

# Mediation and AI: The Silent Revolution

## How Human is ChatGPT?

*By Michael Lardy, 2.7.2024*

The launch of ChatGPT in November 2022 was a furious one: One million people registered within five days. Unprecedented. Facebook took ten months to reach its first million users in 2004, and Instagram still took 2.5 months in 2010. ChatGPT<sup>1</sup> and other AI released over time have massively changed our world in a very short time, made many things possible that previously seemed impossible, and raised many questions.

It didn't take long for mediators to take notice of this new tool and ask to what extent AI can be used in mediation<sup>2</sup> and whether mediators might become redundant in the medium term<sup>3</sup>. The particular strength of a mediator lies in the skilful use of language and active listening that considers body language, facial expressions, gestures, and tone of voice. Successful mediators are masters of language.

AI like ChatGPT are so-called LLM, Large Language Models, AI systems trained with enormous amounts of text, whose particular strength lies in eloquent sentence formulation. Like mediators, LLM are also masters of language.

### Very Dynamic Developments

The current developments in AI are so incredibly dynamic that even top professionals occasionally make grand misjudgements: "Not even GPT 5000 will be able to understand this problem"<sup>4</sup> said Yann LeCun, head of Meta's (formerly Facebook) AI department, on January 22, 2022, in a podcast<sup>5</sup>. At that time, GPT 3.0 was current. Just eleven months later, in November 2022, ChatGPT 3.5 appeared and was able to solve the problem described by LeCun as unsolvable at the beginning of 2022.

Predicting when artificial intelligence will have a particular significance in mediation work would be mere guesswork. It is also questionable whether it is desirable for mediations to be taken over by AI systems in the future. However, the extremely impressive capabilities of AI systems and the incredible speed at which they are constantly being improved should not be ignored by mediators but rather addressed.

Susan Guthrie, a US lawyer and mediator, promotes her Mediation/AI webinars with the slogan: "AI systems will not replace mediators, but they will replace the mediators who do not engage with AI." A far-reaching statement.

Notably, the willingness to engage with AI in mediation varies greatly by country. Mediators in the USA and France are very open and interested. In the USA, there have been regular webinars on this topic for some time, conducted by well-known and prominent lawyers and mediators.

The behaviour in Germany and Austria is rather reserved. Here, it is noticeable that mediators, as well as lawyers, partially ignore current developments and feel a false sense of security. A study by the investment bank Goldman Sachs<sup>6</sup> concludes that in the near future, 44% of the work of lawyers and jurists will be taken over by AI<sup>7</sup>. Unfortunately, there are no studies yet on the consequences for mediators. Mediators and lawyers are partly convinced that their services are tied to human abilities that AI will never be able to provide. Doubts about this assessment are justified.

### AI - The Better Mediator?

Even before the current AI boom, the question of whether such systems could be usefully employed in mediation or alternative dispute resolution was being discussed.

In 2010, David Allen Larson, a professor at the Mitchell Hamline School of Law, published an extensive article<sup>8</sup> titled "Artificial Intelligence: Robots, Avatars and the Demise of the Human Mediator," where

he envisioned, thinking far into the future, humanoid robots with AI as future mediators. In 2018, Maxi Scherer, a professor at Queen Mary University of London, gave a remarkable lecture in Vienna on "International Arbitration 3.0 – How Artificial Intelligence Will Change Dispute Resolution" <sup>9</sup> and in 2021, Marita Katharina Wambach-Schulz, a professor at IUBH, International University Düsseldorf, questioned whether AI systems might be the better mediators. In her publication "Mediation and Artificial Intelligence" <sup>10</sup> she doubts whether mediators possess the essential qualities for successful mediation, namely neutrality and impartiality towards the parties. She refers to studies that question this and writes, "This raises the question of whether these problems can be resolved through gender-neutral robots as AI assistance to mediators in the social human-machine interaction."

### **How Human is ChatGPT?**

After ChatGPT 3.5 became generally accessible in November 2022 and then ChatGPT 4.0 in spring 2023, discussions about the use of AI in mediation, negotiation, and arbitration significantly intensified. At the same time, AI systems became faster and more powerful. This raises the question: How powerful are they?

A very interesting development in the last two years is assessing the capabilities of AI systems with psychological tests, the same psychometric tests typically used for humans. In the study titled "Who is ChatGPT? Benchmarking LLMs Psychological Portrayal Using PsychoBench" <sup>11</sup> this is justified as follows: "Given the potential for exponential advances in artificial intelligence, which could pose an existential threat to humanity (Bostrom, 2014)<sup>12</sup>, researchers have examined the psychology of LLM to ensure they align with human expectations."<sup>13</sup>

The psychological tests used to assess ChatGPT and other AI are very diverse. Among other things, empathy, emotional intelligence, various personality traits, and the Dark Triade (psychopathy, narcissism, Machiavellianism) were examined. This article focuses on the traits particularly relevant to mediation, namely empathy and emotional intelligence.

#### **a) L.E.A.S. Test – Level of Emotional Awareness**

In this standard test<sup>14</sup>, the ability of the test subject, in this case, an LLM, is tested to recognize and describe their own emotions and those of others. Of course, AI has no feelings, but it can very well capture and describe the emotional states of a person as described in a story.

The test is conducted with AI in the same way as with humans: 20 short stories are to be commented on by the AI. The test was conducted with the AI twice at an interval of four weeks. The results of this test are astonishing and were compared with the results of a human control group: Already in the first round, ChatGPT's ability to recognize and describe human emotions was better than that of the humans in the control group. In the second round, it was significantly better. Similar results were obtained in the EIS and WLEIS tests in the previously mentioned study "Who is ChatGPT ...".<sup>15</sup>

#### **b) Emotional Intelligence**

In this test, also described in "Who is ChatGPT ...," the emotional intelligence of the AI is tested. Again, unsurprisingly, ChatGPT scores above average compared to humans.

#### **c) Consequences?**

What conclusions can be drawn from these test results? The statement "Empathy is the soul of mediation" by Christoph Salger from the Strauss Institute for Dispute Resolution takes on new meaning. While it served as a clear distinction between human capabilities and the capabilities of AI in his article "Artificial Intelligence in Mediation-ChatGPT as Mediator 4.0," <sup>16</sup> published in June 2023, the distinction is no longer so clear.

As psychologist Marisa Tschopp noted in an interview<sup>17</sup>, "Empathy can be simulated." This does not

imply that AI has feelings. However, AI can already very well identify and distinguish human emotions in the context of mediation or a mediation simulation.

Tests have shown that an LLM can identify and rank the needs of a mediation party by importance.<sup>18</sup>

### **Human/AI Relationship**

The answer to the question of how and to what extent AI can be used in mediation largely depends on how people perceive their “relationship” with AI systems. The word “relationship” might seem misused here, as it is usually reserved for human-to-human interactions. However, a look into the past shows that this is not necessarily the case. A small time travel to 1965 leads to the USA, where Joseph Weizenbaum, a computer scientist and professor at MIT (Massachusetts Institute of Technology), programmed the computer program ELIZA.

This program<sup>19</sup>, which simulated a “dialogue” (keyboard input/screen output) between a patient and a psychotherapist, was very simple in design according to the computer technology of the time and was based on Carl Rogers' client-centred therapy approach.

Here is a short excerpt from a “conversation” between a client and ELIZA<sup>20</sup>:

**Client:** Men are all alike.

**ELIZA:** In what way?

**Client:** They're always bugging us about something or other.

**ELIZA:** Can you think of a specific example?

**Client:** Well, my boyfriend made me come here.

**ELIZA:** Your boyfriend made you come here.

**Client:** He says I'm depressed most of the time.

**ELIZA:** I'm sorry to hear you are depressed.

And so on...

The program quickly became known around MIT, and many students and university staff used it. Joseph Weizenbaum was surprised to find that his program was being credited with capabilities it didn't remotely possess. Test subjects confided very personal information to the computer. Users anthropomorphized this simple machine in ways that surprised Joseph Weizenbaum greatly. Even professionals were fooled: A group of psychotherapists to whom Weizenbaum presented ELIZA were very impressed. While they recognized some weaknesses in the program, they believed that with improvements, the program could conduct psychotherapy in the near future. They dreamed of a computer being able to treat multiple clients simultaneously. Today, almost 60 years later, there are still no digital psychotherapists.

In his book “Computer Power and Human Reason”<sup>21</sup> Weizenbaum criticizes the uncritical belief in the problem-solving power of computers and warns that this technology, if used unreflectively, could undermine our human values. Statements that are still relevant today.

Another example of the “anthropomorphization” of technology is an electronic toy for children, the Tamagotchi (1997)<sup>22</sup>. Tamagotchis are virtual creatures that need care from the moment they hatch. Tamagotchis have needs such as sleeping, eating, drinking, affection, and develop their own personality. If neglected, they die. But they rarely died, as children took their responsibility seriously and cared intensely for their Tamagotchis.

The tendency of humans to anthropomorphize technology and computer systems is a behaviour that has been studied for many decades. With the widespread adoption of LLM since the appearance of ChatGPT, the topic of anthropomorphization has become even more relevant: It is now possible to have long conversations with an AI that are on the level of normal interpersonal conversations, either in the mother tongue or any foreign language. The voice of the computer is now very human-like and

is improving at an incredible speed. This has brought human/machine interaction to a whole new level. In recent studies, such as in Marisa Tschopp's dissertation<sup>23</sup>, the idea that social relationships can only be human-to-human has been abandoned.

For the first time, how people shape their social relationships with conversational AI was studied. Applied to the topic of mediation and AI, it can be observed that people are willing to engage in a trusting relationship with an AI if they feel emotionally addressed. A very pleasant, trustworthy-sounding voice can be enough to establish a trusting relationship between human and machine. Stanley Kubrick's film classic "2001: A Space Odyssey" with the voice of the supercomputer HAL 9000 shows how a particularly pleasant-sounding voice can contribute to building trust between humans and machines. An AI mediator with the voice of HAL would presumably be very successful.

### **Active Listening**

The American psychologist and psychotherapist Carl Rogers described active listening<sup>24</sup> as a tool for client-centred psychotherapy (talk therapy), and it also plays a central role in mediation. The mediator should be open to everything the parties express through body posture, language, tone of voice, word choice, speech rhythm, facial expressions, etc.

This raises the question of whether LLM (Large Language Models) are also capable of active listening. Studies at the Faculty of Psychology at the University of Basel<sup>25</sup> show that AI systems can very accurately describe a person's emotional state by observing facial expressions. Trained AI was able to reliably recognize emotions such as anger, disgust, fear, joy, sadness, surprise, and love. Additionally, AI could detect micro-expressions—brief emotional expressions lasting milliseconds—that are not perceptible to humans.

Another important means of human expression that conveys much about one's emotional state is the voice. The voice provides valuable information. Speech Emotion Recognition (SER) is a very active research field that applies current machine learning techniques and neural networks. The nuances of emphasis, tone, phrasing, variations in speech speed and continuity, as well as accompanying physical gestures, convey something of the inner impulses and feelings of a person. Numerous studies have shown that LLM can capture and interpret<sup>26</sup> these nuances very well.

### **Status Quo: AI in Mediation**

How can an LLM currently be used meaningfully in mediation work? LLM, like good mediators, are masters of formulation and handling language, which is one of their major strengths. Equally important from a mediator's perspective is the extent to which AI systems can "understand" the context and the involved feelings and needs in a conflict situation between people.

Does an AI understand if one of the disputing parties feels unnoticed, undervalued, or perhaps unloved? Does an AI have the ability to actively listen? The previously mentioned studies that attest to LLM empathy and emotional intelligence have already indicated that AI systems are indeed capable of "understanding" and describing interpersonal feelings or conflicts. They do this very well and are improving at an astonishing speed. An AI, for example, can distinguish very well between positions and needs.

Asked what the difference between needs and positions is, ChatGPT 4.0 responds:

**ChatGPT:** *Positions are the concrete demands or stances a party takes in a conflict. For example: "I want a 10% raise." Needs are the underlying interests, values, or motives behind the positions. For example: The need for recognition or financial security.*

*In summary: Positions are what we want; needs are why we want it.*

It is crucial to understand that these answers are not the result of database queries. That would be mundane and nothing special. An LLM is not a database. An LLM has an "understanding" of what these terms mean and can identify positions, needs, and feelings in a conflict analysis.

The analysis of the following statement by ChatGPT is fascinating:

**Input:** Peter and Maria are driving together to friends; they are invited for dinner. Peter is driving and Maria says to him, "Peter, it is 5:55 PM." How does Peter understand Maria's statement?

**ChatGPT:** *In the described situation, Peter can interpret Maria's statement "Peter, it is 5:55 PM" in various ways, depending on the context and their relationship. Here are some possible interpretations:*

1. **Time urgency:** *Peter might understand that Maria wants to point out that they might be late and that he should possibly drive faster or that they should arrive soon.*
2. **Information sharing:** *Peter could take the statement as pure information, with no particular implication, simply noting the time.*
3. **Reminder:** *Peter might understand that Maria is reminding him that they need to be at their friends' place at a specific time and that they should keep track of it.*
4. **Stress indication:** *If Maria sounds concerned, Peter could interpret the statement as an expression of stress or worry about the timing.*

#### **You cannot not-communicate!**

**Input:** Peter and Maria are driving together to friends; they are invited for dinner. Peter is driving and Maria says to him, "Peter, it is 5:55 PM." Peter does not respond. What might Maria understand?

**ChatGPT:** *Maria's understanding of Peter's lack of response to her statement "Peter, it is 5:55 PM" could have various interpretations, depending on their relationship, the situation, and previous experiences. Here are some possible interpretations:*

1. **He noted the information:** *Maria might think that Peter heard her statement and noted it but does not find it necessary to respond because there is no immediate action required.*
2. **No need for action:** *Maria might assume that Peter is aware of the time and believes there is no reason to change speed or route, hence nothing to say.*
3. **Indifference or ignoring:** *Maria might understand that Peter is ignoring her statement or does not attach any importance to it, which she might perceive as indifference or rudeness.*
4. **Focus on driving:** *Maria could understand that Peter is concentrating on driving and therefore not responding to avoid distraction.*
5. **Dissatisfaction or stress:** *If Peter usually responds to such statements and does not this time, Maria might assume that he is stressed or upset and hence not responding.*

#### **What does this mean for practice?**

Due to the ability to "understand" and describe interpersonal behaviour, there are numerous applications.

## a) In Training

Since training involves working with fictional cases, unlike real cases, data protection issues do not arise. Based on role-play descriptions, it is possible, for example:

- **Conducting mediation simulations:** This works very well. One variant: The trainee takes on the role of the mediator and instructs the AI to assume and speak the roles of disputants A and B. The trainee mediator can conduct a fictional mediation with A and B, similar to the role-plays during training. The advantage is that this exercise can be repeated as often as desired and varied in content. The simulation can be expanded by instructing the AI to provide verbal feedback on the trainee's work at the end of the mediation. In practice, depending on the computer used, the entire "mediation," as in reality, takes place orally. Keyboard inputs and screen reading are out. This makes a mediation simulation very realistic and instructive. While ChatGPT had some weaknesses in these simulations in the early months, the AI has steadily improved. It now never loses track of the assigned roles, even in simulations with larger groups. "Mediations" with seven disputants plus a mediator have been successfully conducted<sup>27</sup>. A particularly exciting test: Can an LLM handle a mediation simulation with role reversal? Can the AI, in the role of disputant A, take on the role and perspective of disputant B and vice versa? Impressively, this is also possible and an indicator of how advanced AI systems have become.
- **Versatile design:** Simulations can be individually designed. A role-play description can serve as a basis, and the individual characters can be "optimized" with personality traits (consensus-oriented, opinionated, argumentative) to make the mediation more or less complicated. For the trainee mediator, this is a very good way to practice what was learned in the course. Tendentially, ChatGPT is quite consensus-oriented. Therefore, it is necessary to make the characters of the disputants somewhat more conflict-oriented to challenge the trainee mediator.
- **Conducting case analyses:**
  - What are the positions of the disputants?
  - What are the needs of the disputants?
  - How can I foster understanding between the disputants?
  - What questions could I ask?
  - What might be the background of the needs? Both fictional and concrete mediation cases can be deeply analysed with an AI. "Conflicts where a mediator is consulted are primarily not fought on the factual level but involve numerous psychological processes"<sup>28</sup> and mediators are generally neither psychologists nor psychotherapists. Therefore, it can be very helpful to use an AI to better understand the background of a conflict. For example, within the framework of a digital brainstorming session.
- **Drafting an ideal first contact and initial conversation:** First impressions count. How important the first contact and initial conversation are for the successful course of counselling or therapy is well described by Manfred Prior in his book "Optimally Preparing Counselling and Therapy"<sup>29</sup>. An AI can be very helpful in putting this knowledge into practice and drafting an initial conversation.
- **Creating a mediation agreement**

**b) In practice as a mediator:** Using an LLM in concrete, real cases currently faces a major problem: The available standard AI are not GDPR-compliant (as of June 27, 2024). While ChatGPT offers two options to prevent the further processing of chats by the AI, these optional settings do not meet data protection requirements. There are ways to use AI systems legally, but it is more complex than booking a premium account with OpenAI for \$20. Special providers or setting up a system based on an Open LLM are recommended. In principle, the application of AI in practice is similar to that in the "In Training" section. In all work with AI systems, the quality of the inputs determines the quality of the outputs. This concerns both the question of what data the AI was trained with and whether this data was content-balanced. It also concerns how good the user's inputs are. Learning to formulate good inputs ("prompting") must be learned and practiced, because no AI can yet read the user's mind to know what they really want. This is where many AI users fail. They misjudge the capabilities of LLM because they get poor results from poor inputs.

## Conclusion

Where does this lead? These developments are fascinating and frightening at the same time. Predictions about the consequences of technical inventions have often been absurdly wrong in history: neither the computer, the telephone, nor electricity were given the correct significance for the future at the time of their introduction. Parallel to the invention of the automobile, doctors warned that speeds over 30 km/h were life-threatening.

"Predictions are difficult, especially when they concern the future." (Winston Churchill)

Good reasons to be cautious about predicting future developments.

I am pleased if my article reaches many readers; therefore, publication, unabridged and with author attribution, is free of charge both online and in print after consultation with me.

Michael Lardy

Mediator / Médiateur Franco-Allemand

Law Studies 2020 - , Univ. Salzburg

Chemistry Studies 1983 – 1988, Univ. des Saarlandes

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

[mail@MichaelLardy.com](mailto:mail@MichaelLardy.com)

LinkedIn: <https://www.linkedin.com/in/michael-lardy-019394213/>

## References

<sup>1</sup> This article focuses on ChatGPT 4.0 Plus, as this AI has consistently delivered the best results in the context of mediation.

<sup>2</sup> Die Künstliche Intelligenz ChatGPT als Mediationsunterstützung? <https://inkovema.de/blog/kuenstliche-intelligenz-chatgpt-mediationsunterstuetzung-im-gespraech-mit-chatgpt/> (12.6.2024).

<sup>3</sup> Mediation im Zeitalter von ChatGPT: Werden menschliche MediatorInnen bald überflüssig? <https://michaellardy.com/blog/index.html> (12.6.2024).

<sup>4</sup> LeCun believed that an AI could not "understand" that an object placed on a moving object would move with it.

<sup>5</sup> Lex Fridman Podcast #258 <https://www.youtube.com/watch?v=SGzMEIJ11Cc> (58:47, 12.6.2024).

<sup>6</sup> Generative AI Could Automate Almost Half of All Legal Tasks <https://www.law.com/legaltechnews/2023/03/29/generative-ai-could-automate-almost-half-of-all-legal-tasks-goldman-sachs-estimates/> (12.6.2024).

- <sup>7</sup> NZZ: Kommt der künstliche Jurist? <https://www.nzz.ch/schweiz/kommt-der-kuenstliche-jurist-36-prozent-der-rechtsgelehrten-arbeiten-bereits-mit-ki-ld.1828484> (12.6.2024).
- <sup>8</sup> Artificial Intelligence: Robots, Avatars and the Demise of the Human Mediator <https://open.mitchellhamline.edu/facsch/351/> (12.6.2024).
- <sup>9</sup> International Arbitration 3.0 – How Artificial Intelligence Will Change Dispute Resolution [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3377234](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3377234) (12.6.2024).
- <sup>10</sup> Mediation und Künstliche Intelligenz <https://www.econstor.eu/handle/10419/231773> (12.6.2024).
- <sup>11</sup> “Who is ChatGPT? Benchmarking LLMs' Psychological Portrayal Using PsychoBench” <https://arxiv.org/abs/2310.01386> (12.6.2024).
- <sup>12</sup> Superintelligence <https://global.oup.com/academic/product/superintelligence-9780198739838?cc=us&lang=en&> (12.6.2024).
- <sup>13</sup> Who is ChatGPT? Benchmarking LLMs' Psychological Portrayal Using PsychoBench <https://arxiv.org/abs/2310.01386> (12.6.2024).
- <sup>14</sup> „ChatGPT outperforms humans in emotional awareness evaluations” <https://doi.org/10.3389/fpsyg.2023.1199058> (12.6.2024)
- <sup>15</sup> “Who is ChatGPT? Benchmarking LLMs' Psychological Portrayal Using PsychoBench” <https://arxiv.org/abs/2310.01386> (12.6.2024).
- <sup>16</sup> Artificial Intelligence in Mediation-ChatGPT as Mediator 4.0 <https://mediate.com/artificial-intelligence-ai-in-mediation-chatgpt-as-mediator-4-0/> (12.6.2024).
- <sup>17</sup> Psychotherapie und KI: „Empathie kann man simulieren“ <https://t3n.de/news/psychotherapie-ki-empathie-simulation-1522127/> (12.6.2024).
- <sup>18</sup> Possibilities & Limitations of Using AI in Mediation, Michael Lardy, unpublished.
- <sup>19</sup> Die Macht der Computer und die Ohnmacht der Vernunft, Joseph Weizenbaum, 1978, Suhrkamp Verlag
- <sup>20</sup> ELIZA – A Computer Program for the Study of Natural Language Communication Between Man and Machine, Joseph Weizenbaum, 1965
- <sup>21</sup> Die Macht der Computer und die Ohnmacht der Vernunft, Joseph Weizenbaum, 1978, Suhrkamp Verlag
- <sup>22</sup> <https://de.wikipedia.org/wiki/Tamagotchi> (14.6.2024).
- <sup>23</sup> Towards a Better Understanding of Human-AI Relationship Perception, Dr. Marisa Tschopp, Dissertation 2024
- <sup>24</sup> [https://de.wikipedia.org/wiki/Aktives\\_Zuh%C3%B6ren#Nach\\_Rogers](https://de.wikipedia.org/wiki/Aktives_Zuh%C3%B6ren#Nach_Rogers) (14.6.2024)
- <sup>25</sup> Machine Learning Facial Emotion ... <https://doi.org/10.1159/000534811> (14.6.2024)
- <sup>26</sup> A review on speech emotion recognition <https://doi.org/10.1016/j.neucom.2023.127015> (14.6.2024).
- <sup>27</sup> Possibilities & Limitations of Using AI in Mediation, Michael Lardy, unpublished.
- <sup>28</sup> Mediation: Psychologische Grundlagen und Perspektiven, Leo Montada, Elisabeth Kals, Beltz Verlag 2013
- <sup>29</sup> Beratung und Therapie optimal vorbereiten, Manfred Prior, Carl-Auer Verlag 2022