



Rada
Európskej únie

V Bruseli 26. októbra 2022
(OR. en)

14095/22
ADD 1

LIMITE

CORLX 995
CFSP/PESC 1432
CONUN 254
CODUN 48
CONOP 112
COTER 256
COARM 217

NÁVRH

Od:	Stefano SANNINO, generálny tajomník, v zastúpení vysokého predstaviteľa Únie pre zahraničné veci a bezpečnostnú politiku
Dátum doručenia:	26. októbra 2022
Komu:	pani Thérèse BLANCHETOVÁ, generálna tajomníčka Rady Európskej únie
Predmet:	Návrh rozhodnutia Rady na podporu vykonávania projektu s názvom Podpora zodpovednej inovácie v oblasti umelej inteligencie v záujme mieru a bezpečnosti, ktorý Rade predložil vysoký predstaviteľ Únie pre zahraničné veci a bezpečnostnú politiku

Delegáciám v prílohe zasielame dokument HR(2022) 238.

Príloha: HR(2022) 238

EURÓPSKA SLUŽBA PRE VONKAJŠIU ČINNOSŤ



PROJEKTOVÝ DOKUMENT

Podpora zodpovednej inovácie v oblasti umelej inteligencie v záujme mieru a bezpečnosti

HR(2022)238

1. KONTEXT

Nedávny pokrok v oblasti umelej inteligencie priniesol nové možnosti na podporu a udržanie mieru a bezpečnosti, napríklad prostredníctvom technologických zlepšení v oblastiach, ako je včasné varovanie v prípade konfliktov a overovanie kontroly zbrojenia a vývozu. Tento pokrok však na druhej strane priniesol nové prostriedky, ktorými sa prispieva k vzniku alebo zhoršovaniu napätia, konfliktov a neistoty medzi štátmi a v rámci nich. Riziká, ktoré predstavujú určité aplikácie umelej inteligencie, ako sú smrtiace autonómne zbraňové systémy, sú z hľadiska komunity pre kontrolu zbrojenia veľkým problémom. Jednou z rizikových možností, ktorá si zasluhuje väčšiu pozornosť a na ktorú môže byť súčasná kontrola zbrojenia a diplomatické úsilie len nedostatočnou reakciou, je odklon a zneužitie výskumu a inovácie v oblasti civilnej umelej inteligencie nezodpovednými aktérmi vrátane zlomyseľných neštátnych subjektov, ako je zneužívanie generatívnych súperiáciich sietí (GAN) na vytváranie „deepfake“ pre dezinformačné kampane.

Umelá inteligencia je podpornou technológiou s veľkým potenciálom všeobecného využitia. Určité subjekty by mohli mať (relatívne jednoduchý) prístup k výskumu a inovácií v oblasti umelej inteligencie, ktoré sa vyvíjajú pre civilné aplikácie, a mohli by zmeniť ich účel na škodlivé alebo disruptívne využitie, čo by mohlo mať dôsledky pre medzinárodný mier a bezpečnosť. Odklon a zneužitie civilných technológií nie je novým javom, ani sa neobmedzuje na umelú inteligenciu. Príkladom toho je v súvisiacej oblasti robotiky nedávne využívanie rekreačných dronov ako zbraní zo strany Dá'iš/ISIS v Sýrii. V oblasti umelej inteligencie však tento problém komplikuje viacero faktorov: nehmotná a rýchle sa meniaci povaha algoritmov a dát umelej inteligencie, ktorá sťažuje kontrolu ich transferu/šírenia; vedúca úloha súkromného sektora v ekosystéme výskumu, vývoja a inovácie a jeho následná potreba chrániť vlastné algoritmy a celosvetová dostupnosť ľudí s odbornými znalosťami a materiálnymi zdrojmi schopnými zmeniť účel technológií umelej inteligencie. Pracovníci v oblasti umelej inteligencie v civilnom sektore si okrem toho príliš často neuvedomujú potenciálne dôsledky, ktoré by mohli mať odklon a zneužitie ich práce pre medzinárodný mier a bezpečnosť, alebo váhajú s účasťou

na existujúcich diskusiách o rizikách spojených s umelou inteligenciou v kruhoch, ktoré sa zaoberajú kontrolou zbrojenia a nešírením zbraní.

Je potrebné podporiť väčšiu angažovanosť komunity civilnej umelej inteligencie na úrovni chápania a zmierňovania rizík pre mier a bezpečnosť spojených s odklonom a zneužitím technológií civilnej umelej inteligencie nezodpovednými aktérmi. V prípade Štokholmského medzinárodného inštitútu pre výskum mieru (SIPRI) a Úradu Organizácie Spojených národov pre otázky odzbrojenia (UNODA) by sa to mohlo dosiahnuť väčšou podporou zodpovednej inovácie v rámci celosvetovej komunity civilnej umelej inteligencie. Minulá práca SIPRI a UNODA ukázala, že zodpovedná inovácia ako mechanizmus vlastného riadenia by mohla poskytnúť celosvetovej komunite civilnej umelej inteligencie praktické nástroje a metódy na identifikáciu rizík, ktoré by odklon a zneužívanie výskumu a inovácie v oblasti civilnej umelej inteligencie mohli predstavovať pre mier a bezpečnosť, a pomoc pri predchádzaní týmto rizikám a ich zmierňovaní. V rámci práce SIPRI a UNODA sa zároveň určili metodiky a niekoľko prebiehajúcich iniciatív zameraných na civilné osoby v oblasti zodpovednej umelej inteligencie, na ktorých by sa mohlo stavať s cieľom zvyšovať informovanosť komunity civilnej umelej inteligencie v otázkach kontroly zbrojenia a nešírenia zbraní, o odborných diskusiách a stanoviskách štátov k zodpovednému vývoju, šíreniu a využívaniu umelej inteligencie, ako aj o poznatkoch získaných z práce obranného sektora v oblasti zodpovednosti¹. Predovšetkým sa v rámci tejto predchádzajúcej práce ako kľúčový prvok každého účinného úsilia v oblasti zodpovednej inovácie jasne identifikovala spolupráca so študentmi vedy, technológie, inžinierstva a matematiky (STEM), ktorí s umelou inteligenciou pracujú stále len vo vzdelávacom formáte.

2. CIELE

Cieľom týchto projektov je podporiť väčšiu angažovanosť komunity civilnej umelej inteligencie pri zmierňovaní rizík, ktoré môžu pre medzinárodný mier a bezpečnosť vyplynúť z odklonu a zneužitia výskumu a inovácie v oblasti civilnej umelej inteligencie nezodpovednými aktérmi. Na dosiahnutie tohto cieľa sa projekty zamerajú v prvom rade na lepšie pochopenie toho, ako môžu rozhodnutia v rámci vývoja a šírenia výskumu a inovácie v oblasti umelej inteligencie ovplyvniť riziká odklonu a zneužitia, a tým vytvárať hrozby alebo príležitosti v oblasti mieru a bezpečnosti, a v druhom rade na podporu procesov, metód a nástrojov zodpovednej inovácie, ktoré môžu pomôcť zabezpečiť mierové uplatňovanie civilných inovácií a zodpovedné šírenie poznatkov v oblasti umelej inteligencie. Budú sa nimi preto podporovať činnosti zamerané na budovanie kapacít, výskum a angažovanosť, ktoré *i)* posilnia schopnosť celosvetovej komunity civilnej umelej inteligencie zohľadňovať a riešiť, prostredníctvom zodpovedných inovačných procesov, riziká pre mier a bezpečnosť, ktoré vyplývajú z odklonu a zneužitia civilnej umelej inteligencie nezodpovednými aktérmi, a *ii)* posilnia prepojenie medzi úsilím

¹ Medzi metodiky patria napríklad odporúčané postupy Inštitútu elektrotechnických a elektronických inžinierov (IEEE) na posudzovanie vplyvu autonómnych a inteligentných systémov na dobré životné podmienky ľudí (IEEE Std 7010-2020), a Zoznam bodov na posudzovanie dôveryhodnej umelej inteligencie (ALTAI), ktorý vypracovala expertná skupina na vysokej úrovni pre umelú inteligenciu. Medzi tieto iniciatívy patrí: Globálna iniciatíva IEEE pre etiku autonómnych a inteligentných systémov, Partnerstvo v oblasti umelej inteligencie a Globálne partnerstvo pre umelú inteligenciu.

o zmiernenie rizika v oblasti zodpovednej umelej inteligencie v civilnej sfére a úsilím, ktoré už prebieha v rámci komunity pre odzbrojenie, kontrolu zbrojenia a nešírenie zbraní na medzivládnej úrovni. Zásadným prvkom je, že ich cieľom nie je stanoviť nové normy, zásady či predpisy ani iným spôsobom zasahovať do oblastí, ktoré patria do právomoci štátov. Namiesto toho je ich cieľom rozvíjať zodpovedné civilné úsilie v oblasti inovácie, ktoré sa zameria aj na riziká, ktoré pre mier a bezpečnosť vyplývajú z odklonu a zneužitia civilnej umelej inteligencie nezodpovednými aktérmi, a poskytovať informácie o existujúcom relevantnom medzivládnom úsilí.

S cieľom účinne osloviť a ovplyvniť komunitu civilnej umelej inteligencie sa v rámci projektov uplatňuje trojstupňový prístup, ktorý sa zameriava na:

- a) spoluprácu s pedagógmi – pracovať s vybranými pedagógmi a zostavovateľmi učebných osnov na vypracovaní a propagácii vzdelávacích materiálov, ktoré možno použiť na zohľadnenie rizík, ktoré pre mier a bezpečnosť vyplývajú z odklonu a zneužitia výskumu a inovácie v oblasti civilnej umelej inteligencie nezodpovednými aktérmi, pri odbornej príprave budúcich odborníkov aktívnych v oblasti umelej inteligencie (napr. na kurzoch o etike umelej inteligencie a zodpovednej inovácii);
- b) spoluprácu so študentmi – oboznámiť vybraných študentov vedy, technológie, inžinierstva a matematiky (STEM) z celého sveta s tým, ako možno identifikovať riziká, ktoré pre mier a bezpečnosť predstavuje odklon a zneužitie vývoja civilnej umelej inteligencie nezodpovednými aktérmi, a ako možno týmto rizikám predchádzať a zmierňovať ich v rámci procesu výskumu a inovácie alebo prostredníctvom iných procesov riadenia, a
- c) spoluprácu s odvetvím umelej inteligencie – pracovať s profesijnými komorami a normalizačnými orgánmi, ako je Inštitút elektrotechnických a elektronických inžinierov (IEEE), s cieľom: i) šíriť prispôsobené vzdelávacie materiály a činnosti zamerané na zapojenie technických odborníkov; ii) podporovať pozitívne využívanie umelej inteligencie v záujme mieru a bezpečnosti a iii) uľahčiť dialóg a výmenu informácií medzi odborníkmi z akademickej obce, súkromného sektora a verejnej správy o tom, ako možno zmierniť riziko odklonu a zneužitia výskumu a inovácie v oblasti civilnej umelej inteligencie nezodpovednými aktérmi.

Takýto prístup umožňuje, aby sa projekty dostali do komunity umelej inteligencie na všetkých úrovniach, a to nielen medzi v súčasnosti aktívnych odborníkov, ale aj ich budúce generácie. Umožňuje tiež zapojenie akademickej obce, odvetvia a iných uzavretých skupín a podporuje udržateľnosť budúceho úsilia vytváraním sietí, ktoré prekračujú tieto hranice.

Cieľom projektov je tiež využiť vplyv a skúsenosti SIPRI a UNODA, a tak ovplyvniť komunitu umelej inteligencie na celom svete, nielen zainteresované strany v EÚ. SIPRI a UNODA majú jedinečné postavenie na to, aby oslovili aktérov v oblasti umelej inteligencie z Afriky, Ázie a Tichomoria, Európy a Severnej a Južnej Ameriky a uľahčili spoluprácu medzi nimi. Oba subjekty zároveň majú skúsenosti s prácou v iných oblastiach vedy a techniky, ktoré čelia podobným výzvam v oblasti dvojakého použitia a šírenia, a to vrátane biotechnológií. Ďalším cieľom projektov je využiť podmienky existujúce v Európskej únii, ako napríklad

a) pokročilé procesy so zapojením viacerých zainteresovaných strán v oblasti zodpovednej umelej inteligencie; b) vysoká miera angažovanosti a odborných znalostí v otázkach odzbrojenia, kontroly zbrojenia a nešírenia zbraní v EÚ; c) rozmanitosť väzieb, ktoré majú akademické a výskumné organizácie a organizácie súkromného sektora v EÚ s inými regiónmi, najmä na globálnom Juhu, ktorý bude zároveň hlavným cieľom spolupráce, a d) rozmanitá štátna príslušnosť študentov, pedagógov a inžinierov na univerzitách, vo výskumných inštitúciách a v súkromnom sektore.

Začlenenie bude kľúčovým hľadiskom pri vykonávaní činností projektov. S cieľom účinne podporovať komunitu umelej inteligencie sa v projektoch uznáva, že túto komunitu tvorí rôznorodá škála aktérov a že konkrétne rod je jedným z dôležitých faktorov.

- a) Z tohto dôvodu sa bude uplatňovať rodové hľadisko v súlade so všeobecnými stratégiami OSN pre uplatňovanie hľadiska rodovej rovnosti a rovnocennosti. V rámci všetkých činností každého projektu sa bude podporovať a vyžadovať účasť žien, ako aj
- b) začlenenie osôb so zdravotným postihnutím, pričom sa zabezpečí primerané prispôsobenie ich potrebám. To bude zahŕňať odstraňovanie prekážok účasti osôb so zdravotným postihnutím a podniknutie krokov na vypočutie a zohľadnenie vecných názorov a skúseností osôb so zdravotným postihnutím.

3. PROJEKTY

Nasledujúce tri projekty majú byť komplementárne a vzájomne sa podporovať počas 36 mesiacov trvania ich jednotlivých prvkov.

Projekt 1 – Vypracovanie vzdelávacích materiálov a materiálov na budovanie kapacít pre komunitu civilnej umelej inteligencie

3.1.Účel projektu

Projekt 1 sa zameriava na poskytnutie znalostí a prostriedkov aktérom v oblasti civilnej umelej inteligencie, aby mohli vyhodnotiť riziká, ktoré môžu pre medzinárodný mier a bezpečnosť vyplývať z odklonu a zneužitia výskumu a inovácie v oblasti civilnej umelej inteligencie nezodpovednými aktérmi, a zmierňovať ich. Jeho cieľom je vypracovať vzdelávacie materiály a materiály na budovanie kapacít, ktoré poskytnú odborníkom v oblasti umelej inteligencie zo všetkých regiónov, úrovni a odvetví (vrátane pedagógov zameraných na umelú inteligenciu, zostavovateľov učebných osnov, študentov STEM a inžinierov a výskumných pracovníkov v oblasti umelej inteligencie v akademickej obci a súkromnom sektore) informácie a nástroje potrebné na

- a) pochopenie, ako by sa výskum a inovácia v oblasti civilnej umelej inteligencie mohli odkloniť a zneužiť tak, že by mohli predstavovať riziko pre medzinárodný mier a bezpečnosť, a ako môžu rozhodnutia pri vývoji a šírení výskumu a inovácie zvýšiť alebo znížiť riziko takéhoto odklonu a zneužitia;

- b) pochopenie úsilia, ktoré už vyvinula komunita pre odzbrojenie, kontrolu zbrojenia a nešírenie zbraní s cieľom zmierniť riziká odklonu a zneužitia civilného výskumu a inovácie, a
- c) realizáciu zodpovednej inovácie spôsobom, ktorý zmiernuje riziko odklonu a zneužitia pri vývoji a šírení výskumu a inovácie.

3.1.1. Opis projektu

V rámci tohto projektu sa vypracujú tri samostatné súbory vzdelávacích materiálov a materiálov na budovanie kapacít.

- a) *Príručka (1)* – V príručke sa zhromaždia základné znalosti a prostriedky, aby aktéri v oblasti umelej inteligencie mohli v rámci procesu výskumu a inovácie vyhodnotiť riziká odklonu a zneužitia technológií v oblasti civilnej umelej inteligencie nezodpovednými aktérmi a znížiť ich. Bude obsahovať odpovede na otázky prečo a ako môžu rozhodnutia počas vývoja a šírenia výskumu a inovácie v oblasti umelej inteligencie ovplyvniť riziká odklonu a zneužitia, a tým vytvárať riziká alebo príležitosti pre mier a bezpečnosť. Okrem aspektov bezpečnosti a ochrany, o ktorých sa diskutuje vo vojenských kruhoch, ako aj v kruhoch, ktoré sa zaoberajú odzbrojením, kontrolou zbrojenia a nešírením zbraní, sa v nej uvedú aj príslušné záväzky vyplývajúce z medzinárodného práva a povinnosti týkajúce sa kontroly vývozu. Predstaví sa v nej tiež príklady procesov a nástrojov na realizáciu zodpovednej inovácie, ako sú metodiky posudzovania vplyvu technológií a vzory na posudzovanie rizík.
- b) *Séria podcastov (cca 10)* – Tieto podcasty budú pre aktérov v oblasti umelej inteligencie prístupným a zaujímavým médium, z ktorého sa dozvedia, prečo a ako možno procesmi zodpovednej inovácie v oblasti umelej inteligencie podporiť medzinárodný mier a bezpečnosť tým, že sa zmiernia riziká, ktoré predstavuje odklon a zneužitie umelej inteligencie na civilné využitie nezodpovednými aktérmi. V podcastoch sa bude pozornosť venovať dôležitým témam (napr. modelom odklonu a zneužitia výskumu a inovácie s dvojakým použitím/všeobecným využitím; humanitárnym, strategickým a politickým výzvam spojeným s výskumom a inováciou v oblasti civilnej umelej inteligencie a ich potenciálnym zneužitím; výzvam, ktorým v rámci úsilia o zmierňovanie rizík čelia kruhy zaoberajúce sa odzbrojením, kontrolou zbraní a nešírením zbraní; otázkam, ako zabezpečiť zodpovednú inováciu prostredníctvom posudzovania rizika; dodržiavania pravidiel kontroly vývozu; znížovaniu rizika už v štádiu návrhu; zodpovednému publikovaniu; významu poznať svojich zákazníkov a skúsenostiam so simulačnými cvičeniami), pričom budú mať formu rozhovorov medzi projektovým tímom a zástupcami príslušných komunít.
- c) *Séria blogov (9 až 10)* – Tím vypracuje sériu organizovaných blogov, ktorých cieľom je zviditeľniť úsilie o zblíženie „zodpovednej umelej inteligencie“ zameranej na civilné osoby s komunitami pre kontrolu zbrojenia a nešírenie zbraní. Séria blogov poskytne platformu na šírenie poznatkov, nápadov a riešení, pokiaľ ide o identifikáciu a odstraňovanie rizík spojených s odklonom a zneužitím civilnej umelej inteligencie

v procese výskumu a inovácie. Cieľom blogov bude prezentovať rozmanitosť myslenia a perspektív v odvetví umelej inteligencie.

Tieto materiály sa budú verejne šíriť prostredníctvom webových sídel vykonávajúcich aktérov, ich prítomnosti v sociálnych médiách a prostredníctvom priamej komunikácie s príslušnými akademickými subjektmi, profesijnými komorami v oblasti civilnej umelej inteligencie a inými vhodnými skupinami.

3.1.2. Očakávané výsledky projektu

V rámci tohto projektu sa očakáva vytvorenie nového súboru materiálov, pomocou ktorých by odborníci v oblasti civilnej umelej inteligencie mohli získať informácie o a) spôsobe, akým by sa výskum a inovácia v oblasti civilnej umelej inteligencie mohli odkloniť a zneužiť tak, že by mohli ohroziť medzinárodný mier a bezpečnosť, b) ako sa takéto riziká odstraňujú v rámci komunity pre odzbrojenie, kontrolu zbrojenia a nešírenie zbraní a c) ako by odborníci aktívni v oblasti umelej inteligencie mohli ďalej prispievať k zmierňovaniu takýchto rizík prostredníctvom procesov zodpovednej inovácie.

Očakáva sa, že sa tým podporí väčšia angažovanosť komunity civilnej umelej inteligencie pri zmierňovaní rizík, ktoré môže odklon a zneužívanie civilnej umelej inteligencie predstavovať pre medzinárodný mier a bezpečnosť; zlepši sa schopnosť technických odborníkov zapájať sa do príslušných procesov v rámci spolupráce s komunitou pre odzbrojenie, kontrolu zbrojenia a nešírenie zbraní a podporí sa interakcia s novým publikom, ktoré nie je tradične zahrnuté do úsilia o vzdelávanie v oblasti odzbrojenia a nešírenia zbraní.

Zároveň by sa prostredníctvom tohto materiálu mala podporiť aj realizácia ostatných projektov a mal by slúžiť ako základ pre činnosti v oblasti vzdelávania a budovania kapacít v rámci projektu 2, ako aj pre dialóg a činnosti v oblasti angažovanosti v rámci projektu 3. Očakáva sa, že tieto činnosti následne poskytnú spätnú väzbu na prepracovanie a zlepšenie tohto materiálu. Takýto opakovaný prístup by mohol prispieť k odstraňovaniu možných prekážok brániacich propagácii, šíreniu a používaniu materiálu v rámci komunity umelej inteligencie, a to aj v súvislosti s otázkami týkajúcimi sa jazyka, obsahu, kontextu a dostupnosti, ktoré by mohli brániť vplyvu na celosvetovej úrovni, najmä na globálnom Juhu.

3.2. Projekt 2 – Činnosti v oblasti vzdelávania a budovania kapacít pre budúcich odborníkov aktívnych v oblasti umelej inteligencie

3.2.1. Účel projektu

Účelom projektu 2 je podporiť začlenenie problematiky odklonu a zneužitia výskumu v oblasti civilnej umelej inteligencie nezodpovednými aktérmi do vzdelávania budúcich generácií odborníkov aktívnych v oblasti umelej inteligencie. Z dlhodobého hľadiska to zabezpečí, aby si študenti vedy, technológie, inžinierstva a matematiky (STEM), ktorí formujú budúcnosť umelej inteligencie, boli vedomí negatívnych vplyvov, ktoré by mohol mať odklon a zneužitie ich práce nezodpovednými aktérmi na medzinárodný mier a bezpečnosť, a aby získali základné nástroje potrebné na identifikáciu a zmierňovanie takýchto rizík v procese výskumu a inovácie.

V rámci tohto projektu sa uskutoční séria vzdelávacích seminárov a seminárov zameraných na budovanie kapacít s pedagógmi a študentmi v spolupráci s vybranými medzinárodnými univerzitami a subjektmi z odvetvia. Cieľom projektu je rozvíjať týmto spôsobom činnosti zamerané na budovanie kapacít, ktoré by pedagógovia a zostavovatelia učebných osnov mohli využiť s cieľom začleniť do odbornej prípravy budúcich odborníkov aktívnych v oblasti umelej inteligencie (napr. do kurzov o etike v oblasti umelej inteligencie a zodpovednej inovácii) otázky týkajúce sa rizík odklonu a zneužitia výskumu a inovácie v oblasti civilnej umelej inteligencie nezodpovednými aktérmi a prepojiť tieto riziká vo všeobecnejšom mierovom a bezpečnostnom kontexte. Prostredníctvom týchto seminárov sa projekt zameria aj na identifikáciu siete zainteresovaných pedagógov, zostavovateľov učebných osnov a študentov, ktorí by podporovali šírenie a propagáciu vzdelávacích materiálov a činností zameraných na budovanie kapacít tohto projektu vo vzdelávacej komunite aktívnej v oblasti umelej inteligencie a v komunite odborníkov aktívnych v oblasti umelej inteligencie. Cieľom tejto zložky týkajúcej sa vytvárania kontaktov je zabezpečiť udržateľnosť projektov aj po ich skončení a umožniť budovanie silnejších väzieb na podporu civilnej technickej angažovanosti v záujme širších cieľov v oblasti mieru, bezpečnosti, odzbrojenia a kontroly zbrojenia.

3.2.2. Opis projektu

V rámci tohto projektu sa uskutoční séria vzdelávacích seminárov a seminárov zameraných na budovanie kapacít s pedagógmi a študentmi z vybraných univerzít z celého sveta. Išlo by o kombináciu prednášok a interaktívnych činností, ktoré pedagógom a študentom poskytnú príležitosť zamyslieť sa nad tým, ako možno výskum a inováciu v oblasti civilnej umelej inteligencie odkloniť a zneužiť tak, že by mohli predstavovať riziká pre medzinárodný mier a bezpečnosť, a ako možno takéto riziká identifikovať, ako im možno predchádzať alebo ako ich možno zmierňovať v rámci procesu výskumu a inovácie alebo prostredníctvom iných procesov riadenia. Tieto činnosti by vychádzali z predchádzajúcej pilotnej práce menšieho rozsahu, ktorú vykonal UNODA a v rámci ktorej sa experimentovalo s metódami na zapojenie študentov STEM a zlepšenie ich informovanosti o tom, že je dôležité, aby zohľadnili širší vplyv svojej práce, ako aj na využívanie odborných znalostí z iných oblastí. Konkrétne by išlo o:

- a) regionálne semináre zamerané na budovanie kapacít pre pedagógov a študentov (4)
 - na regionálnych seminároch sa budú vykonávať a podporovať činnosti, ktoré môžu pedagógovia využiť na budovanie kapacít študentov STEM, pokiaľ ide o zodpovednú inováciu v oblasti umelej inteligencie, s osobitným zameraním na to, ako vyhodnotiť riziká odklonu a zneužitia technológií civilnej umelej inteligencie nezodpovednými aktérmi a zmierňovať ich. Každý seminár sa zorganizuje v spolupráci s univerzitou so sídlom v EÚ a s prominentnou univerzitou z iného regiónu sveta, čím sa vždy zabezpečí prepojenie rôznorodého súboru účastníkov z EÚ s rôznorodým súborom z iného regiónu mimo EÚ. Semináre sa teda budú týkať Latinskej Ameriky a Karibiku, Severnej Ameriky, Afriky a Ázie a Tichomoria. To umožní účasť študentov (na magisterskej a doktorandskej úrovni) z celého sveta vrátane globálneho Juhu. Semináre by sa viedli primárne v angličtine, ale ak by to bolo možné, mohla by sa účastníkom poskytnúť príležitosť zapojiť sa do aktivít založených na alternatívnych jazykových zoskupeniach, a

- b) medzinárodný seminár o udržateľnom budovaní kapacít (1) – seminár bude vychádzať zo skúseností získaných počas regionálnych seminárov a uľahčí výmenu informácií a skúseností medzi pedagógmi a vybranými študentmi z univerzít zapojených do projektu. Na tomto seminári sa bude diskutovať o tom, ako vylepšiť činnosti a nástroje vypracované v priebehu projektu a rozšíriť ich aj mimo skupiny zúčastnených univerzít. Zároveň sa na ňom prediskutujú možnosti, ako podporiť angažovanosť študentov v oblasti zodpovednej umelej inteligencie, ktorá rieši riziká odklonu a zneužitia ohrozujúce medzinárodný mier a bezpečnosť, po tom, ako vstúpia na trh práce.

Na uľahčenie a podporu činností sa podľa potreby využijú siete v rámci SIPRI a UNODA a ich prítomnosť v Afrike, Ázii a Tichomorí, Európe, Severnej a Južnej Amerike.

3.2.3. Očakávané výsledky projektu

Očakáva sa, že počas projektu sa vytvoria modely činností v oblasti budovania kapacít a angažovanosti, ktoré by pedagógovia a zostavovatelia učebných osnov mohli využiť s cieľom zvýšiť povedomie budúcich odborníkov aktívnych v oblasti umelej inteligencie o hrozbách odklonu a zneužitia civilnej umelej inteligencie nezodpovednými aktérmi a o tom, ako môžu pomôcť zmierňovať tieto hrozby prostredníctvom procesov zodpovednej inovácie. Od účastníkov (pedagógov, ale aj študentov STEM) sa očakáva, že po dokončení projektových činností budú schopní využívať a propagovať nástroje, metódy a koncepcie zodpovednej inovácie na identifikáciu a zmierňovanie rizík odklonu a zneužitia pri vývoji a šírení výskumu a inovácie v oblasti civilnej umelej inteligencie.

Očakáva sa tiež, že prostredníctvom projektových činností sa vytvorí sieť pedagógov, zostavovateľov učebných osnov a študentov, ktorí budú projektové činnosti propagovať v rámci vzdelávacích a odborných komunít aktívnych v oblasti umelej inteligencie (napr. počas konferencií Spoločnosti počítačovej inteligencie (*Computational Intelligence Society*) inštitútu IEEE), a zároveň budú pripravení technickými kapacitami prispievať k procesom medzinárodného riadenia pod vedením jednotlivých štátov (napr. proces v rámci Dohovoru o zákazoch alebo obmedzeniach použitia určitých konvenčných zbraní týkajúceho sa vznikajúcich technológií v oblasti smrtiacich autonómnych zbraňových systémov).

Krátkodobá a dlhodobá hodnota týchto činností sa preukáže prostredníctvom prieskumov pred ich začiatkom a po ich skončení.

3.3. Projekt 3 – Uľahčenie dlhodobejšieho udržateľného rozvoja, šírenia a vplyvu zodpovednej inovácie v oblasti umelej inteligencie v záujme mieru a bezpečnosti

3.3.1. Účel projektu

Účelom projektu 3 je uľahčiť dlhodobejší udržateľný rozvoj, šírenie a vplyv zodpovednej inovácie v oblasti umelej inteligencie ako prostriedku na zmiernenie rizík, ktoré môžu predstavovať odklon a zneužitie výskumu a inovácie v oblasti civilnej umelej inteligencie pre mier a bezpečnosť. Na dosiahnutie tohto cieľa sa využijú okrúhle stoly s odvetvím umelej inteligencie, dialógy medzi viacerými zainteresovanými stranami, vypracovanie verejnej správy a ciele činnosti v oblasti šírenia informácií. Cieľom projektu je zabezpečiť, aby vykonaná práca, najmä činnosti v oblasti vzdelávania, budovania kapacít a angažovanosti, mala dosah a vplyv na komunitu umelej inteligencie vo všeobecnosti, a to na všetkých úrovniach (od študentov po inžinierov a iných odborníkov aktívnych v oblasti umelej inteligencie) a v rámci všetkých zemepisných, odvetvových a iných oblastí. Aby sa zvýšila možnosť dosiahnuť široký a hlboký vplyv, je nevyhnutné spolupracovať v tomto priestore s profesijnými organizáciami, ako je IEEE, a zapájať na viacerých úrovniach aktérov z akademickej obce, odvetvia a iných uzavretých skupín. Takéto úsilie poskytne zainteresovaným zástupcom z rôznych komunit v oblasti umelej inteligencie príležitosť prevziať zodpovednosť za túto problematiku a poskytnúť vlastné názory na to, ako udržateľným spôsobom rozvíjať a podporovať úsilie o zmiernenie rizík v rámci celosvetovej komunity umelej inteligencie. Pre štáty, medzivládne organizácie a iných aktérov je dôležité aj z hľadiska dlhodobej hodnoty projektu, aby odborníci aktívni v oblasti umelej inteligencie mohli získať znalosti od vládnych expertov zapojených do zmiernenia rizík v kontexte odzbrojenia, kontroly zbrojenia a nešírenia zbraní a spolupracovať s nimi. Z hľadiska udržateľnosti je takisto rozhodujúce zabezpečiť, aby sa poznatky získané prostredníctvom činností zameraných na angažovanosť primerane analyzovali, konsolidovali a šírili.

3.3.2. Opis projektu

Tento projekt pozostáva z týchto kľúčových zložiek:

- a) dialógy medzi viacerými zainteresovanými stranami o zodpovednej inovácii v oblasti umelej inteligencie v záujme mieru a bezpečnosti (najviac 9) – v rámci tejto série virtuálnych dialógov by sa stretli odborníci z akademickej obce, výskumu, súkromného sektora a tradičnej kontroly zbrojenia z EÚ aj mimo nej, aby diskutovali o
 - i. technologických trendoch, ktoré môžu spôsobiť riziko difúzie, odklonu a zneužitia s vplyvom na medzinárodný mier a bezpečnosť;
 - ii. možnostiach zapojenia sa do zmiernenia rizík prostredníctvom procesov, metód a prostriedkov zodpovednej inovácie a príležitostiach a výzvach týkajúcich sa dialógu a výmen poznatkov medzi komunitami zainteresovaných strán vrátane tých, ktoré pôsobia v iných odvetviach, ako je napríklad biologický a chemický priemysel, a

- iii. potenciálnej hodnote, účele a formáte sebestačnej siete expertov a činností týkajúcich sa dialógu. Expertná skupina sa bude stretávať niekoľkokrát ročne a bude pracovať na organizácii dvoch verejných podujatí pre širšiu komunitu.

Z deviatich virtuálnych stretnutí by dve mali byť otvorené pre verejnosť, aby sa uľahčila širšia konzultácia;

- b) okráhlye stoly súkromného sektora (najviac 6) – touto sériou virtuálnych okrúhlych stolov sa začne dialóg so subjektami, ktoré pracujú s procesmi zodpovednej inovácie v oblasti umelej inteligencie v súkromnom sektore (napr. partnerstvo v oblasti umelej inteligencie), o tom, ako môžu prispieť k minimalizácii rizík odklonu a zneužitia technológií civilnej umelej inteligencie nezodpovednými aktérmi, a preskúmajú sa možné stimuly v rámci rozvoja súkromného sektora v prospech takýchto príspevkov. Témy budú zahŕňať:
 - i. význam medzinárodnej bezpečnosti a odzbrojenia pre súkromný sektor;
 - ii. právne prostredia, v ktorých sa vo svete uskutočňuje vývoj, zavádzanie a prevádzka umelej inteligencie;
 - iii. ako nadviazať na mechanizmy posudzovania rizík a iné opatrenia, ktoré sú súčasťou procesov zodpovednej inovácie a programov dodržiavania predpisov na úrovni podnikov alebo by sa do nich mohli začleniť, a
 - iv. poznatky získané z iných odvetví, procesov a rámcov týkajúcich sa kontroly zbrojenia (napr. biologický a chemický priemysel);
- c) správa o perspektívach komunity umelej inteligencie, pokiaľ ide o kontrolu zbrojenia a zmierňovanie rizík v oblasti umelej inteligencie so zameraním na komunitu umelej inteligencie a komunitu pre kontrolu zbrojenia (1) – vypracovaním tejto správy sa zachytia zistenia a odporúčania projektu a skonsolidujú sa do jedného referenčného dokumentu zameraného na komunitu civilnej umelej inteligencie, ako aj na komunitu pre kontrolu zbrojenia. V správe by sa uviedli spôsoby, ako identifikovať, vyhodnotiť a odstraňovať riziká v oblasti medzinárodného mieru a bezpečnosti spojené s odklonom a zneužitím výskumu a inovácie v oblasti civilnej umelej inteligencie;
- d) podujatia týkajúce sa šírenia informácií so zameraním na konzultácie a spoluprácu s komunitou umelej inteligencie a komunitou pre kontrolu zbrojenia (zatiaľ neurčené) – tím bude počas celého trvania projektu hľadať príležitosti na informovanie o svojej práci a zisteniach a odporúčaniach. Formát podujatí a obsah prezentácií by sa prispôbili potrebám cieľových skupín. Môžu zahŕňať zasadnutia pracovnej skupiny CONOP, zhromaždenia Európskej aliancie pre umelú inteligenciu, skupiny vládnych expertov pre vznikajúce technológie v oblasti smrtiacich autonómnych zbraňových systémov, medziagentúrnej pracovnej skupiny pre umelú inteligenciu (IAWG-AI), iniciatívu Medzinárodnej telekomunikačnej únie (ITU) „AI for Good“ (umelá inteligencia pre dobro), každoročný dialóg inštitútu UNIDIR o inováciách a zasadnutia Inštitútu elektrotechnických a elektronických inžinierov. Tím by sa tiež snažil nadviazať dvojstrannú spoluprácu s príslušnými zainteresovanými stranami na úrovni verejnej správy, akademickej obce a súkromného sektora.

3.3.3. Očakávané výsledky projektu

Očakáva sa, že týmto projektom sa položia základy udržateľného rozvoja, šírenia a vplyvu procesov zodpovednej inovácie v oblasti umelej inteligencie so zameraním na riešenie rizík difúzie, odklonu a zneužitia a ich dôsledkov pre mier a bezpečnosť po uplynutí účinnosti rozhodnutia Rady.

Dialóg medzi viacerými zainteresovanými stranami by mal poskytnúť model výmeny informácií a spolupráce v oblasti zmierňovania rizík nielen v rámci celosvetovej komunity umelej inteligencie, ale aj medzi komunitou zodpovednej civilnej umelej inteligencie a komunitami pre odzbrojenie, kontrolu zbrojenia a nešírenie zbraní. Takýto model by sa mohol použiť na oboznámenie tvorcov politik s kľúčovým technologickým a vedeckým pokrokom relevantným pre zodpovednú inováciu v oblasti umelej inteligencie a tiež na oboznámenie technického publika s prostredím, v ktorom sa v súčasnosti angažujú tvorcovia politik. Očakáva sa, že projekt uľahčí udržateľné vzťahy a spoluprácu medzi zainteresovanými aktérmi v rámci týchto rôznych komunít a medzi nimi. Takéto heterogénne sieťové účinky by mali umožniť väčší rozvoj a širokú podporu zodpovednej inovácie v oblasti umelej inteligencie v záujme mieru a bezpečnosť aj po skončení realizácie projektu.

Očakáva sa, že dialóg so súkromným sektorom umožní väčšiu a hlbšiu angažovanosť súkromného sektora v oblasti umelej inteligencie pri identifikovaní rizík, predchádzaní rizikám a zmierňovaní rizík, pokiaľ ide o mier a bezpečnosť, ktoré vyplývajú z odklonu a zneužitia výskumu a inovácie v oblasti civilnej umelej inteligencie. Projekt by mal umožniť kľúčovým aktérom v procesoch súkromného sektora lepšie pochopiť problémy, ktoré sa snaží riešiť, a prevziať za ne zodpovednosť. Jeho cieľom je okrem toho uľahčiť (všeobecnejšie) prijatie a uplatňovanie procesov, metód a prostriedkov zodpovednej inovácie v existujúcich mechanizmoch a postupoch riadenia rizík na úrovni podnikov.

Očakáva sa tiež, že dialóg medzi viacerými zainteresovanými stranami a okrúhly stôl súkromného sektora poskytnú prehľad o viacerých zásadných otázkach vrátane a) toho, ako možno metódy a prostriedky zodpovednej inovácie ďalej zdokonaľovať a zavádzať s cieľom identifikovať riziká, ktoré predstavuje odklon a zneužitie výskumu a inovácie v oblasti civilnej umelej inteligencie, predchádzať im a zmierňovať ich; b) ako možno pozitívne využiť výskum a inováciu v oblasti umelej inteligencie na podporu cieľov v oblasti mieru a bezpečnosti (napr. aplikácie na včasné varovanie pred konfliktmi a pre humanitárnu pomoc), a c) ako uľahčiť intenzívnejší dialóg a výmenu informácií o úsilí o zmierňovanie rizika, ktoré sa vyvíja v komunite zodpovednej civilnej umelej inteligencie (t. j. rôzne iniciatívy pod vedením IEEE), a úsilím, ktoré už prebieha v rámci komunít pre odzbrojenie, kontrolu zbrojenia a nešírenie zbraní na medzivládnej úrovni.

V správe a činnostiach v oblasti šírenia informácií sa budú analyzovať, konsolidovať a šíriť poznatky získané prostredníctvom projektov 1, 2 a 3, čím sa podporí propagácia zistení a odporúčaní projektových činností v rámci celosvetovej komunity umelej inteligencie, ako aj v rámci politickej komunity. Takisto sa očakáva, že pomôžu zabezpečiť udržateľnosť vplyvu aj po skončení realizácie projektov.

4. TRVANIE

Celkový odhadovaný čas realizácie projektov je 36 mesiacov.

