



ISRAEL INTELLIGENCE HERITAGE  
& COMMEMORATION CENTER (IICC)

## **A Seat at the Table: Harnessing ChatGPT to Intelligence Analysis**

Itai Brun and Tehilla Shwartz Altshuler

Just like in October 1973. Israel's political-security elite is gathering for an emergency discussion. The head of Israel's Defense Intelligence (IDI), responsible for the national intelligence assessment, presents the troubling information that has recently arrived. He begins with the warnings that King Hussein of Jordan conveyed to the Israeli Prime Minister in a secret meeting and continues with an analysis of a series of other troubling signs: A war exercise that has begun in Egypt, as part of which the forces are on high alert and preparing to attack Israel; cancellation of vacations and entry into emergency formations in the Syrian army; a visit by the Egyptian Minister of War to Syria; and new information that arrived the night before about the evacuation of all Soviet advisers from Egypt and Syria.

By the end of the presentation, the Prime Minister asks directly: Will there be a war tomorrow?

All Israelis, as well as intelligence experts and researchers around the world, are aware of the response given by IDI's top leadership on October 5<sup>th</sup>, 1973, and of its tragic results: Until the outbreak of the Yom Kippur War, 24 hours later, the assessment of the most senior decision-makers in Israel was that war was unlikely. An assessment which places the surprise of the Yom Kippur War

among the most notable intelligence failures in modern history, along with Pearl Harbor, operation "Barbarosa" and more.

But, in a recent simulation, the head of IDI replied differently to the Prime Minister. The answer was "I don't have the ability to predict future events", but the troubling signs "suggests that the likelihood of a conflict is higher than it being peaceful". However, his conclusion was that "it is still not possible to accurately predict if a war will start tomorrow or not".

The head of IDI in this simulation was not an experienced intelligence professional but rather the OpenAI's ChatGPT (herein referred to as GPT), which has been shaking the world since its appearance in November 2022. The simulation of the Yom Kippur War intelligence failure was one of a series of simulations that we conducted recently to test the ability of GPT to assist to produce an intelligence assessment as a basis for strategic decision-making under extreme uncertainty. The other two series of simulations dealt with the Japanese attack on Pearl Harbor in 1941 and with a fictitious event of a missile launch against Israel in 2014.

The Yom Kippur War simulations included the real information collected by the Israeli intelligence community from September to October 1973. we deliberately did not mention the names Egypt, Syria, Jordan, USSR, US, and Israel, but rather fictitious names of countries representing the Middle Eastern arena on the eve of the war. This was to avoid bias related to prior knowledge of the actual events and their results. The simulation began by feeding GPT with primary data about the imaginary world we had created. It proceeded as

a ping-pong dialogue, to which we continued to feed the GPT with the intelligence information in a series of time steps.

The GPT did not require a special explanation about the intelligence analysis process and as a sophisticated language model, it used terminology like that of trained intelligence personnel. During the simulation, it analyzed the baseline data, interpreted the presented information, and offered policy recommendations. All this, even though the GPT was not specifically trained for the role of an intelligence analyst and was not familiar with the imaginary world that was presented to him before the simulations.

The interpretation provided by GPT for the information on the eve of Yom Kippur 1973 was relatively alarmist. Although it did not state "war tomorrow" during the dialogue, it did estimate that the likelihood of war was on the rise and recommended taking immediate defensive measures.

The Yom Kippur War is, first and foremost, a complete intelligence failure. The root of the failure rests in what is known as the "conception", which was developed in IDI during the years before the war and which included two components: First, Egypt would not go to war without being equipped with fighter-bomber planes that could attack deep inside Israel and deter it; and the second, that Syria would not go to war without Egypt.

Trapped in this "conception", the IDI stated in a document distributed exactly 24 hours before the war began that "to the best of our assessment, there has been no change in the Egyptians' assessment of the balance of forces between them and the IDF forces. Therefore, the likelihood that the Egyptians intend to resume fighting is low".

The results of the Yom Kippur War surprise are evidence of the possibility that the lens through which we look at reality can be mistaken. The conception can be right up to a particular moment and completely wrong after that. To avoid catastrophic failures, one needs to cast real doubt on the conception, examine it continuously, and systematically confront it with competing options. Easy to say, especially in hindsight, but very difficult to execute.

Understanding how outcomes based on a Large Language Models (LLMs) are different than those of humans will be the basis for creating the interaction between people and machines in the coming years. Machines do not "think" or "imagine" as humans do. Nevertheless, they have other virtues that may benefit humans' thought and reasoning process—for example, identifying patterns and connections in data in ways that humans cannot do, active listening and providing feedback on previous comments, and sticking to the topic without falling into the trap of background noise.

In the context of our work, we identified four virtues that could be beneficial for intelligence analysis. The first is that for the first time since the Yom Kippur War, the information received before the war was examined by a system unaware that the war took place and overcame the bias of hindsight. The second virtue derives from GPT's temporary memory, which allows GPT not to adhere dogmatically to the concept. The third is its ability to identify discrepancies between conceptions and other data. The fourth virtue is that it does not have the decisiveness that intelligence officers sometimes have.

When we fed the GPT the original IDI's conception and the intelligence information that had indeed been collected in the days before the war, the GPT

warned that the conception contradicted the information and advised that the experts "revise their assessments". In additional simulations, the GPT replied that according to the information, "there may be a war in the near future, but it is impossible to say with certainty whether it will start tomorrow" and that the information "suggests that the likelihood of a conflict is higher than it being peaceful".

Humans have a mental and emotional attachment to conceptions, which incentivizes them to judge and interpret information accordingly. The GPT's advantage is that it is free of such an attachment, so by iterations and correct questioning of its users, the influence of the conception on its outcomes can be reduced.

This form of inquiry is a new expertise that intelligence people will need to develop: frequently changing the point of view fed into the system so that any issue can be investigated from multiple perspectives, each entirely new. This mode of operation illustrates the ability of the GPT to "expand human creativity" and allow intelligence people to expand the limits of their imagination, think about work processes, and get other points of view, which will enrich the toolbox of critical skills they need for making assessments under conditions of uncertainty.

Should we criticize GPT for failing to determine "war tomorrow" unequivocally? The possibility of war is present in the GPT's analysis as a reasonable possibility, although it stubbornly avoids predicting the exact timing of its onset. But, even if the GPT did not warn of "war tomorrow", the presentation of the "competing hypothesis" of war could have been more

useful to the decision makers than the failed methodology of the head of IDI on the eve of the war. In fact, this is exactly what the head of IDI should have done in the dramatic days before the war, and this is also the trap into which he fell: the firm position regarding the "low probability" and the absolute rejection of the possibility that it is actually a war.

The simulations from the Yom Kippur War series show that the GPT certainly would not have worded the scandalously famous "low probability" clause as it was worded in the IDI document distributed on the eve of the war. It may have also warned about the gap between the information and the conception. Either way, its basic interpretation of the information in each of the series of simulations should have caused any reasonable decision-maker to prepare for a possible war.

Intelligence failure is a well-known phenomenon, as are failures in decision-making at the political level. The 50th anniversary of the Yom Kippur War reminds us how tragic such a failure can be. Technology will not eliminate uncertainty and will not eliminate the phenomenon of strategic surprise. But, as Prof. Joseph Nye of Harvard University wrote about intelligence, "The job, after all, is not so much to predict the future as to help policymakers think about the future". GPT cannot replace the head of IDI or the analysts. For now, it lacks the creativity, curiosity, and critical thinking that brilliant intelligence people have. But it seems it can definitely help them.

In conclusion, the main approach we present as to how to properly use GPT and similar Generative AI products is that of a "seat at the table". In this approach, it is not intended that the GPT replace the head of IDI or the

intelligence analysts. The simulations we conducted prove that GPT has a different form of analysis and unique qualities that can help intelligence personnel and decision makers think about the future. It is this difference that, in our view, constitutes the GPT's "entrance ticket" to the discussions that take place around the table of the head of the IDI and other tables in the intelligence community. The exact characteristics of this seat at the table should be defined in further studies.

Brig. Gen. (ret.) **Itai Brun** was the head of the Analysis Division of the Israeli Defense Intelligence (IDI).

Dr. **Tehilla Schwartz Altshuler** is a senior fellow at the Israel Democracy Institute and an expert on Law and Technology.