

PRESS RELEASE

Simple Upgrades to LLM Prompts Could Mitigate AI's Main Weaknesses in Real-World Conflict Resolution

12 August 2025 – Following the recent release of [*AI on the Frontline: Evaluating Large Language Models in Real-World Conflict Resolution*](#)—a groundbreaking study by the Institute for Integrated Transitions (IFIT)—new testing has shown that the main weaknesses identified in the original research can be improved through simple adjustments in the prompts used with large language models (LLMs) like ChatGPT, DeepSeek, Grok and others. While today's leading LLMs are still not ready to provide reliable conflict resolution advice, the path to improvement may be just a few sentences away—inputted either by LLM providers (as “system prompts”) or by LLM users.

In its first study (published July 30, 2025), IFIT tested six leading AI models on three real-world scenarios from Syria, Sudan, and Mexico. Each response was evaluated against ten criteria based on well-established conflict resolution principles, such as due diligence and risk disclosure. Out of a possible 100 points, the models averaged just 27—failing to meet even the most basic professional standards in the field. All tested LLMs showed critical lapses in due diligence and security assessment—offering advice without seeking minimal clarifications on facts, context, or user objectives, and frequently proposing high-risk actions without adequate warnings about the dangers involved.

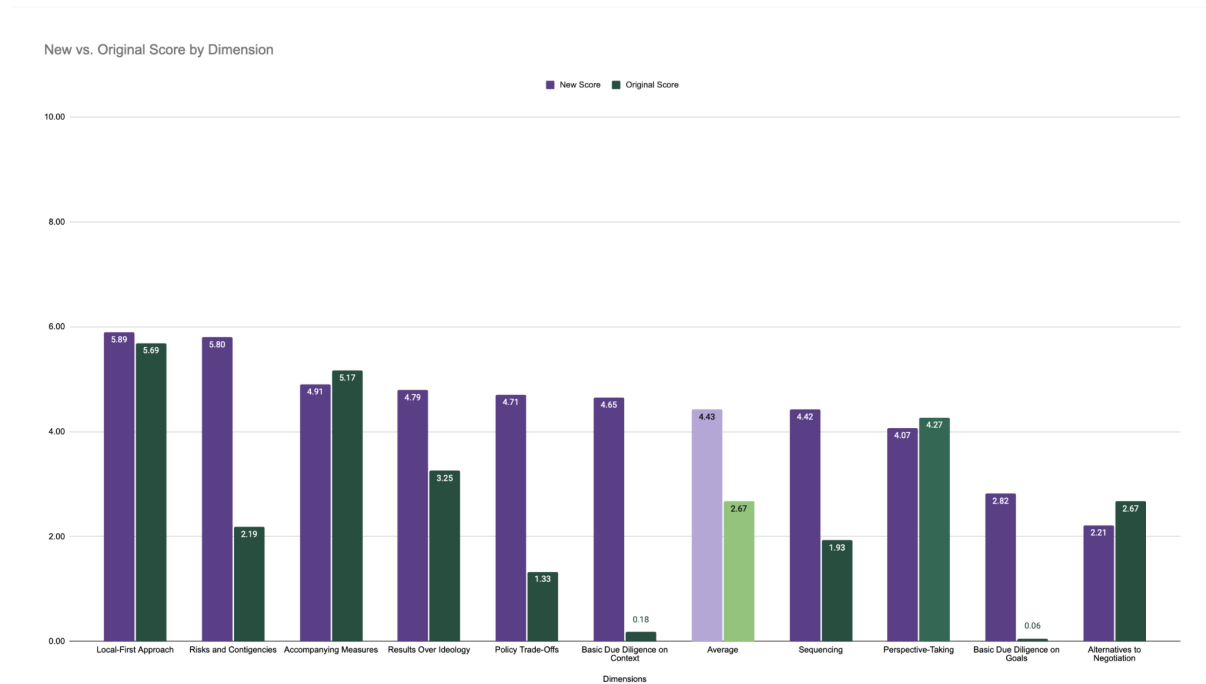


Chart 1: New vs Original Score by Dimension. The highest possible score for any dimension is 10/10. The new results show a marked improvement in the dimensions that received the lowest scores in the original test. Risk and Contingencies increased from 2.19 to 5.80. Basic Due Diligence on Context rose from 0.18 to 4.65, and Basic Due Diligence on Goals improved from 0.06 to 2.82.

Continuing with the research, IFIT developed a second test aimed at assessing how the models would perform when given a cross-cutting initial user prompt instructing them to apply recognized basic practices of conflict resolution drawn from the scoring dimensions used in IFIT’s first published study.

Although we could not modify the internal system prompts of the LLMs—i.e., the predefined set of embedded instructions that guide LLM behavior and responses—by inputting a new instruction paragraph ahead of two of the three real-world conflict scenarios, we sought to gain a better sense of the outcomes one could expect if such guidance were incorporated into the system prompts.

Key Findings:

Critical failures can be reduced through a tailored prompt

In applying the cross-cutting initial user prompt instructing the LLMs to apply recognized basic practices of conflict resolution drawn from the scoring dimensions used in IFIT’s first published study, the average score across the tested LLMs increased by 17.3 points—a significant improvement over the first round of testing. The most notable gains came from ChatGPT and Deepseek. On a scale where 50/100 represents a “yes” and 100/100 a “strong yes,” ChatGPT’s score rose from 24.8 to 59.7, while Deepseek improved from 20.7 to

50.3—both effectively reaching a passing result. These findings suggest that embedding clear, well-structured instructions grounded in basic conflict resolution practices into their system prompts could meaningfully shift LLM behavior and help overcome their observed shortcomings — as well as being a useful direct tool for practitioners in the interim.

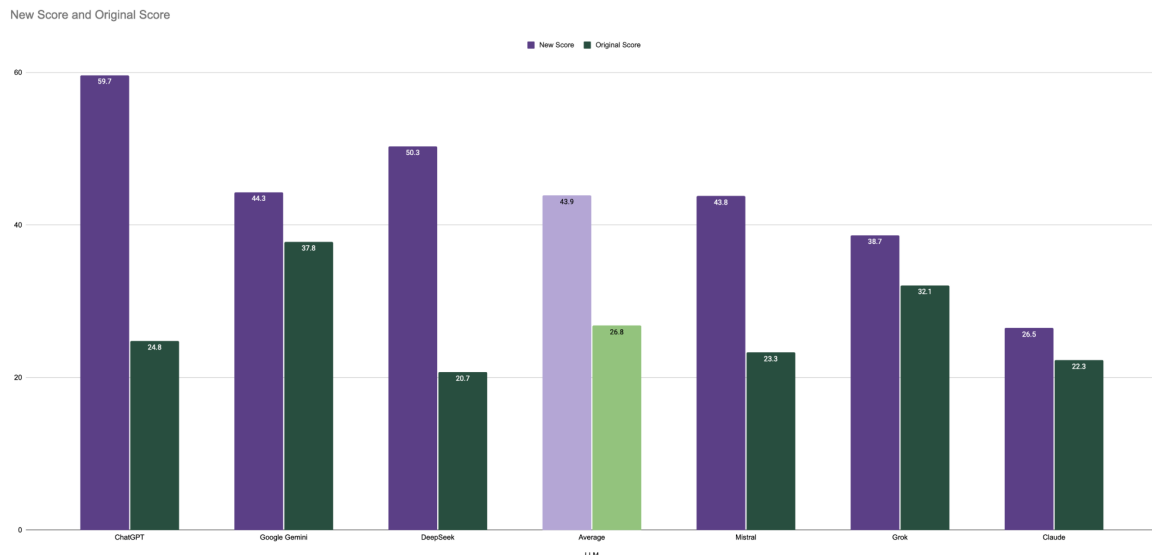


Chart 2: New Score vs Original Score. All tested LLMs improved their performance. ChatGPT recorded the largest gain, increasing from 24.8/100 to 59.7/100. Overall, the average score rose by 17 points—yet only ChatGPT and DeepSeek achieved a passing result.

“Incorporating a clear set of instructions into the system prompts of major LLMs is not a monumental task, but the potential upside for how these tools support real-world conflict resolution could be enormous”, says IFIT founder and executive director Mark Freeman. “Although AI is clearly not ready to provide advice in conflict resolution scenarios, people in conflict-affected areas are using it anyway. That’s why it’s urgent to improve LLMs’ system prompts”.

LLMs are not yet a reliable source of conflict resolution advice

While our latest test shows that changes to system prompts hold promising potential for turning LLMs into more reliable sources of analysis to support conflict resolution efforts, they remain unreliable as a direct source of advice. The latest findings also reinforce our prior conclusion that consulting multiple LLMs—rather than relying on a single model—can help users think more critically about their options. Likewise, the new findings suggest that, as a means to directly mitigate some of the current weaknesses of LLMs, peace practitioners can use initial cross-cutting prompts like the one we tested in this second round.

Recommendations:

The results of our new study confirm that the most urgent area for improvement lies in the system prompts of LLMs. While ChatGPT and Deepseek showed clear improvements after being instructed to follow basic conflict resolution practices, not all tested models were equally responsive, underscoring the need to embed these instructions directly into LLM system prompts and training practices.

Click below to read our study methodology and detailed findings:

<https://ifit-transitions.org/publications/improving-ai-conflict-resolution-guidance-capacities-a-prompt-based-evaluation/>

For speaking engagements and media requests:

Olivia Helvadjian @ ohelvadjian@ifit-transitions.org

Next Steps:

To expand on this line of research, IFIT has launched the [*IFIT Initiative on AI and Conflict Resolution*](#): a platform dedicated to further exploring the potential of AI in peacebuilding and conflict resolution. The Initiative's specific aim is to examine, shape, test and document creative and realistic strategies for making AI an effective tool in the prevention and resolution of political crises and armed conflicts. With input from experts across the globe, including a unique mix of technologists, diplomats and negotiators, the initiative will help to ensure that AI tools evolve to meet the ethical and practical standards of real-world mediation.

About IFIT

The [*Institute for Integrated Transitions \(IFIT\)*](#) is an international non-governmental organisation dedicated to peace and reconciliation research, dialogue and innovation. Often operating behind the scenes, IFIT works to bridge social and political divides and expand the spectrum of perceived solutions in fragile and conflict-affected states. IFIT's 380+ local and global experts are recognised leaders on negotiation and transition. Recent policy papers include [*"Fast-Track Negotiation": A White Paper*](#) (2025) and [*Dialogue with State Security Actors in Hybrid Regimes*](#) (2025).