of two New York attorneys who filed pleadings containing non-existent court rulings. As it later turned out, these had been generated by ChatGPT and incorporated into the document without any prior verification. As a result, the court imposed a \$5,000 fine on the lawyers.

Similar situations have also occurred in Europe – a UK judge recently criticised lawyers who prepared pleadings using AI-generated content without verifying its accuracy. In his ruling, he also warned that such actions could result in criminal liability and undermine trust in the justice system.

In Poland, on the other hand, the case of our colleague Mikołaj Stępkowski, who received a trademark cancellation application referring to fictitious judgments, has likewise attracted attention.

Such cases are the result of so-called AI *hallucination* and excessive, uncritical reliance on AI (*overreliance on AI*).

AI improves the work of lawyers – it helps analyse data, summarise content, and organise information. However, it cannot replace their knowledge and experience.



Therefore, in our daily work, we should remember that:

- before using AI-generated content, it must be carefully verified – both for accuracy and for relevance to the specific factual situation;
- the final text should be written by us – AI can help us, but it should not decide for us;
- responsibility for the results of our work rests solely with us – not with AI.





#2: You remain fully responsible for the outcomes of your work — not the AI

While AI can support lawyers at many stages of their work, any professional responsibility for the final outcome still rests solely with the user - the lawyer.

Remember: AI is a tool, not a subject of the law. No legislation or rules of professional conduct provide for a 'transfer' of responsibility to the language model or technology provider.

- It is the lawyer, not the AI, who understands the needs of a particular client and knows the entire situational context.
- It is the lawyer, not AI, who takes part in the negotiations and recommends the final draft of the contract.
- It is the lawyer, not AI, who signs the pleading.
- It is the lawyer, not the AI, who suffers the consequences of a mistake.

AI can be compared to a knife: it can be used to prepare dinner, create a priceless sculpture, or cause harm. It all depends on the person holding it.

In the work of a lawyer, AI can be a tool to improve the work or communication with the client, but it can also lead to mistakes that can give rise to legal consequences. For example, when unverified data or legislation is used to prepare an analysis or legal advice (yes, LLMs can give you a reference to articles from the Code of Civil Procedure that do not exist but would fit perfectly to solve your problem).

💡 What does this mean in practice?

- AI does not understand the context of a case like you do – it may put something 'suggestively' but not necessarily accurately.
- AI may hallucinate – generate convincing but completely false information.
- AI will not take into account the nuances of client relationships, litigation strategy or the reputational impact of decisions.





#3: AI lacks your experience, expertise, and understanding of context. You alone are responsible for strategic decisions

Principle 3 reminds lawyers of their crucial role in the decision-making process - no matter how advanced AI becomes:

- AI does not 'understand' a case the way a lawyer does.
- AI can perform case tasks based on the lawyer's instructions, but it will not 'solve' the case.
- AI can make the lawyer act more efficiently and effectively, but it is the lawyer who draws the conclusions and decides which way to go.



What does this mean in practice?

AI works on the basis of data. It does not know the full legal, business and social background in which the client operates, it does not know the culture of the client's organisation, the client's values, and therefore it does not know the client's objectives and the risks involved in the decisions taken. The final decisions on what action to take, how to handle the case, when to go to court and when to negotiate a settlement - must rest with the lawyer.

Example: Leakage of personal data at a client, report to the DPA or not? AI will carry out an analysis of the prerequisites from the GDPR for the need to notify and present the result - theoretically negative. However, the lawyer will take into account that a year earlier the client was fined by the DPA for failing to report a similar leak, the lawyer has participated in communications/discussions with the DPA and knows the background of the decision.

AI does not have the experience and full legal knowledge. AI can instead support lawyers in analysing documents, generating drafts of pleadings, finding rulings, articles or summarising them. However, it is the lawyer who has a full understanding of the nuances of the case, the client's legal environment, negotiation strategy and context.

Example: We are negotiating a contract, we ask the AI to generate proposals for clauses that are favourable to the client. The lawyer will extend what AI has generated with the specific clauses required by the corporation in question according to its standards.

AI can help do 'more' and faster, but the conclusions and decisions should be up to the lawyer. Decisions made by the lawyer are also a matter of professional responsibility. The lawyer is responsible for the quality and accuracy of the advice given. AI can suggest possible solutions, but only the lawyer gives them meaning, priority and final form.

Example: A client is considering whether to invest in a new industry. AI can help analyse the industry regulations. In a prompt, AI can be asked about planned changes in legislation, but AI will not 'know' about the mood/trends in parliament during the legislative process.

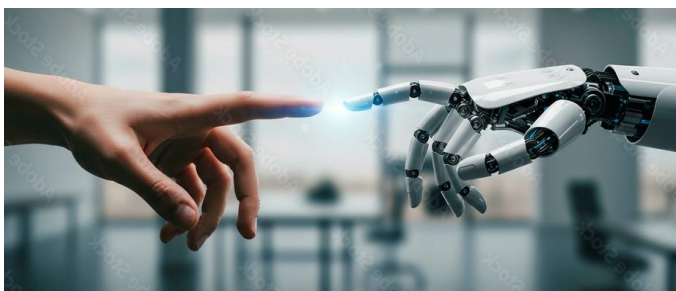




#4: Always verify the information AI provides

There is no place for "hallucinations" in the legal world - they are one of the biggest risks of using AI tools. Why? Because AI can create content that sounds convincing but is completely false, e.g. AI can quote non-existent laws or rulings or misinterpret existing regulations.

This is no longer a theory - in the US or UK, lawyers are being punished for using AI to create fictitious rulings. These are real consequences, not hypothetical scenarios. Sandra Słowik has written about this in more detail. Also recently, our colleague Mikołaj Stępkowski wrote about fictitious judgments found in an application for invalidation of trademark.



The recommendations in **"AI in the work of an attorney-at-law"** are clear:

- AI is a support tool, not a replacement for the lawyer.
- Information generated by AI should be verified by a human (the use of standard and proven tools may limit this verification).
- Responsibility for content always lies with the lawyer - not AI.

How can you tell if AI is "hallucinating"?

- Verify sources and references - if AI gives a case reference, the name of a legal act, etc. - check if it actually exists. Sometimes AI provides links to sources that are completely irrelevant to the issue under investigation. Very specific references (e.g. to article numbers) without citing sources should also be suspect.
- Ask questions about sources - you can ask the AI: "What legal act are you basing this answer on?" or "Give the case reference you are relying on?"
- Check dates and timeliness - AI may give outdated information. Pay attention to dates of publication of laws, issuance of rulings.
- Beware of answers that are too general - "hallucinations" often appears in answers that sound too good but lack specifics. Lack of names, dates, signatures - this is a warning signal.
- Test AI with follow-up questions - ask a question you know the answer to and see if AI answers correctly. This will help you assess its reliability in a given area. It's also worth setting a metaprompt: "Write if you don't know, don't hallucinate."



Why is this important?

Client trust is built on reliability. As professionals, we are obliged to provide the highest level of service, especially when we are paid for it.



#5: Learn how to give clear and effective prompts

Much has been written about how to prompt well. Today I am sharing my practical and proven experience. For anyone who wants to go into the subject in more depth, I also recommend reading Appendix 1 of the AI Recommendations for the Work of Counsel, which provides specific tips for optimal prompting.

Practice makes perfect

Most of us need to learn prompting from scratch. It is worth taking a few minutes a day to perfect your questions. They don't have to be solely about work; you can ask the AI about sights or restaurants in your holiday destination, for example. Pay attention to the prompts the tool gives you. They are often a great starting point for subsequent prompts.

Not sure how to formulate a prompt? Ask AI!

If your prompt doesn't work or you don't know how to attach a document, just ask AI how to do it. It really does work.

Think like a mentor

When creating a prompt, imagine you are giving instructions to a junior lawyer:

- indicate what needs to be done (the goal)
- explain the context of the task
- what you want to achieve, i.e. your expectations
- include sample material
- highlight the key points that need attention.

Example:

[OBJECTIVE] Prepare a draft personal data processing entrustment agreement based on the [A] [SOURCE] agreement compliant with the GDPR.

[CONTEXT] Act for the data controller.

[ISSUE] Enter the controller's objection in the event of sub-trust. The finished draft should be consistent with the intent of the contract, safeguard the controller's interest, specifically protect control over the data and provide for full liability of the processor.

Add context

If possible, use metaprompts or context settings. Some tools allow you to record permanent information, such as:

- "I am a lawyer at a large law firm".
- "Always cite the source".
- "Don't hallucinate"
- "If you don't know - don't make it up".

This way you don't have to write the general context every time, the one specific to the case will suffice.

Be specific and precise

We often hear that AI "did not prepare a legal opinion". And what did the prompt look like? Usually: "Prepare a legal opinion on this contract". This is not enough. AI needs precise instructions. If the topic is complex, break it into smaller parts, e.g. ask separately about the statute of limitations and separately about the validity of the contract. The results will be much better!

#6: Ensure the AI tool you're using protects data confidentiality

Does your AI tool truly protect data confidentiality? Make sure to verify this before you start using it!

In an era of growing popularity of AI-based tools, users - and in particular attorneys-at-law and advocates - should take extra care to ensure the security and confidentiality of the information entrusted to them by their clients.

💡 How can you ensure that chosen AI tool guarantees data confidentiality?

Start by reviewing the contract and documentation provided by the AI tool vendor, focusing on:

- commitments to maintain the confidentiality of input and output data,
- prohibition on using such data to train AI models,
- vendor liability for potential breaches of confidentiality,
- options for implementing additional safeguards, e.g. disabling data storage features for vendor review or limiting the tool's operation to the organisation's source material only.

Such a solution is envisaged, for example, by M365Copilot, where the user can choose whether to use the AI tool in a "closed environment" (Work), allowing grounding of documents/information containing client data, or in an open environment (Internet), where confidential client data should not be present.



Implement organisational security measures, including communication protocols and staff instructions. Regularly train those using AI tools in the organisation.

Remember, one of the biggest threats to data confidentiality remains **human error**. The now famous example of an engineer from a Korean company shows how easily confidentiality can be breached when AI tools are used irresponsibly. In this case, ChatGPT became the source of the leak of:

- confidential source codes of the company employees used ChatGPT to quickly fix bugs, while revealing confidential information;
- test sequences intended to identify defects - which were secret but got exposed when employees entered test patterns into ChatGPT to optimise faulty chip detection;
- a meeting summary document - an employee uploaded this document to the AI tool to prepare a presentation.

All of these situations could have been avoided: by verifying the AI tool's capabilities, applying appropriate restriction within its functionality, implementing organisational technical measures and staff training.

Remember - the duty of confidentiality and professional secrecy does not end at the organisation's door. In the age of AI, it starts with clicking 'Sign in' on a computer.



#7: Start by using AI for routine or repetitive tasks, giving you more time to focus on work that requires your judgment and creativity

A large part of a lawyer's work consists of repetitive and technical tasks that require a significant amount of time to complete. However, they are often essential tasks that cannot be avoided. At the same time, this time could be allocated to activities that require more intellectual effort and creativity. In such situations, the use of artificial intelligence (AI) may be a good solution, as it can become our digital assistant.

For which tasks is it worth engaging AI?

Examples include:

- transcribing meetings and automatically creating notes, including a list of follow-up tasks;
- drafting an email to a client with a summary of the meeting;
- simplifying the language of legal advice;
- preparing for a meeting with a client, including drawing up a list of questions;
- conducting preliminary research to collect basic information on a given issue;
- summarising long documents and organising information from large data sets.



AI can support us in these activities. However, it is important to remember that the results it generates must be verified by humans. The material generated can serve as a starting point for further work, which then needs to be completed independently. At the current stage of development of artificial intelligence, it is not possible for it to do all our work for us – its role is (and should be) to support us.

In our experience, AI works best where we already have the knowledge and only need to speed up the work. In areas we are familiar with, we are able to assess the quality and accuracy of the generated content (and importantly: we know which sources we can use to verify it). Consequently, we need to have the necessary knowledge ourselves to use AI results with full awareness and responsibility.

It is also worth emphasising that the more precisely we formulate the prompt (by specifying, among other things, the context, the audience, and the expected outcome), the better the result we are likely to obtain – in line with the *garbage in, garbage out* principle, according to which the quality of the output depends on the quality of the input. Sometimes, however, the best result is achieved not the first time, but through several successive prompts that make the task more specific and our expectations more precise.

Of course, in all this, one must not forget to comply with the principles of legal ethics and to ensure the confidentiality of information.



#8: Don't neglect your writing and critical thinking skills while using AI

In the article 'The risk of letting AI do your thinking', the Financial Times reports about an MIT Media Lab experiment that showed that people using large language models to write essays performed more poorly (linguistically, cognitively and behaviourally) than those who worked without AI, and over time increasingly pasted finished passages without much modification rather than writing them themselves. Another study found a significant negative correlation between frequent use of AI tools and critical thinking skills. At the same time, the authors stress that more research is needed and that potential risks can be mitigated by introducing simple rules that optimise the use of AI. Similar findings are highlighted by Microsoft's April 2025 research, which confirms that AI can significantly increase efficiency, but that with too much trust in the system there is a risk of reduced engagement in critical thinking.

For lawyers, this information is not necessarily a cause for concern, but certainly a signal for action (including a change in approach to educating young law students). What will set us apart from AI in the future is not the speed of writing, but strategic and critical thinking, which AI will not replace.

 **Remember:** critical thinking is the ability to analyse problems objectively, evaluate arguments and draw reasoned conclusions, a skill that is absolutely crucial in the work of a lawyer. AI can help with this, as long as it is used judiciously: it acts as a 'mirror' for our reasoning, pointing out gaps and alternatives that we might not see ourselves.

How to develop these competences using AI?

- Write independently, revise with AI - create your own draft first, then use AI to analyse and improve your text.
- Learn from revisions - analyse changes suggested by AI and learn from them for the future.
- Keep the final edit to yourself - this is a chance to consolidate your own style and skills.
- Generate counter-arguments - ask the AI to impersonate the opposing side and provide a retort to your thesis.
- Ask follow-up questions - ask about assumptions, exceptions and situations where an argument could be wrong.
- Simulate scenes and discussions - practise your argument in a 'conversation' with an AI playing the role of a judge or demanding client.

The AI can be your best sparring partner, as long as you decide the final shape of the argument.





#9: Stay up to date on technological developments and your organisation's internal policies

When implementing AI tools, it is worth remembering that it is not just a technology - it is a dynamic ecosystem that changes from week to week. Therefore, not only the implementation itself is crucial, but also:

- constant monitoring of AI tools being used;
- keeping up to date with market offerings and industry practices;
- regular reviewing and updating internal recommendations / guidelines / policies.

Why this matters.

- AI market is evolving rapidly - every month there are new models, new features within existing AI tools, and consequently also new risks. What was safe and effective a year ago, today may no longer meet either an organisation's needs or current regulations.
- Market practices are evolving - organisations are testing, comparing, implementing and abandoning tools that do not meet expectations. AI is changing, but so are business objectives and the regulatory environment.
- Therefore, it is worth treating AI not as a 'ready-made solution', but as an organisational competence that needs to be developed.



What does this mean in practice?

- Regularly evaluate AI tools you use - not only in terms of functionality, risks, but also compliance with regulations and organisational needs.

- Review changes made by AI tool provider (changes to regulations and price lists) and adapt internal guidelines/recommendations for the use of AI tools to these changes.
- Involve the team in the evaluation process of AI tools - because awareness is the first step to accountability.
- Keep track of new solutions available in the market - not only global, but also local - test, compare, learn.
- Observe the practices of other organisations - how they implement AI, how they train teams, how they manage risk.
- Organise periodic reviews of internal policies / guidelines / recommendations and update them - so that they are relevant to new functionalities, risks and obligations arising from the ever-changing legal environment. Determine the frequency of such reviews, the terms and conditions under which they are conducted and the person responsible for the process.
- Educate the team - because technology without knowledge is risk. The development of an organisation depends on the informed use of AI by all team members. A bottom-up approach makes technology of real benefit to the whole company.
- Document processes and decisions - this builds trust and transparency.

AI is not "buy and forget". It is a continuous process of adapting to technological, legal and organisational changes. Regular evaluation is not just an obligation - it's a competitive advantage.



#10: Be curious — experiment with how AI can support your work

Principle 10 reminds lawyers of the supportive nature of AI and the need to try, test and find appropriate AI tools. It is an encouragement for an open, creative and informed approach to working with AI.

How can this principle be understood?

- Don't be afraid to try new AI tools
- Test different applications of previously used AI tools, explore new features
- Learn by doing
- Look for new ways to increase your effectiveness



What does this mean in practice?

- More and more AI tools are appearing on the market for general use, but also dedicated to lawyers - not every tool is 'fit for purpose'. The dynamics of AI development, on the other hand, means that a tool that yesterday did poorly with contract comparisons is now doing well. It is therefore also worthwhile to keep up to date with the opinions of fellow professionals, e.g. following the 'impressions' of lawyers using AI on LinkedIn.
- Sometimes it's not worth 'catching several pigs by the tail', especially if you have already incurred the expense of using a particular tool. Experiment with what you have. It's possible that you don't yet know all the features of a tool or a tool can be used in a way you haven't yet thought of. Maybe the results will be better if you use a different function or work on the prompt - the time spent testing will definitely pay off.

- Every interaction with AI teaches you something new. You can't hone this skill by reading about AI on LinkedIn. By exploring the tool you learn about the possibilities and limitations, e.g. knowing what to look out for in a prompt. This is where practice is needed, this is where you need to test by example what and how AI will deal with.
- There can be many ways to increase efficiency with AI, e.g. writing emails faster, analysing data, automating repetitive tasks. Everyone has a different working style and performs different tasks, so there is no single way to increase efficiency - it requires self-analysis and adapting AI to one's needs.



We look forward to hearing from you!



Agata Szeliga

Partner, attorney-at-law

☎ +48 698 660 648

✉ agata.szeliga@skslegal.pl



Anna Tujakowska

Senior Counsel, attorney-at-law

☎ +48 698 694 645

✉ anna.tujakowska@skslegal.pl



Sylwia Macura-Targosz

Senior Associate, attorney-at-law

☎ +48 694 415 447

✉ sylwia.macura-targosz@skslegal.pl



Katarzyna Paziewska-Młodawska

Senior Associate, attorney-at-law

☎ +48 883 391 774

✉ katarzyna.paziewska-mlodawska@skslegal.pl



Agnieszka Jurcewicz-Androsz

Associate, attorney-at-law

☎ +48 600 782 823

✉ agnieszka.jurcewicz-androsz@skslegal.pl



Sandra Słowik

Associate, attorney-at-law

☎ +48 883 391 732

✉ sandra.slowik@skslegal.pl