1.У найближчому, та й далекому майбутньому використання високопродуктивних універсальних архітектур, в тому числі універсальної архітектури для штучного інтелекту, буде неефективним.

2.Про апаратний «штучний мозок» почали серйозно говорити у світі тоді, коли складність експериментальних зразків такого «мозку» перейшла межу близько 10^{11} транзисторів. Тепер такий «мозок» реалізується на більш-менш універсальній архітектурі графічного акселератора і використовується в експериментальних системах типу ADAS та має собівартість порядку 10^4 дол. США. Не викликає сумніву, що збільшення кількості транзисторів удвічі не збільшить удвічі рівень інтелекту такого мозку, а лише на десяток відсотків. При цьому вартість та енергоспоживання збільшаться удвічі. Отже, суттєво покращити інтелект такої системи можна лише створивши архітектуру, яка є спеціалізованою до конкретного виду штучного інтелекту. Звідси випливає також таке.

3.Через гальмування закону Мура апаратний «штучний мозок» не буде з часом зменшувати свою вартість при тій самій або трохи зростаючій продуктивності. Тобто його стане невигідним замінювати на новіший через 3-5 років, як це трапляється досі з побутовими та серверними процесорами. Цей фактор, а також використання «мозку» в транспортних та інших роботизованих засобах з підвищеною небезпекою потребують довготривалого терміну служби такого пристрою. Щобільше, такий вартісний «мозок» будуть вилучати зі старих та пошкоджених моделей для повторного використання (зелена планета!). При цьому виникає вимога кардинального збільшення надійності апаратного «штучного мозку». Але ця вимога розходиться з тенденцією зменшення надійності інтегральних схем при зменшенні їхніх проєктних норм у новітній технології ІС. Єдиним виходом стає спеціалізація архітектури для такого штучного інтелекту.

Отже, загальні висновки такі:

Розробка універсальної архітектури штучного інтелекту та елементної бази для неї є марнотратством і не є доцільною.

Комп’ютерна інженерія штучного інтелекту буде розвиватись у напрямі пошуку нових ефективних алгоритмів та спеціалізованих структур для їх реалізації. Також багато уваги приділятимуть підвищенню ефективності програмування поширених архітектур, які адаптовані до завдань штучного інтелекту, як наприклад TensorFlow.

Слюсар В. І.

До підрозділу 7.1. «Штучний інтелект у сфері безпеки та оборони»

Тенденція до роботизації, що охопила різні аспекти діяльності людства, особливо помітна у військовій сфері. Провідні країни світу докладають значних зусиль щодо оснащення військових підрозділів роботизованими системами різного призначення та підвищення ефективності їх бойового застосування. Досвід військово-технічного співробітництва з державами-членами НАТО та відповідними країнами-партнерами свідчить, що військові аналітики розглядають штучний інтелект як проривну технологію для розвитку спроможностей військ. Упровадження штучного інтелекту є важливим трендом у розвитку систем управління полем бою та озброєнням, у тому числі роботизованими платформами [Stanley-Lockman and Hunter 2021].

У сфері військового управління технології штучного інтелекту розглядають як важливе доповнення до людських ресурсів за цілим спектром напрямів, зокрема: розширення ситуаційної обізнаності та обмін даними; координація командування підрозділами; розподіл цілей; координація функціонування датників і засобів ураження; виявлення та ідентифікація загроз, скорочення часу реакції на них; оцінка намірів; напівавтономний вибір зброї; робота з меншими ресурсами, з частковим вилученням людини з процесу прийняття рішень тощо. У перспективі оптимальний вибір комбінації сенсорів і засобів ураження, залежно від загроз,

має здійснюватися за допомогою штучного інтелекту, роль якого буде постійно зростати як при вирішенні завдань формування ситуаційного уявлення, так і підтримки прийняття рішень.

У 2017–2018 роках у НАТО розпочато процес вирішення завдань стандартизації ШІ. Наразі на цьому шляху вже пройдено кілька етапів. Перший з них стосувався термінологічних аспектів. На початковому етапі фахівці НАТО використовували кілька неофіційних означень штучного інтелекту. Зокрема, учасники досліджень NIAG SG-238 запропонували два альтернативні тлумачення терміна «штучний інтелект» [Слюсар 2020], згідно з якими це:

- спроможність, яку надають алгоритми оптимального або неоптимального вибору з широкого простору можливостей, для досягнення цілей шляхом застосування стратегій, які можуть включати навчання або адаптацію до навколишнього середовища;
- системи, які діють у фізичному або цифровому світі, враховуючи складну мету, сприймаючи своє середовище, інтерпретуючи зібрані структуровані або неструктуровані дані, обґрунтовуючи отримані з цих даних знання і обираючи найкращі дії (відповідно до заздалегідь визначених параметрів), які необхідно виконати для досягнення поставленої мети.

У заключному звіті з дослідження взаємосумісності спільної стратегії повітряних сил *BI-SC Final Report on the Joint Air Power Strategy Interoperability Study (JAPS-IS)* від 15 січня 2020 року використано означення, запропоноване NIAG SG-231: «Штучний інтелект (ШІ) – здатність небіологічної системи досягти будь-якої складної мети за допомогою процесів, порівняних із когнітивними процесами людини, таких як сприйняття, дедукція, розпізнавання, запам’ятовування та навчання».

Перше з офіційних означень НАТО (*NATO adopted*) було включене в настанову *AJP-3.10 Ed. B, Ver. 1. Allied Joint Doctrine for Information Operations*. У її проєкті, датованому травнем 2021 року, в переліку термінів зазначено, що ***штучний інтелект – це «розділ інформатики, присвячений розробці систем аналізу даних, які виконують функції, зазвичай пов’язані з людським інтелектом, такі як міркування, навчання та самовдосконалення».***

Паралельно з унормуванням означень ШІ експерти НАТО розпочали процес узгодження відповідних акронімів. Наприклад, при розробці настанови *AJP-3.3.2 (B) «Allied Joint Doctrine for Close Air Support and Air Interdiction, Study Draft 1»* було прийнято рішення вилучити скорочення *AI* для термина *Air Interdiction*, залишивши його стосовно штучного інтелекту (*Artificial Intelligence*). Разом з тим, гармонізація акронімів ще не досягла остаточного вирішення.

Віддзеркаленням поглядів аналітиків на можливі сфери військового застосування ШІ є поступово зростаючий за кількістю документів кластер стандартів, у яких відображені окремі аспекти ролі та місця засобів штучного інтелекту у тих чи інших місіях. Суттєво, що ці документи зі стандартизації вже охоплюють усі середовища ведення мультидоменних операцій – суходіл, повітря, морський та кібернетичний простори. Крім того, інтеграція відповідних нормативних положень щодо ШІ поступово поширюється на всі складові оборонних потенціалів *DOTMLPFI* (*Doctrine, Organization, Training, Materiel, Leadership, Personnel, Facilities, Interoperability*), а саме: доктрини, організацію, тренування, матеріальне забезпечення, лідерство, персонал, засоби, взаємосумісність. Наприклад, у настанові ATP-49(G) «Use of Helicopters in Land Operations» Ed. G Ver. 1 передбачено використання ШІ в наземній операції для управління підрозділом безпілотних літальних апаратів (*UAV*) (Рис. 2), у тому числі в складі інтегрованого з вертолітною групою комбінованого пілотажно-безпілотного підрозділу (*MUM-T*).

У настанові ASCP-01 Ed. A Ver. 1 (квітень 2020 року) «*NATO Stratcom Training Standards*» (*Annex F, page F-5*) стандартизовано вимоги до загальних компетенцій спеціалістів з оцінки інформаційного середовища (*Information Environment Assessment Specialists*), які повинні мати здатність розуміти і застосовувати штучний інтелект і машинне

навчання для оцінки інформаційного середовища (*IEA*): «*Understand and apply Artificial Intelligence/Machine Learning in IEA. (PL 200)*».

Дорожня карта реалізації спіралей Федеративної мережі місій (*FMN*) у редакції 2019 року визначила метою 6-ї спіралі *FMN* покращити процеси аналізу та прийняття рішень шляхом включення в них ШІ. Відповідний графік реалізації 6-ї спіралі передбачає, що в листопаді 2022 року буде розпочато процес формування вимог до експлуатації та безпеки з завершенням у 2024 році розробки фінальних специфікацій, у тому числі технічних. При цьому початок експлуатації відповідних технологій штучного інтелекту в рамках *FMN* запланований на 2027 рік з їх масовим оперативним використанням у 2028–2029 роках.

Значні сподівання щодо використання ШІ в медичній сфері відображені в настанові *AJMedP-5 Ed. B, Ver. 1 «Allied Joint Doctrine for Medical Communications and Information Systems»*. Зокрема на с. 5-10 міститься окремий параграф «*Automation and Artificial Intelligence*», в якому вказано, що автоматизація процедур і використання нинішнього та майбутнього штучного інтелекту дозволять поліпшити командування і управління медичними підрозділами, особливо в ситуації з масовими жертвами, будь то в бойовій чи в гуманітарній місіях (*The automation of procedures and the use of current and future Artificial Intelligence will enable better command and control particularly in Mass Casualty situation whether in combat or on humanitarian missions*).

Той факт, що штучний інтелект вже може суттєво впливати на ведення наземних операцій, констатовано в настанові ATP-3.2.1.1 «*Conduct of Land Tactical Activities*» (Ведення наземних тактичних дій), додаток D «*Considerations on countering UAS threat*».

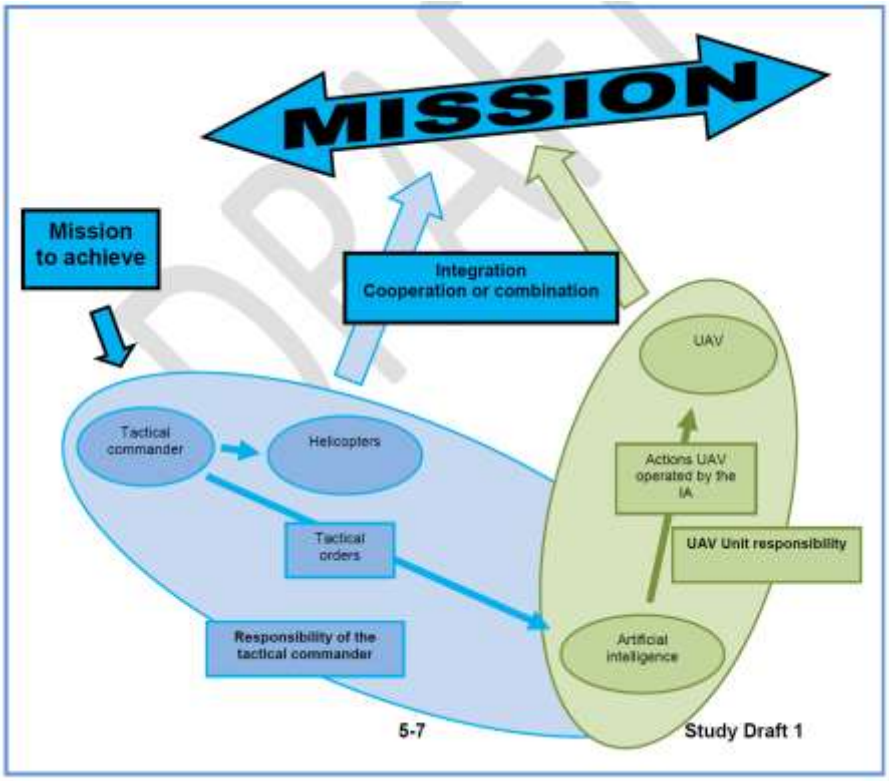

Рис. 2. Використання штучного інтелекту в складі MUM-T (ATP-49(G)).

Проте найрадикальнішим є підхід, запроваджений у доктрині AJP-3.10.2 «*Allied Joint Doctrine for Operations Security and Deception*», *Ed. A Ver. 1*, яка стосується операційної безпеки і обману. В ній вперше ШІ фігурує нарівні з людиною як суб'єкта, що приймає рішення: «*For the purposes of AJP-3.10.2, a decision-maker is understood to be a person or artificial intelligence responsible for decision making within an adversary or population's hierarchy*». При цьому ШІ як особа, що приймає рішення, може перебувати на будь-якому

рівні ієрархії та в будь-якому середовищі і має бути в змозі впливати на посилення поведінки або креативно змінювати її.

На черзі – початок технічної стандартизації засобів штучного інтелекту, зокрема вимог з безпеки застосування ШІ в системах озброєння тощо. Першим технічним стандартом НАТО, який враховує застосування ШІ, може стати настанова *AOP 4452*. Згідно з презентацією голови підгрупи з безпеки дизайну боєприпасних систем (*SG/B*) на засіданні Групи НАТО з питань безпеки боєприпасів (*AC/326, CASG*) у червні 2021 року, до проєкту оновленої версії цієї настанови планується включити розгляд вимог з безпеки використання ШІ в боєприпасних системах. Річ у тім, що інтеграція модулів ШІ до розумних (*smart*) боєприпасів розглядають як важливий тренд у розвитку засобів ураження. Такі модулі будуть здатні аналізувати поле бою, забезпечувати виявлення та ідентифікацію цілі в попередньо визначеному районі й обирати відповідний ефект, специфічний для ідентифікованої цілі. Зокрема, боєприпас зі штучним інтелектом повинен відрізнити важку броньовану ворожу машину від піхотного загону в пішому порядку і надати в першому випадку кумулятивний ефект (*HEAT, High Explosive Anti Tank*), а в другому – ефект осколкової або несмертельної дії (наприклад, потужний електромагнітний імпульс). Баражувальний боєприпас або ударний безпілотний літальний апарат (БПЛА), оснащений зазначеним модулем ШІ, збільшить здатність пригнічувати або нейтралізовувати дії супротивника, щоб зменшити кількість нанесених залпів, мінімізувати час уразливості від контрбатарейного вогню противника, максимізувати ефективність боєприпасів із мінімальними супутніми збитками.

Разом з тим, аналіз поточного механізму стандартизації в НАТО дає змогу зробити висновок, що стандартизація технологій штучного інтелекту у військовій сфері має здійснюватися як складова процесу формування системи систем стандартів (*System of Systems of Standards, S3*) [Слюсар, 2017]. Ця *S3* повинна являти собою ієрархічну, багатовимірну та багатофункціональну, взаємоузгоджену інтеграцію системоутворювальних кластерів нормативних документів. Потреба саме в такій структурі системи стандартів зумовлена тим, що вона повинна забезпечувати розробку, випробування та обслуговування всього життєвого циклу системи систем озброєння та військової техніки із застосуванням ШІ і бути її віддзеркаленням. У цьому контексті заслуговує на увагу концепція міжвидових (*cross-domain*) стандартів НАТО [Слюсар, 2018], яка передбачає, що такі стандарти могли б поєднувати в одному документі опис специфіки використання ШІ в інтересах сухопутних військ, військово-повітряних та військово-морських сил, наприклад у вигляді окремих розділів або додатків. Як варіант, частину стандартів можна застосовувати без змін в усіх перелічених видах військ, що також має бути спеціально застережено та схвалено основними групами Конференції національних директорів озброєнь, КНДО (*CNAD*). Такий підхід дозволить уникнути дублювання в стандартизації, гармонізує використання стандартів стосовно ШІ у різних видах збройних сил, дасть можливість скоординувати роботу експертних спільнот основних груп КНДО НАТО за спорідненими напрямами стандартизації технологій ШІ.

До ключових напрямів стандартизації ШІ у сфері безпеки і оборони належать:

- операційні сценарії та типові приклади використання (*use cases*) ШІ;
- мінімальні військові вимоги до систем зі штучним інтелектом;
- режими роботи систем зі штучним інтелектом;
- загальна архітектура, основні технічні характеристики систем зі штучним інтелектом та протоколи обміну великими даними;
- інтерфейс взаємодії користувача зі штучним інтелектом.

Якщо розглядати типові сценарії використання ШІ, то як прилад можна вказати, що на допомогу водію-механіку бронетехніки ШІ може виконувати низку типових функцій:

- попередження про можливість перекидання та визначення безпечного шляху;
- виявлення раптових загроз, які перешкоджають руху;
- візуальне оповіщення для маркування зон, які потребують особливої уваги;

- аналіз гіперспектральних зображень ґрунту для ідентифікації змін на його поверхні, що є ознакою штучного маскування саморобних вибухових пристроїв чи мін;
- ідентифікація камуфляжу на тлі природного ландшафту тощо.
- Подібні переліки мають бути створені для всіх можливих сфер військового використання ШІ.

Стосовно інтерфейсу важливо зауважити, що як засіб комунікації між штучним інтелектом та людиною доцільно розглядати технологію доповненої реальності (ДР), оскільки результати опрацювання інформації штучним інтелектом найзручніше донести оператору за допомогою візуальних, акустичних і тактильних символів ДР. У такий же спосіб доречно також ставити завдання системі ШІ, тим більше, що стандартизувати символи ДР значно легше, ніж достягти повної технічної сумісності систем різних виробників. Зокрема, зворотну взаємодію людини зі штучним інтелектом на основі доповненої реальності можна здійснювати шляхом призначення для ШІ зон, які підлягають аналізу, використання різних варіантів графічного інтерфейсу для введення вихідних даних, перетворення голосових повідомлень на команди переміщення тривимірних об’єктів ДР, їхньої орієнтації тощо.

Через ДР, зокрема через хмаринний сервіс, можна організувати взаємодію кількох систем ШІ між собою. Наприклад, спеціально призначена система ШІ може синтезувати тривимірну картину місцевості за контурними двовимірними зображеннями, отриманими з багаторакурсних знімків, зроблених рознесеними у просторі платформами ШІ. З метою максимальної реалізації потенціалу ДР як інтерфейсу систем ШІ важливо визначити вимоги до відповідної функціональності. Крім того, необхідна стандартизація дизайну символів ДР з метою взаємосумісності результатів роботи ШІ з операторами та іншими системами ШІ.

Слід підкреслити, що ШІ слід залучити до генерації контурних символів цілей у процесах цілевказування шляхом передачі лише оболонок цілей як символів ДР, які далі накладатимуться на реальні зображення навколишнього середовища. Все це потребує розгортання масштабних робіт щодо формування відповідних наборів даних. У такий же спосіб за допомогою ШІ слід формувати тривимірні символи доповненої реальності, здійснювати їх анімацію, забезпечувати мінімізацію ефектів оклюзії при візуалізації кольорових символів ДР на дисплеї. Важливо, що вибір оптимальної світності символів ДР для різних фонових об’єктів є доволі актуальним завданням. Залежно від позиції та просторової орієнтації поєднаних у мережу бойових платформ, навколишній фон може бути різним і досить часто збігається за яскравістю та кольором із символами ДР. Це призводить до часткової або повної втрати функціональності систем ДР. За допомогою ШІ цю проблему можна подолати шляхом адаптивного вибору кольору та яскравості символу ДР при накладанні його на фонову картину. Засоби ШІ повинні оцінювати фонову ситуацію і призначати оптимальній колір для символів ДР, вмикати динамічну зміну їхньої яскравості та кольору при візуалізації, або ж активувати їхню пульсацію (мерехтіння), обертання та інші ефекти анімації. У такий же спосіб використання ШІ дозволить запровадити допоміжну, напівпрозору кольорову підкладину, яка була б перехідним буфером між кольоровою палітрою фону та символом візуалізації даних. У цьому випадку адаптивний вибір сполучення кольорів допоміжної підкладини, фону та символів ДР слід також покласти на ШІ.

Алгоритми ШІ можуть створювати не лише контурні символи цілей, але й візуалізувати моделі їх уразливостей, які зараз використовуються для моделювання та симуляції. Ці візуалізовані моделі вразливості сегментують ворожі об’єкти на кілька зон для ураження, забезпечуючи більш ефективний їх вибір з метою досягнення максимальної ймовірності нейтралізації або руйнації цілі. Інформацію про такі зони можна розподілити як символи ДР між об’єднаними в мережу бойовими машинами всередині підрозділу для колективного обстрілу складних цілей. Рівень сегментації контурних символів ДР можна змінювати залежно від відстані до цілі, а стан такої декомпозиції слід використовувати як додаткову інформацію про поточну відстань до об’єкту вогневого впливу.

Доступні технології ШІ здатні забезпечити генерацію зображень за їх голосовим або текстовим описом, перетворювати текстові звіти та повідомлення в анотації та символи доповненої реальності або ж, за потреби, озвучувати їх чи трансформувати в аудіосимволи. У цьому контексті актуальним завданням загальнонаціонального рівня є створення мовної моделі української мови. Це дозволить створювати інтелектуальних помічників командирів, генерувати сценарії навчань з реалістичним наповненням їх сюжетів і послідовністю дій умовного противника, значно полегшить збір інформації з неструктурованих текстів тощо. У перспективі на цій основі можливий синтез ДР та синтетичного віртуального середовища за допомогою ШІ, що дозволить суттєво поліпшити якість тренувань і військових навчань.

Разом з тим, синергія між ШІ та автономією у військовій сфері створює низку проблем з безпеки застосування автономних систем озброєнь, заснованих на ШІ, якими переймаються не тільки експерти НАТО, але й всього світу. Зокрема, реальні загрози появи летальних автономних систем зі штучним інтелектом на полі бою та можливі внаслідок цього ризики для цивільного населення спонукали до перегляду пріоритетів у діяльності Управління Організації Об'єднаних Націй (ООН) з питань роззброєння (*UNODA*). Відповідно також було скореговано зміст діяльності Інституту ООН з дослідження проблем роззброєння (*UNIDIR*), а у складі *UNODA* було створено спеціальну Групу урядових експертів з питань летальних автономних систем (*GGE LAWS*) [Слюсар, 2021].

Під орудою *UNIDIR* було проведено серію командно-штабних навчань у кількох регіонах світу з різними сценаріями можливого застосування автономних систем озброєнь зі штучним інтелектом. Їх результати є важливими в контексті формування стратегічних підходів і тому заслуговують ретельного розгляду.

До навчань були залучені технічні, військові та юридичні експерти і дипломати. Розроблені сценарії розглядаються як інструмент для формування більш практичного і детального уявлення про відповідний оперативний і тактичний контекст. При цьому було використано два методологічні підходи, що стосувалися:

- процесу цілевказування (*targetting*) з різним рівнем залучення автономних систем озброєння;
- дослідження впливу на процес прийняття рішень різного рівня контролю з боку людини над автономними системами озброєнь залежно від типу цілей, географічних та інших умов, з урахуванням потенційних ризиків для цивільних осіб. Головну увагу приділяли не повністю автономним системам зі штучним інтелектом, у яких усі етапи націлювання здійснюються без будь-якого контролю з боку людини, а так званій «сірій» зоні, коли автономія застосовується для виконання обмеженої кількості конкретних завдань, що технічно можуть бути здійснені. При цьому було виявлено багато нових відкритих питань, що потребують розгляду та юридичного врегулювання.
- Основні припущення, які були покладені в основу навчань, такі:
- технологічні розробки розглядалися в еволюційному розвитку, оскільки існує невизначеність даних стосовно потенційних спроможностей революційних технологій (наприклад, у віддаленій перспективі впровадження квантових обчислень здійснить безпрецедентний вплив на розвиток ШІ), тому важливо було зосередитися на сучасному стані справ та реалістичній оцінці покрокового розвитку технологій у часі;
- експерти обмежилися розглядом виключно автономних систем озброєнь зі штучним інтелектом, які можуть бути розгорнуті у фізичному середовищі і здатні завдати кінетичного ефекту, тому кібернетична зброя лишилася за рамками досліджень; крім того, експертів не цікавили багатовимірні системи, які можна використовувати для розвідки і спостереження у військових і цивільних інтересах;
- з розгляду було вилучено системи підтримки прийняття рішень на основі технології ШІ, які вже можна використовувати з метою планування;

- навчання зосереджували на рішеннях щодо розгортання летальних автономних систем озброєння, при цьому політичні рішення щодо придбання або розробки таких технологій лишилися за рамками досліджень.

Аналітичною основою досліджень слугувала типова схема врахування людського чинника при прийнятті рішень щодо застосування сили, опублікована *UNIDIR* у 2020 році [Human element 2019]. Ця інфографіка є лише верхівкою айсберга, оскільки процес прийняття рішення щодо законного застосування сили є дуже складним і починається набагато раніше фактичного використання сили. Передусім, цей процес бере початок із наведеної внизу схеми політичного рішення щодо необхідності військового втручання, потім проходить кілька рівнів планування та оцінки і фактично доходить до того моменту, коли система озброєння розгортається на полі бою.

Незважаючи на загальне намагання прискорити проведення навчань, у ході них було застосовано саме такий широкий підхід. За рамками видимої частини айсберга лишилися важливі рішення та ключові параметри, які мають визначальний вплив на застосування сили. Наприклад, на стратегічному рівні це визначення правил ведення бойових дій і загальних цілей для ураження – зокрема які цілі допустимі для ураження, а які ні. На операційному рівні об'єкти ураження уточнюють, їх перелік стає більш детальним, його ретельно аналізують і затверджують. Особи, які приймають рішення, визначають найкращий тип систем озброєння для досягнення бажаного ефекту. Крім того, оцінюють супутні втрати з урахуванням усіх чинників. Усю цю інформацію передають далі на тактичний рівень, де забезпечується виконання місії з детальним плануванням усіх необхідних заходів. На всіх перелічених етапах першочергове значення має контекст із урахуванням параметрів, обставин і обмежень.

Процес ураження на тактичному рівні поділявся на 5 етапів:

1. Пошук цілей: навігація і маневр на полі бою для виявлення цілей, спираючись на наявну інформацію, розвідку та зібрані у реальному часі дані.
2. Фіксація (виявлення) цілей і взяття на супровід: після виявлення цілей сенсори використовують для визначення координат об'єктів ураження, подальшого моніторингу навколишнього середовища і забезпечення впевненості в тому, що позитивна ідентифікація цілей підтримується у просторі та часі.
3. Цілевказування: остаточна перевірка цілей перед ураженням з урахуванням різних форм оцінки ризиків і відповідності правилам ведення бойових дій, міжнародному гуманітарному праву.
4. Ураження: атаку здійснено із застосуванням зброї або призупинено чи скасовано.
5. Оцінка результатів: оцінювання ефективності атаки та прийняття рішення щодо подальших дій – у тому числі, за необхідності, проведення повторної атаки чи відслідковування або очікування прояву наслідків і перехід до пошуку нових цілей.

Для кожного з цих етапів було розглянуто 4 можливі рівні управління зброєю зі штучним інтелектом:

1. Повний прямий контроль.
2. Людина в контурі управління: система виконує поставлене завдання автономно, проте необхідне втручання людини для перевірки і виконання специфічних дій.
3. Людина над контуром управління: система виконує поставлене завдання автономно, але під наглядом людини-оператора, який може за потреби втручатися, корегувати або переривати специфічні дії.
4. Людина вилучена з контуру управління: система зі штучним інтелектом виконує поставлене завдання повністю автономно, без нагляду і втручання людини-оператора.

У ході навчань для кожного зі сценаріїв експертів запросили створити ідеальну, на їх погляд, конфігурацію управління, подавши її у вигляді таблиці (табл. 1). У наведеній версії комірки заповнені для прикладу. При цьому одиницею позначено наявність відповідного рівня контролю на тому чи іншому етапі, а його відсутність – нулем.

У ході навчань при заповненні таблиці технічні експерти зробили свій висновок, спираючись на власне розуміння технічної здійсненності застосування можливих рівнів контролю для різних завдань у різних контекстах. Військові експерти підійшли до цього питання більше з позицій військової доцільності або досягнення військової переваги. У той же час з правової точки зору відповідь на питання більше фокусувалася на тому, які юридичні наслідки або юридичні міркування щодо допустимості летальних дій будуть додатково застосовуватися у конкретному випадку.

Таблиця 1. Матриця конфігурації управління ШІ

Етап	Рівень контролю			
	Повний	Людина в контурі	Людина над контуром	Людина поза контуром
Пошук цілей	0	0	1	1
Виявлення і супровід	0	1	1	1
Цілевказування	1	1	1	0
Ураження	1	1	1	0
Оцінка результатів	1	1	1	1

Фахівці *UNIDIR* також запросили експертів визначити, які чинники вплинули на їхню оцінку, і в результаті зафіксували дуже широкий спектр категорій впливу – від типу цілі, середовища, в якому знаходиться ціль, до домену (на суші, на морі, в повітрі), типу місії і її параметрів. Крім того, час атаки, її місце та бажаний ефект впливали на оцінку ризиків для цивільного населення і своїх військ, а також на вибір технічних характеристик систем озброєння.

Під час навчань розглядали *4 типові сценарії бойових дій*.

1. Виявлені ворожі безекіпажні ракетні установки без військового або цивільного персоналу, споряджені і готові для стрільби. Бажаний ефект – руйнація пускових установок.
2. Ворожі БПЛА оснащені зброєю і активні цілодобово протягом тижня, останнє їхнє місцеперебування було відоме за 12 годин до початку навчань. Бажаний ефект – нейтралізація БПЛА.
3. Лінії комунікацій застосовуються ворожими силами для поповнення запасів озброєння і боєприпасів, дороги використовуються цивільними, поряд розташовані житлові будинки. Бажаний ефект – руйнація ліній комунікацій.
4. Ворожий конвой у русі, його позиція на дорозі невідома, дорога також використовується цивільними. Бажаний ефект – знищення конвою до досягнення ним меж міста.

Зазначені сценарії не були призначені для того, щоб охопити всі можливі випадки. Вони були лише інструментом, який мав стимулювати обговорення чогось більш конкретного і вимірюваного. Аналіз сценаріїв свідчить, що в двох із них місцеперебування цілей було відоме (сценарії 1 та 3), а в двох інших – невідоме. Було передбачено різний типаж цілей – фіксовані і мобільні, військові платформи, інфраструктура та конвой. Іншим важливим параметром сценаріїв був клас збитків, ризик супутніх збитків, потенційна участь цивільного населення.

Результати опитування трьох категорій експертів (технічні, військові та юридичні фахівці) для кожного із зазначених сценаріїв дали змогу зробити цікавий висновок, що коли експертам надають широкий спектр можливостей для контролю над автономною зброєю зі штучним інтелектом, більшість із них часто схиляється до якогось варіанта, який дозволив

би людям зберегти деяку форму контролю або участі. Експерти дуже рідко вибирають крайні випадки повного контролю чи вилучення людини з контуру управління. Звичайно, при цьому було виявлено певні регіональні аспекти та регіональні варіації. Не кожен регіон тим чи іншим чином відреагував так само, як інший, хоча, дивлячись на агреговані дані, слід вказати, що існує багато варіантів і ступенів відмінностей у рамках тих самих експертних спільнот у різних регіонах і між експертними співтовариствами. Тому дуже складно виділити одну конкретну тенденцію, яку можна застосувати виключно до однієї експертної спільноти.

Іншою важливою характеристикою є графічні залежності розподілу думок експертів стосовно допустимого рівня автономності зброї зі штучним інтелектом для кожного з етапів процесу ураження цілей на полі бою. Суттєво, що для кожного з етапів бойових дій що плоскіший характер має крива залежностей, то більше варіантів розбіжностей існує у поглядах експертів у рамках одної групи. Що більше піків видно на кривій, то більше експертів схиляються до відповідної спільної позиції. Аналіз свідчить, що насправді результати навчань дали змогу зафіксувати багато відмінностей у підходах між технічними та військовими експертами і правниками. Проте для сценарію нейтралізації БПЛА стосовно етапу пошуку цілей розподіл думок різних категорій експертів є дуже схожим і зводиться переважно до доцільності повної автономності систем пошуку.

Важливо також урахувати, що багато експертів підкреслили можливість компромісних рішень, коли рівень автономності, делегований конкретній системі зі штучним інтелектом, може змінюватися залежно від того, яка автономність передбачена для подальших етапів або передувала поточній ситуації. Більш детальний аналіз отриманих даних викладено в офіційному звіті *UNIDIR*.

За результатами навчань було сформульовано найважливіші технічні міркування. Зокрема, ознака «автономний» не є синонімом понять «автоматизований» (automated) і «автоматичний» (automatic). Різницю між цими трьома концепціями можна проілюструвати на прикладі втрати зв’язку між наземною станцією управління і дроном.

Коли цей зв’язок розривається, для досягнення компромісу спрацьовує запрограмований автоматичний тригер, який спонукає дрон зробити тайм-аут, наприклад на одну хвилину, і затриматися на місці, очікуючи, поки канал зв’язку не буде відновлено. Якщо після закінчення тайм-ауту зв’язок не відновлюється, то автоматично запускається зворотний відлік часу, який ініціює автоматичний процес польоту в інше місце, щоб дочекатися встановлення зв’язку. Якщо навіть цей процес зазнає невдачі, є інший тригер, який автоматично спрацьовує і дрон летить в означену зону для виконання контрольованої посадки. Це все є прикладами автоматичних тригерів, які використовує система для реалізації автоматизованих процесів, замість того, щоб повідомляти дрону, куди йти, щоб чекати сигналу, або де приземлитися у випадку, якщо сигнал пропадає, скажімо, більш ніж на годину. Якщо всі ці дії сплановано заздалегідь, автоматизована система буде виконувати їх в автоматичному режимі.

Прикладом автономії є ситуація, коли у разі надходження повідомлення про втрату зв’язку дрону визначають лише параметри безпечного місця приземлення. Наприклад, це можуть бути вимоги віддалення такого місця не менше, ніж на 50 км від району операції, щоб поблизу не було житла і виконувалась низка інших обмежень. При цьому системі дозволено самій визначити краще місце для автономного приземлення. Таким чином, складне завдання, що ставиться машині, запускає більш складний процес прийняття рішень і аналізу даних, який система повинна робити, щоб керувати своєю поведінкою.

Технологія ШІ забезпечує автоматизацію і автономність, але вона не автономна сама по собі. Якщо розглядати автономію як заключну ланку розвитку спроможностей прийняття рішень людиною, то технологія ШІ найбільш придатна для автономії, але не обов’язково єдина. Знову ж таки, слід визнати, що спроможність ШІ перевершувати людей сьогодні обмежується конкретними завданнями з розпізнавання об’єктів. Відповідні технології ШІ більш просунуті, ніж розуміння поведінки. Тим не менше, все ще існують значні обмеження, коли справа доходить до реального використання ШІ, який добре працює в контрольованих

середовищах або лабораторіях. Відповідна проблема має назву «моделювання реального розриву». Значний акцент роблять також на нездатності ШІ до узагальнень і перерозподілу продуктивності між різними завданнями залежно від їх контексту, оскільки ШІ здатен добре виконувати конкретне завдання в дуже специфічному контексті. Це означає, що той же самий алгоритм не буде однаково ефективно працювати для інших завдань.

Багато експертів висловили думку, що впровадження автономії буде поступовим і обмежиться тими завданнями й функціями, які продемонструють, що ШІ зможе задовольнити вимоги передбачуваності та надійності. Остаточне рішення лишається за кінцевим користувачем. Це не виключає, що деякі нерозсудливі актори вирішать інтегрувати ШІ у критичні ролі, як тільки вони це зможуть. Але якщо виключити такі випадки, експерти погоджуються, що це буде, швидше за все, поступове введення.

Слід підкреслити важливість даних, які вважають ключовими для навчання, тестування і розгортання ШІ. Тестування та оцінка точності є фундаментальною частиною процесу встановлення або калібрування довіри між людьми-операторами та різними системами ШІ. Але зараз не вистачає ресурсів, через що велику частину талантів і значну частину коштів задіяно на розробку нових типів, нових моделей і нових засобів ШІ, а не у вивчення, тестування та перевірку наявних рішень на основі ШІ.

Звичайно, є деякі очевидні переваги введення ШІ і при виконанні конкретних завдань. На думку військових експертів, передбачуваними перевагами застосування ШІ під час місій є швидкість, точність, ефективність ресурсного менеджменту, здатність діяти в умовах відсутності зв’язку. Проте навіть за таких переваг доцільність застосування систем ШІ буде залежати від контексту й параметрів місій. У будь-якому випадку для військових експертів не було ніякого стимулу припустити, що збройні сили будуть зацікавлені в розгортанні автономних систем озброєння, які вони не можуть контролювати за допомогою більш цілісного контролю, особливо при розгляді критичних функцій. Наявність технологічних рішень не означає, що вони будуть відразу інтегровані у війська. Існує ціла структура, яка повинна бути розроблена, щоб пов’язати технологію з військовим потенціалом, і вона включає в себе доктрини, організаційні структури, навчання і под. (*DOTMLPFI*). Немає причин вважати, що можливості технологій ШІ повинні скоротити будь-який з цих важливих кроків. У всякому разі, вони зроблять їх ще більш актуальними і важливими.

Крім того, використання автономних систем зброї ніяк не знижує юридичну відповідальність за наслідки для тих осіб, які приймають рішення. Проте деякі експерти вказали на те, що можуть виникнути нові проблеми, пов’язані з неправильним призначенням відповідальності людини або кримінальної відповідальності за використання зброї, що виникла в результаті помилок безекіпажних систем зі штучним інтелектом.

Варто зауважити, що організатори навчань спростили опис сценаріїв заради економії часу. Для відповідності цілям досліджень сценарій більш високого рівня був достатнім з точки зору того, які елементи принципів міжнародного гуманітарного права виконуються. У будь-якому випадку сценарії не були занадто складними, проте вони дозволили зберегти необхідний рівень достовірності. Разом з тим, це тільки початок і в майбутньому можливо буде запускати набагато складніші і більш реальні сценарії.

Усі задіяні експертні групи були з різних причин стурбовані якістю даних, якими оперує автономна система зі штучним інтелектом, проте найбільше цим переймалося технічне співтовариство, яке вважає дані ключовим фактором для тестування та розгортання автономних систем озброєння. Військове співтовариство було стурбоване проблемою даних, оскільки воно знає, що від їхньої якості може явно залежати продуктивність системи ШІ. Військовим експертам відома важливість тестування і перевірки машини тими ж оперативними даними, які потім будуть використовуватися, коли система розгортається на полі бою.

І, звичайно, юридична спільнота, залежно від регіонів, більш-менш усвідомлювала юридичну важливість даних. Проте серед неї не було розуміння того, що зрештою автономна система зброї зі штучним інтелектом взаємодіятиме з даними, і якщо ці дані будуть

неоптимальними або відрізнятимуться від тих, для яких система була підготовлена, то можуть виникати додаткові питання, що стосуються передбачуваності та надійності системи.

Навпаки, військові експерти висловлювали побоювання з приводу надійності систем ШІ. Вони підкреслюють першочергову важливість міжнародного гуманітарного права і, зокрема, відносну важливість принципів, які необхідно буде прив'язати до конкретного сценарію і конкретного виду зброї.

Відповідні дослідження щодо застосування автономних систем озброєнь зі штучним інтелектом будуть продовжені і в подальші роки з урахуванням упровадження нових технологій. З огляду на це, важливо забезпечити участь представників України у роботі КНДО НАТО та Групи урядових експертів з питань летальних автономних систем (ГУЕ ЛАСО) Управління ООН з питань роззброєння. Ефективним методом співробітництва слід вважати активну участь у дискусіях та виступи з коментарями щодо порушених питань з урахуванням творчого використання досвіду національних експертів для впровадження новітніх ідей у відповідні нормативні документи.

Стрижак О. Є.

До розділу «Вступ»

Економічний розвиток будь-якої країни у XXI-му столітті повністю залежить від рівня представлення на світовому ринку систем знань, які певним чином циркулюють в усіх ланках соціально-економічних відношень у суспільстві. Ключовим проявом цього явища є становлення ***економіки знань*** (англ. *knowledge economy*), яка формується на основі міждисциплінарно пов`язаних між собою процесів створення, опрацювання, зберігання, поширення та використання знань. Це зумовлює те, що когнітивно-комунікативні сценарії взаємодії в усіх сферах соціально-економічної діяльності держави також цілком залежать від її спроможності ефективно опрацьовувати наявні знання та комплексно використовувати вже накопичені інформаційні ресурси. Але інформаційні ресурси, які репрезентують системи знань за сукупністю та характером викладу, належать до класу великих даних (англ. *big data*). Усі вони також характеризуються багатоаспектністю, множинними латентними зв'язками тощо.

Вирішення проблем розвитку та ефективного використання систем знань у різних галузях людської діяльності, як свідчить світовий досвід, лежить у площині застосування сучасних інформаційних технологій, що реалізуються на засадах штучного інтелекту як однієї з ключових технологій сучасності.

Агентство передових оборонних дослідницьких проєктів США (DARPA) визначило XXI століття початком ери трансдисциплінарних досліджень. Цілком забезпечити реалізацію цього меганапряму наукового й науково-технічного розвитку як складової економіки знань можливо на засадах використання усіх засобів штучного інтелекту. При цьому забезпечується формування логічних метарамок, за допомогою яких знання, що відображають результати трансдисциплінарних досліджень, можуть бути консолідовано інтегровані у різні галузеві напрямки розвитку інформаційного суспільства й, як наслідок, сприятимуть розбудові економіки знань.

Такі особливості сучасного етапу становлення суспільства знань зумовлюють актуальність проблеми створення інтелектуальних інструментів і засобів, спроможних узяти на себе принаймні частину основних когнітивних функцій людини. Тому подальший розвиток економіки знань, особливо в нашій країні, практично цілком залежить від того, наскільки ефективно буде реалізовано й використано досягнення в сфері інформаційних технологій і штучного інтелекту.

Висновки

Таким чином, Стратегія розвитку штучного інтелекту в Україні є важливим документом, який регулює основні напрями проведення фундаментальних досліджень і отримання нових знань для створення проривних технологій у цій сфері з урахуванням українських реалій. Вважаємо прийняття такого документа важливою державною справою.

Література

1. Шевченко А. І. До питання щодо створення штучного інтелекту. Штучний інтелект, № 1, 2016. С. 7-15.
2. Stanley-Lockman, Zoe, and Christis, Edward Hunter. 2021. An artificial intelligence strategy for NATO. 25 October 2021. URL: https://www.nato.int/docu/review/articles/2021/10/25/an-artificial-intelligence-strategy-for-nato/index.html.
3. Концепція розвитку штучного інтелекту в Україні. 2020. URL: https://zakon.rada.gov.ua/laws/show/1556-2020-%D1%80#n8.
4. Вашкевич А. Електронна особа. 2016. URL: https://zbruc.eu/node/51750.
5. Каткова Т. Г. Закони про роботів: сучасний стан і перспективи розвитку. 2017. URL: http://aphd.ua/publication-345/.
6. Слюсар В. І. Щодо стратегії формування системи стандартів НАТО // Зб. матеріалів V міжнародної науково-практичної конференції “Проблеми координації воєнно-технічної та оборонно-промислової політики в Україні. Перспективи розвитку озброєння та військової техніки”. Київ, 11-12 жовтня 2017 р. С. 84–86.
7. Слюсар В. И. Концепция межвидовых стандартов НАТО // Тези доповідей 14-ї наукової конференції “Новітні технології – для захисту повітряного простору”, 11-12 квітня 2018 року. Харків: ХНУПС. С. 46–47.
8. Слюсар В. И. Роль искусственного интеллекта в кросс-платформенном распределении данных дополненной реальности // Зб. матеріалів VIII міжнародної науково-практичної конференції “Проблеми координації воєнно-технічної та оборонно-промислової політики в Україні. Перспективи розвитку озброєння та військової техніки”. Київ, 2020. С. 417–420.
9. Слюсар В. И. Концепция виртуализации поля боя 2050 года // Озброєння та військова техніка. 2021. № 3 (31). С. 111–112.
10. Calo, Ryan. 2016. Robots in American law. Legal studies reserch paper No. 2016-04. 44 p.
11. Dash, S., Shakyawar, S.K., Sharma, M. et al. 2019. Big data in healthcare: management, analysis and future prospects. J Big Data 6, p. 54. https://doi.org/10.1186/s40537-019-0217-0.
12. Dehaene, S., Lau, H., and Kouider, S. 2017. What is consciousness, and could machines have it? In: Science 358, pp. 486-492.
13. Graziano, M. 2017. The attention schema theory: A foundation for engineering artificial consciousness. In: Frontiers in Robotics and AI 4, art. 60, pp. 1-9.
14. Graziano, M., and Webb, T. 2017. Understanding consciousness by building it. Part three: Metaphilosophy of consciousness studies. In: Bloomsbury companion to the philosophy of Consciousness, pp. 185-210.
15. Gröne, O., and Garcia-Barbero, M. 2002. Trends in integrated care – Reflections on conceptual issues (EUR/02/5037864). Copenhagen: World Health Organization.
16. Hawking, Stephen. 2016. Automation and AI is going to decimate middle class jobs. URL: http://www.businessinsider.com/stephen-hawking-ai-automation-middle-class-jobs-most-dangerous-momenthumanity-2016-12.
17. Hennessy, J. L., and Patterson, D. A. 2019. A new golden age for computer architecture. In: Communications of the ACM, vol. 62, issue 2, pp. 48–60.
18. The human element in decisions about the use of force. 2019. URL: https://www.unidir.org/publication/human-element-decisions-about-use-force.
19. Ng, Andrew. 2016. What artificial intelligence can and can’t do right now. Harward business review. URL: https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now.
20. Seth, A. K., Baars, B. J., and Edelman, D. B. 2005. Criteria for consciousness in humans and other mammals. In: Consciousness and cognition 14, pp. 119-139.
21. Stanley-Lockman, Zoe, and Christis, Edward Hunter. 2021. An artificial intelligence strategy for NATO. 25 October 2021. URL: https://www.nato.int/docu/review/articles/2021/10/25/an-artificial-intelligence-strategy-for-nato/index.html.

Стаття надійшла до редакції 07.05.22
Після обробки 05.06.22