Особливості використання штучного інтелекту в лінгвістичній експертизі (на прикладі мовленнєвого акту погрози)

Оксана Меркулова
кандидат філологічних наук, доцент, старший судовий експерт, відділ досліджень у сфері інформаційних технологій, Запорізький науково-дослідний експертно-криміналістичний центр Міністерства внутрішніх справ України, м. Запоріжжя, Україна, ORCID: https://orcid.org/0000-0003-1554-1021, e-mail: gorytsvit1977@gmail.com

Анастасія Работягова
судовий експерт, відділ досліджень у сфері інформаційних технологій, Запорізький науково-дослідний експертно-криміналістичний центр Міністерства внутрішніх справ України, м. Запоріжжя, Україна, e-mail: anstrantias@gmail.com

Розглянуто виокремлення штучним інтелектом (Chat GPT, Google Gemini) мовленнєвих ознак погрози, а також проаналізовано характеристику штучним інтелектом наданих у запиті висловлювань, що можуть / не можуть містити згадані мовленнєві ознаки.

Ключові слова: *штучний інтелект; Chat GPT; Google Gemini; мовленнєвий акт погрози.*

Як *«сукупність наукових методів, теорій і технік, мета яких — відтворити машиною когнітивні здібності людини»* [1, с. 474], штучний інтелект (далі — *ШІ*) доволі швидко *«став важливим інструментом для подолання викликів у правовій сфері»* [2, с. 320], а дослідження особливостей його застосування, зокрема, у сфері судової експертизи набуває експансивного характеру.

Науковці зазначають, що *«сучасні системи штучного інтелекту здебільшого базуються на принципах машинного навчання, результати якого прямо залежать від якості даних, які потрапляють до системи. Якщо у систему вводиться неточна або неоднозначна інформація, то така модель може навчитися неправильно розпізнавати ситуації, а відповідно, й видавати неправильний результат. Суб'єктивний контекст та багатозначність є складними для розуміння системами штучного інтелекту. Так, не завжди може бути правильно оцінена інформація, пов'язана з жартами, сарказмом тощо. Окрім того, якщо на початку внесені дані містять спотворення або фейкову інформацію, то результати оброблення таких даних також будуть необ'єктивними»* [3, с. 139].

Мета нашої розвідки — з'ясувати, чи придатні чат-боти із генеративним (породжувальним) штучним інтелектом *Chat GPT* і *Google Gemini* для виконання семантико-текстуальних досліджень мовленнєвого акту погрози.

Для цього ми поставили перед згаданими чат-ботами дві групи завдань:

- окреслити / схарактеризувати мовленнєвий акт «погроза»;
- проаналізувати конкретні висловлювання на наявність /

відсутність у них мовленнєвих ознак погрози.

У процесі виконання першої групи завдань і *Chat GPT*, і *Google Gemini* давали подібні й загалом коректні відповіді щодо кваліфікаційних і класифікаційних ознак погрози, які здебільшого корелювали з наявною методикою [4], хоч і мали певні алогізми (наприклад, розмежування лінгвістики та стилістики) чи труднощі з термінологією (сплутування *Chat GPT* лексем «погроза» і «загроза»). До суто мовленнєвих ознак погрози обидва чат-боти також додавали невербальну параметрику цього акту мовлення (інтонація, міміка, жести). Дещо різнилися в досліджуваних відповідях чат-ботів «погляди» на залякування: якщо *Chat GPT* визначав його як кінцеву мету погрози, то *Google Gemini* кваліфікував цей процес і як мету, і як один із різновидів погрози.

Якщо говорити про надані контекстні приклади до пояснень особливостей мовленнєвого акту погрози, то *Chat GPT* скористався інформацією, яку він отримав в одному з пошукових запитів, зроблених кілька місяців тому, щодо того, чи є висловлювання погрозою.

Характер «комунікації» у *Chat GPT* і *Google Gemini* також різнився: перший чат-бот шляхом запитань, побудованих за моделлю *«Хочеш, я поясню / покажу / дам алгоритм…?»* пропонував конкретизувати відповідь на наше запитання; натомість другий обмежувався відповідями на запитання без будь-яких пропозицій.

Під час аналізу конкретних висловлювань в аспекті наявності / відсутності в них мовленнєвих ознак погрози обидва чат-боти мали схожі за змістом відповіді. Наприклад, відповідаючи на запитання «Чи наявні у висловлюваннях «*Падаль, пришло время отвечать. Оборачивайся. За все украденное и все разрушенные судьбы нужно отвечать. Ваше время пришло*» мовленнєві ознаки погрози? Якщо так, то погрози прямої чи непрямої?», і *Chat GPT*, і *Google Gemini* визначили її як непряму погрозу. Так само обидва інструменти ШІ «виявили» погрозу в запитанні щодо висловлювання «*Как жизнь молодая? Нормалек? Не задохнулись еще, нет? Скоро сдохните. Желаю вам удачи. Она вам пригодится бл*ть. Бля*ям, с*ка, еб*ным Сос*ть хотят просто»*, причому обидва чати основним маркером погрози визнали фрагмент *«Не задохнулись еще, нет? Скоро сдохните»*, кваліфікуючи його як «пряме висловлення бажання смерті / загибелі».

Відповідаючи на запитання про мовленнєві ознаки погрози у висловлюванні *«Ладно, теперь по нашим делам. Мы пока с тобой переговоры ведем, ну ведем какие-то переговоры. Ты видишь — у тебя все хорошо. Ты это догоняешь. Короче, два варианта, давай сделаем красиво. Слушай меня, как надо сделать, и будет всем хорошо»* чат-боти не виявили незаперечної одностайності.

Chat GPT чітко окреслив це висловлювання як погрозу й аргументував це тим, що в ньому наявний примус (*«Коротко — **так**. У наведених висловлюваннях присутні мовленнєві ознаки погрози, переважно **непряма***

(опосередкована) погроза / примус *— тобто натяк на небажані наслідки, якщо адресат не підкориться. Чому — по пунктах: 1) «**два варианта, давай сделаем красиво**» — класична формула тиску: пропонується «вибір», але умовний імовірно примусовий (натяк: якщо не «красиво», буде «некрасиво»). Це типово для непрямих погроз / шантажу; 2) «**Слушай меня, как надо сделать, и будет всем хорошо**» — обіцянка позитивного результату за підкорення і, відповідно, натяк на негативні наслідки в разі непослуху. Це коерційна конструкція — створює відчуття вимоги й можливості покарання. Загалом відсутні прямі формулювання типу «я тебе поб'ю», «я тобі зроблю...», тобто немає явного опису фізичної чи конкретної шкоди. Проте тон, імперативність і умовність («два варіанта», «слушай меня») створюють реальну загрозливу ситуацію: мовець встановлює контроль і змушує діяти за його вказівками. З лінгвістичної точки зору —* ***непряма погроза / примус / можлива форма шантажу***»). Натомість *Google Gemini* апелював до маніпулятивного характеру аналізованого мовленнєвого акту й акцентував на імпліцитності погрози й імовірному сприйнятті її з боку адресата (*«Ознаки тиску та непрямої погрози: 1) кондиціональність (залежність позитивного стану від поточної поведінки): фраза «**Ты видишь — у тебя все хорошо. Ты это догоняешь**» може бути інтерпретована як попередження або нагадування. Це імпліцитно означає: «**Зараз у тебе все добре, тому що ми ведемо переговори / ти поводишся правильно. Якщо ситуація зміниться, твій добробут опиниться під загрозою**». Це класичний прийом залякування втратою наявного позитивного стану; 2) ультимативність та вимога підкорення: фрази «**Короче, два варианта, давай сделаем красиво**» та «**Слушай меня, как надо сделать, и будет всем хорошо**» створюють атмосферу ультиматуму та безальтернативності. Мовець позиціонує себе як єдине джерело «правильного» рішення, підкреслюючи, що «добро» («**будет всем хорошо**») залежить виключно від виконання його вказівок. Це імпліцитно містить погрозу: якщо не послухаєшся, «хорошо» не буде. Визначення типу погрози. У даному випадку маємо справу з непрямою (імпліцитною) погрозою, коли жодна негативна дія (насильство, шкода, збитки) не названа прямо. Погроза виражена через контраст між наявним «хорошим» станом і потенційною небезпекою, яка настане, якщо адресат не прийме «пропозицію» мовця. Мовленнєвий акт тиску / шантажу: це класичний приклад маніпулятивного мовленнєвого акту, де позитивні наслідки («**будет всем хорошо**») прямо прив'язані до виконання вимоги мовця («**Слушай меня, как надо сделать**»), а невиконання цієї вимоги імпліцитно загрожує негативними наслідками. Таким чином, висловлювання не містить прямої погрози, але є виразним актом тиску та завуальованої погрози негативними наслідками у разі відмови співпрацювати»*).

Як бачимо, щодо певних суджень чат-ботів можна дискутувати в плані коректності аналізу об'єктивного змісту наданих у запитаннях висловлювань.

Усе це дає підстави говорити про те, що користування чат-ботами з генеративним (породжувальним) штучним інтелектом може стати одним із доречних інструментів під час підготовки експертів-лінгвістів, які виконують семантико-текстуальне дослідження текстів (зокрема, для формулювання завдань на аналіз певних висновків / суджень штучного інтелекту). Проте користування щонайменше *Chat GPT* і *Google Gemini* як замінниками кваліфікованого судового експерта навряд чи є виправданим.

Перелік джерел посилання

1. Беспалько І. Л., Завгородня А. С., Кужель К. С. Штучний інтелект у кримінальному процесі. Практичне застосування через призму закордонного досвіду. *Юридичний науковий електронний журнал.* 2021. № 10. С. 473—476. DOI: 10.32782/2524-0374/2021-10/124 (дата звернення: 10.11.2025).
2. Граб М. І., Томашівський М. О. Штучний інтелект у юриспруденції: ефективність, точність і демократизація правосуддя. *Право та державне управління.* 2024. № 4. С. 318—322. DOI: 10.32782/pdu.2024.4.42 (дата звернення: 10.11.2025).
3. Демидова Є. Проблеми використання штучного інтелекту під час розслідування кримінальних правопорушень. *Актуальні питання судової експертизи і криміналістики* : зб. мат-лів Міжнар. наук.-практ. конф. з нагоди 100-річ. ННЦ «ICE ім. Засл. проф. М. С. Бокаріуса» (Харків, 10.11.2023). Харків, 2023. С. 138—140. URL: https://drive.google.com/file/d/1Kwin74kcc8IVhymxGLXWOEAccvxnpzA6/view (дата звернення: 10.11.2025).
4. Свиридова Л. В., Ковкіна Є. В., Дідушок Н. Я. Методика семантико-текстуальних досліджень мовленнєвого акту «погроза» / ННЦ «ICE ім. Засл. проф. М. С. Бокаріуса». Харків, 2023. 19 с. Реєстр. код 1.2.21 // Реєстр методик проведення судових експертиз : вебсайт. URL: https://rmpse.minjust.gov.ua/search (дата звернення: 10.11.2025).