

Are we entitled to damages if a „chatbot” gives us false information?

Artificial Intelligence (AI) appears in more and more areas of our daily life. Numerous sites use so called chatbots which –using AI- answer our questions, we can also encounter with „persons” powered by AI „who” try to direct us to the correct administrative pathway using short questions and our answers given to these during a call with a telephone customer service. But what happens if we suffer damages in situations like these due to potential misinformation or incorrect recommendation?

Unfortunately we have to wait a bit longer for the transposition of the European directive created for such issues as the Member States still have to approve it, but from its draft version a few details can already be known.

Firstly we have to clarify what does AI mean. The proposal for the 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 . This latter draft regulation defines the „artificial intelligence-system” as software that is developed with one or more of the techniques and approaches listed in Annex of the Codex and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with.” .”

Given the definition we can conclude that artificial intelligence has 4 key elements: (1) a softver which (2) was developed using specified technology, which (3) can create goals given by man and (4) the contents, predictions, recommendations or decisions generated in such way can interact with their environment.

To return to the question above, if we will be able to request damages if such a system creates a content, prediction, recommendation or decision which is wrong and harms us, the answer is yes.

According to the Policy damages can be enforced for harms which,

- come directly from the AI systems’ outputs (content, prediction, recommendation or decision), or
- for those, which originate from that, that the malfunctioned AI system did not create the wanted output.

It is an interesting fact, that in case the person operating the AI system cannot or do not want to make the evidence available to the court –that he or she is not at fault- than it has to be presumed that he or she did not act due diligence in connection with the damage. The causation link also has to be presumed if the AI was designed in a way that it does not allow effective supervision by a natural person.

All in all there are obviously many questions that still need to be clarified, but the rapid development of technology and rights of users of AU systems undoubtedly call for a complex regulation.