



Bruxelles, 26 octombrie 2022
(OR. en)

14095/22
ADD 1

LIMITE

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

PROPUNERE

Sursă:	Înaltul Reprezentant al Uniunii pentru afaceri externe și politica de securitate, sub semnătura dlui Stefano SANNINO, Secretar General
Data primirii:	26 octombrie 2022
Destinatar:	Dna Thérèse BLANCHET, Secretară Generală a Consiliului Uniunii Europene
Subiect:	Propunerea Înaltului Reprezentant al Uniunii pentru afaceri externe și politica de securitate adresată Consiliului cu privire la o decizie a Consiliului de sprijinire a punerii în aplicare a proiectului „Promovarea inovării responsabile în inteligența artificială pentru pace și securitate”

În anexă, se pune la dispoziția delegațiilor documentul HR(2022) 238.

Anexă: HR(2022) 238

SERVICIUL EUROPEAN DE ACȚIUNE EXTERNĂ



DOCUMENT DE PROIECT „Promovarea inovării responsabile în IA pentru pace și securitate”

HR(2022)238

1. CONTEXT

Progresele recente înregistrate în domeniul inteligenței artificiale (IA) au deschis noi posibilități de a sprijini și de a susține pacea și securitatea, de exemplu, prin realizarea de îmbunătățiri tehnologice în domenii precum alerta timpurie în situații de conflict și verificarea controlului armamentelor și al exporturilor. Pe de altă parte, aceste progrese au permis apariția unor noi modalități de generare – sau de agravare – a tensiunilor, conflictelor și insecurității dintre state și în interiorul acestora. Riscurile pe care le prezintă anumite aplicații ale IA, cum ar fi sistemele de arme letale autonome, au devenit un motiv major de îngrijorare pentru comunitatea care susține controlul armamentelor. O sursă de risc ce merită o atenție sporită și în privința căreia eforturile diplomatice și cele referitoare la controlul armamentelor din prezent ar putea să nu ofere un răspuns suficient este deturnarea și utilizarea abuzivă a cercetării și inovării civile în domeniul IA de către agenți iresponsabili, inclusiv de către agenți nestatali rău-intenționați, precum utilizarea abuzivă a rețelelor adversariale generative (GAN) pentru a produce materiale *deepfake* în vederea unor campanii de dezinformare.

IA este o tehnologie generică cu un potențial ridicat de utilizare generală. Activitățile de cercetare și inovare în domeniul IA realizate pentru dezvoltarea unor aplicații civile ar putea fi (relativ ușor) accesate și reorientate de anumiți agenți în vederea unor utilizări dăunătoare sau perturbatoare care ar putea avea implicații pentru pacea și securitatea internațională. Deturnarea și utilizarea abuzivă a tehnologiilor civile nu sunt fenomene noi și nici nu sunt specifice IA. În domeniul conex al roboticii, acest fenomen a fost exemplificat recent de utilizarea ca arme a dronelor de agrement de către Daesh/ISIS în Siria. În cazul IA însă, problema este complicată de mai mulți factori: natura imaterială și în rapidă evoluție a algoritmilor și a datelor IA, ceea ce face ca transferul/proliferarea acestora să fie dificil de controlat; rolul de prim-plan jucat de sectorul privat în ecosistemul de cercetare, dezvoltare și inovare și nevoia acestuia, în consecință, de a proteja algoritmiile care fac obiectul unor drepturi exclusive, precum și existența la nivel global a expertizei umane și a resurselor materiale capabile să reorienteze utilizarea tehnologiilor IA. De asemenea, persoanele care lucrează în domeniul IA în sectorul civil nu sunt adesea conștiente de posibilele implicații pe care le-ar putea avea deturnarea și utilizarea abuzivă a activității lor pentru pacea și securitatea internațională sau ezită să participe la discuțiile existente privind riscurile legate de IA în cercurile care susțin controlul armamentelor și neproliferarea.

Este necesar să fie susținută o mai mare implicare a comunității civile din domeniul IA în înțelegerea și atenuarea riscurilor la adresa păcii și securității care sunt legate de deturnarea și utilizarea abuzivă a tehnologiei IA civile de către agenți iresponsabili. Din punctul de vedere al Institutului internațional pentru cercetare în domeniul păcii de la Stockholm (SIPRI) și al Biroului pentru Probleme de Dezarmare al Organizației Națiunilor Unite (UNODA), acest lucru ar putea fi realizat prin promovarea într-o mai mare măsură a inovării responsabile în cadrul comunității civile mondiale din domeniul IA. Lucrările desfășurate anterior de SIPRI și UNODA au demonstrat că inovarea responsabilă, ca mecanism de autoguvernare, i-ar putea oferi comunității civile mondiale din domeniul IA instrumente și metode practice pentru a identifica riscurile la adresa păcii și securității pe care le-ar putea reprezenta deturnarea și utilizarea abuzivă a cercetării și inovării civile în domeniul IA, precum și pentru a contribui la prevenirea și atenuarea acestor riscuri. Lucrările desfășurate de SIPRI și de UNODA au identificat, de asemenea, o serie de metodologii și mai multe inițiative în curs axate pe sectorul civil referitoare la IA responsabilă, care ar putea fi utilizate pentru a sensibiliza comunitatea civilă din domeniul IA cu privire la aspectele legate de controlul armamentelor și de neproliferare, la dezbaterile experților și la pozițiile statelor privind dezvoltarea, difuzarea și utilizarea responsabile ale IA, precum și la învățămintele care pot fi desprinse din activitățile privind responsabilitatea sectorului apărării¹. Foarte important, aceste lucrări desfășurate anterior au identificat în mod clar faptul că implicarea studenților din domeniul științelor, tehnologiei, ingineriei și matematicii (STIM), care încă se familiarizează cu IA într-un cadru educațional, reprezintă un element central al oricăror eforturi eficace în sensul inovării responsabile.

2. OBIECTIVE

Aceste proiecte urmăresc să susțină o implicare mai mare a comunității civile din domeniul IA în ceea ce privește atenuarea riscurilor pe care le pot reprezenta pentru pacea și securitatea internațională deturnarea și utilizarea abuzivă a cercetării și inovării civile în domeniul IA de către agenți iresponsabili. Ele urmăresc să realizeze acest lucru, în primul rând, prin obținerea unei mai bune înțelegeri a modului în care deciziile privind dezvoltarea și difuzarea cercetării și inovării în domeniul IA pot avea impact asupra riscurilor de deturnare și de utilizare abuzivă și, prin urmare, pot genera riscuri sau oportunități pentru pace și securitate și, în al doilea rând, prin promovarea unor procese, metode și instrumente de inovare responsabile, care să contribuie la asigurarea aplicării pașnice a inovațiilor civile și la diseminarea responsabilă a cunoștințelor din domeniul IA. În acest scop, proiectele vor sprijini activitățile de consolidare a capacităților, de cercetare și de participare care i) vor consolida capacitatea comunității civile mondiale din domeniul IA de a include și de a aborda riscurile în materie de pace și securitate pe care le prezintă deturnarea și utilizarea abuzivă a IA civile de către agenți iresponsabili, prin procese de inovare responsabile și ii) vor consolida legătura dintre eforturile de atenuare a riscurilor depuse în sfera civilă în ceea ce privește IA responsabilă și eforturile aflate deja în desfășurare la nivel interguvernamental în cadrul comunității care susține dezarmarea, controlul armamentelor și neproliferarea. Un aspect esențial este acela că proiectele nu sunt menite să stabilească noi

¹ Metodologiile includ, de exemplu, practicile recomandate ale Institutului Inginerilor Electrotehniști și Electroniști (IEEE) pentru evaluarea impactului sistemelor autonome și inteligente asupra bunăstării umane (IEEE Std 7010-2020), precum și Lista de evaluare pentru o inteligență artificială de încredere (ALTAI) a Grupului de experți la nivel înalt privind inteligența artificială. Printre aceste inițiative se numără: Inițiativa globală privind etica sistemelor autonome și inteligente a IEEE, Parteneriatul privind IA și Parteneriatul mondial privind IA.

standarde, principii sau reglementări și nici să intervină în domenii care țin de competența statelor. În schimb, ele urmăresc să dezvolte activități de inovare responsabilă pe plan civil, să ia în calcul riscurile pentru pace și securitate pe care le prezintă deturnarea și utilizarea abuzivă a IA civile de către agenți iresponsabili și să ofere formare cu privire la eforturile interguvernamentale actuale care sunt relevante în acest sens.

Pentru a ajunge în mod efectiv la comunitatea civilă din domeniul IA și a avea un impact asupra acesteia, proiectele pun în aplicare o abordare tridirecțională, urmărind:

- (a) să colaboreze cu formatorii – să lucreze împreună cu anumiți formatori și cu anumite persoane care elaborează programe academice în legătură cu elaborarea și promovarea de materiale educaționale care pot fi utilizate pentru a integra în programele de formare a viitorilor profesioniști din domeniul IA (de exemplu, în cursuri privind etica IA și inovarea responsabilă) considerațiile referitoare la riscurile la adresa păcii și securității care decurg din deturnarea și utilizarea abuzivă a cercetării și inovării civile în domeniul IA de către agenți iresponsabili; să colaboreze cu studenții – să asigure familiarizarea anumitor studenți din domeniul științelor, tehnologiei, ingineriei și matematicii (STIM) selectați din întreaga lume cu modul în care riscurile la adresa păcii și securității pe care le prezintă deturnarea și utilizarea abuzivă de către agenți iresponsabili a dezvoltării civile a IA pot fi identificate, prevenite sau atenuate în cadrul procesului de cercetare și inovare sau prin alte procese de guvernare
- (b) și
- (c) să colaboreze cu industria IA – să colaboreze cu asociații profesionale și cu organisme de standardizare, precum Institutul Inginerilor Electrotehniști și Electroniști (IEEE), în vederea i) diseminării de materiale pedagogice adaptate și de activități de mobilizare în rândul profesioniștilor din domeniul tehnic; ii) sprijinirii utilizărilor pozitive ale IA pentru pace și securitate și iii) facilitării dialogului și a schimbului de informații între experții din mediul academic, sectorul privat și cel guvernamental cu privire la modul în care poate fi atenuat riscul de deturnare și de utilizare abuzivă a cercetării și inovării civile în domeniul IA de către agenți iresponsabili.

O astfel de abordare permite ca proiectele să ajungă la comunitatea din domeniul IA la toate nivelurile, incluzând nu numai profesioniștii actuali, ci și generațiile viitoare. Aceasta permite, de asemenea, stabilirea de relații între sectorul universitar, cel industrial și sectoare de altă natură și sprijină sustenabilitatea eforturilor viitoare prin crearea de rețele care trec dincolo de aceste limite.

Proiectele urmăresc, de asemenea, să utilizeze experiența și capacitatea SIPRI și UNODA de a strânge laolaltă actorii relevanți pentru a avea un impact asupra comunității din domeniul IA la nivel mondial, nu doar asupra părților interesate din UE. SIPRI și UNODA se află într-o poziție unică pentru a ajunge la actorii din domeniul IA din Africa, Asia-Pacific, Europa, America de Nord și America de Sud și pentru a facilita cooperarea dintre aceștia. Ambele entități au, de asemenea, experiență în alte domenii ale științei și tehnologiei care se confruntă cu provocări similare, în materie de dublă utilizare și proliferare, inclusiv în domeniul biotehnologiei. De asemenea, proiectele urmăresc să profite de condițiile existente în Uniunea Europeană, precum a) existența unor procese avansate care implică mai multe părți interesate în domeniul IA responsabile; b) nivelul ridicat de implicare și de expertiză din UE în ceea ce privește chestiunile legate de dezarmare, controlul armamentelor și neproliferare; c) diversitatea legăturilor pe care

organizațiile universitare, de cercetare și din sectorul privat din UE le au cu alte regiuni, în special cu cele din sudul global, care vor constitui, de asemenea, un obiectiv major al eforturilor de implicare și d) diversitatea naționalităților studenților, formatorilor și inginerilor din universități, instituții de cercetare și sectorul privat.

Incluziunea va reprezenta o considerație esențială pentru desfășurarea activităților din cadrul proiectelor. Pentru a sprijini în mod eficace comunitatea din domeniul IA, proiectele țin seama de faptul că comunitatea din domeniul IA este formată dintr-o gamă diversă de actori și, în special, că

- (a) genul este un factor extrem de relevant. Din acest motiv, egalitatea de gen va fi integrată în conformitate cu strategiile existente la nivelul sistemului ONU privind integrarea perspectivei de gen și a parității. Participarea femeilor la toate activitățile din cadrul proiectului va fi încurajată și necesară; și
- (b) incluziunea persoanelor cu handicap, precum și luarea în considerare în mod rezonabil a nevoilor acestora vor fi asigurate pe întreaga durată a procesului. Acest lucru va include eliminarea obstacolelor din calea participării persoanelor cu handicap, precum și asigurarea faptului că sunt luate măsuri pentru implicarea acestor persoane și pentru facilitarea reprezentării opiniilor și experiențelor substanțiale ale acestora.

3. PROIECTE

Cele trei proiecte descrise mai jos sunt concepute să fie complementare și să se sprijine reciproc, elementele urmând să se deruleze pe parcursul celor 36 de luni.

Proiectul 1 – Producerea de materiale pedagogice și de materiale referitoare la consolidarea capacităților pentru comunitatea civilă din domeniul IA

3.1.Scopul proiectului

Proiectul 1 urmărește în principal să le furnizeze actorilor civili din domeniul IA cunoștințele și mijloacele pentru a evalua și atenua riscurile pe care le pot reprezenta pentru pacea și securitatea internațională deturnarea și utilizarea abuzivă a cercetării și inovării civile în domeniul IA de către agenți iresponsabili. Scopul său este de a produce materiale pedagogice și materiale referitoare la consolidarea capacităților care le vor oferi profesioniștilor din domeniul IA din toate regiunile, de toate nivelurile și din toate sectoarele (inclusiv formatorii axați pe IA, persoanele care elaborează programe de învățământ, studenții STIM și inginerii și cercetătorii din domeniul IA din mediul academic și din sectorul privat) informațiile și instrumentele necesare pentru:

- (a) a înțelege cum cercetarea și inovarea civilă în domeniul IA ar putea fi deturnate și utilizate în mod abuziv în moduri care ar putea prezenta riscuri pentru pacea și securitatea internațională și cum deciziile luate în cadrul procesului de dezvoltare și difuzare a cercetării și a inovării pot crește sau reduce riscul de deturnare și de utilizare abuzivă;

- (b) a înțelege eforturile deja depuse de comunitatea din domeniul dezarmării, al controlului armamentelor și al neproliferării pentru a reduce riscurile de deturnare și utilizare abuzivă a cercetării și inovării civile și
- (c) a practica o inovare responsabilă într-un mod care să atenueze riscul de deturnare și de utilizare abuzivă în cadrul procesului de dezvoltare și difuzare a cercetării și a inovării.

3.1.1. Descrierea proiectului

Acest proiect va produce trei seturi distincte de materiale pedagogice și de materiale referitoare la consolidarea capacităților.

- (a) Manual (1) – Manualul va sintetiza cunoștințele de bază și modalitățile prin care actorii din domeniul IA vor putea să evalueze și să reducă, în procesul de cercetare și inovare, riscurile de deturnare și utilizare abuzivă a tehnologiei IA civile de către agenți iresponsabili. În manual vor fi examinate cauzele și modul în care deciziile referitoare la dezvoltarea și difuzarea cercetării și inovării pot avea un impact asupra riscurilor de deturnare și de utilizare abuzivă și, prin urmare, pot genera riscuri sau oportunități pentru pace și securitate. Manualul va conține, de asemenea, o introducere referitoare la obligațiile relevante în materie de drept internațional și de control al exporturilor, pe lângă considerațiile privind siguranța și securitatea aflate în discuție în cercurile militare, precum și în cele care susțin dezarmarea, controlul armamentelor și neproliferarea și va prezenta exemple de procese și instrumente care permit practicarea unei inovări responsabile, precum metodologiile de evaluare a impactului tehnologiilor și modelele de evaluare a riscurilor.
- (b) O serie de podcasturi (~ 10) – Aceste podcasturi vor reprezenta un mediu accesibil și atractiv pentru a le permite actorilor din domeniul IA să afle de ce și în ce mod procesele de inovare responsabilă în domeniul IA pot sprijini pacea și securitatea internațională prin atenuarea riscurilor reprezentate de deturnarea și utilizarea abuzivă de către agenți iresponsabili. Seria de podcasturi va trece în revistă teme importante (de exemplu, modelul deturnării și al utilizării abuzive a cercetării și inovării cu dublă utilizare/cu utilizare generală; provocările umanitare, strategice și politice asociate cu posibila utilizare abuzivă a cercetării și inovării civile în domeniul IA; provocările cu care se confruntă cercurile care susțin dezarmarea, controlul armamentelor și neproliferarea în cadrul eforturilor de atenuare a riscurilor; modul în care trebuie realizată inovarea responsabilă prin intermediul evaluării riscurilor; conformitatea controlului exporturilor; reducerea riscurilor încă din stadiul conceperii; editarea responsabilă; cunoașterea clienților și experiența rezultată în urma exercițiilor de simulare teoretică) și va fi structurată în jurul interviurilor pe care echipa de proiect le va desfășura cu reprezentanți ai comunităților relevante.
- (c) O serie de postări pe blog (9-10) – Echipa va realiza un blog care va conține o serie de postări selecționate, menite să sporească vizibilitatea eforturilor care încearcă să depășească granițele dintre comunitatea din domeniul „IA responsabile” axate pe utilizări civile și comunitatea care susține controlul armamentelor și neproliferarea. Seria de postări pe blog va oferi o platformă pentru diseminarea de informații, idei și soluții

privind identificarea și gestionarea riscurilor asociate deturnării și utilizării abuzive a IA civile în procesul de cercetare și inovare. Blogul își va propune să reprezinte diversitatea ideilor și perspectivelor prezente în sectorul IA.

Aceste materiale vor fi difuzate public prin intermediul site-urilor web ale actorilor de implementare, al prezenței lor pe platformele de comunicare socială și prin comunicare directă cu entitățile relevante din mediul universitar, cu asociațiile profesionale din domeniul IA civile și cu alte grupuri corespunzătoare.

3.1.2. Rezultatele preconizate ale proiectului

Acest proiect ar trebui să producă un nou set de materiale prin care profesioniștii din domeniul IA civile să poată fi sensibilizați cu privire la: a) felul în care cercetarea și inovarea civilă în domeniul IA ar putea fi deturnate și utilizate în mod abuziv în moduri care ar putea prezenta riscuri pentru pacea și securitatea internațională, b) felul în care sunt gestionate aceste riscuri de comunitatea care susține dezarmarea, controlul armamentelor și neproliferarea și c) felul în care profesioniștii din domeniul IA ar putea contribui într-o mai mare măsură la atenuarea acestor riscuri prin intermediul unor procese de inovare responsabilă.

Acest lucru ar trebui să permită intensificarea implicării sectorului IA civile în ceea ce privește atenuarea riscurilor pe care le pot reprezenta pentru pacea și securitatea internațională deturnarea și utilizarea abuzivă a IA civile, să sporească capacitatea profesioniștilor care desfășoară activități tehnice de a se implica în procesele relevante desfășurate în cadrul comunității care susține dezarmarea, controlul armamentelor și neproliferarea, precum și să sprijine implicarea unor noi categorii de public care nu sunt incluse în mod tradițional în eforturile educaționale privind dezarmarea și neproliferarea.

Aceste materiale ar trebui, de asemenea, să sprijine punerea în aplicare a celorlalte proiecte și vor servi drept bază pentru activitățile educaționale și de consolidare a capacităților din cadrul proiectului 2, precum și pentru activitățile de dialog și de mobilizare desfășurate în cadrul proiectului 3. Aceste activități ar trebui, la rândul lor, să poată fi folosite pentru producerea și ameliorarea materialelor. O astfel de abordare iterativă ar trebui să contribuie la eliminarea eventualelor obstacole din calea promovării, difuzării și utilizării lor în cadrul comunității din domeniul IA, inclusiv în ceea ce privește aspectele legate de limbă, conținut, context și disponibilitate, care ar putea afecta impactul acestora la nivel mondial, în special în sudul global.

3.2. Proiectul 2 – Activități de formare și de consolidare a capacităților pentru viitorii profesioniști din domeniul IA

3.2.1. Scopul proiectului

Scopul proiectului 2 este de a sprijini integrarea problemei deturnării și utilizării abuzive a cercetării civile în domeniul IA de către agenți iresponsabili în formarea generațiilor viitoare de profesioniști în domeniul IA. Pe termen lung, acest lucru va asigura faptul că studenții din domeniile științelor, tehnologiei, ingineriei și matematicii (STIM) care modelează viitorul IA sunt conștienți de impactul negativ pe care deturnarea și utilizarea abuzivă a activității lor de

către agenți iresponsabili l-ar putea avea asupra păcii și securității internaționale și că vor dispune de instrumentele de bază necesare pentru a identifica și a atenua un astfel de risc în procesul de cercetare și inovare.

Acest proiect va organiza o serie de ateliere educaționale și de consolidare a capacităților cu formatori și studenți, în colaborare cu universități internaționale selectate și actori din industrie. Proiectul urmărește prin aceasta să dezvolte activități de consolidare a capacităților pe care formatorii și persoanele care elaborează programe academice le-ar putea folosi pentru a include în formarea viitorilor profesioniști din domeniul IA (de exemplu, cursuri privind etica IA și inovarea responsabilă) considerații privind riscurile de deturnare și utilizare abuzivă a cercetării și inovării civile în domeniul IA de către agenți iresponsabili și conectarea acestor activități la contextul mai larg al păcii și securității. Prin intermediul acestor ateliere, proiectul va urmări, de asemenea, să identifice o rețea de formatori, persoane care elaborează programe de învățământ și studenți interesați care ar sprijini diseminarea și promovarea materialelor educaționale ale proiectului și a activităților de consolidare a capacităților în comunitatea educațională din domeniul IA și în comunitatea profesioniștilor din domeniul IA. Această componentă de colaborare în rețea urmărește să asigure sustenabilitatea proiectelor dincolo de durata lor imediată și să permită construirea unor legături mai puternice în sprijinul implicării sectorului tehnic civil în vederea atingerii unor obiective mai ample în materie de pace, securitate, dezarmare și control al armamentelor.

3.2.2. Descrierea proiectului

Acest proiect va organiza o serie de ateliere educaționale și de consolidare a capacităților cu formatori și studenți din universități selectate din întreaga lume. Aceste ateliere ar cuprinde o combinație de conferințe și activități interactive care vor oferi formatorilor și studenților oportunități de a reflecta asupra modului în care cercetarea și inovarea civilă în domeniul IA ar putea fi deturnate și utilizate în mod abuziv în direcții care pot genera riscuri pentru pacea și securitatea internațională și asupra modului în care aceste riscuri pot fi identificate, prevenite sau atenuate în cadrul procesului de cercetare și inovare sau prin alte procese de guvernare. Aceste activități se vor baza pe activitățile pilot anterioare la scară mai mică desfășurate de UNODA, care a experimentat metode de implicare și sensibilizare a studenților din domeniile STIM cu privire la importanța luării în considerare a impactului mai larg al activității lor, precum și a colaborării cu expertiza din afara domeniilor lor de origine. Concret, aceste activități ar urma să constea în:

- (a) ateliere regionale de consolidare a capacităților pentru formatori și studenți (4) – atelierele regionale vor desfășura și vor promova activități pe care formatorii le pot utiliza pentru a consolida capacitatea studenților STIM în inovarea responsabilă în domeniul IA, cu un accent deosebit pe modul de evaluare și atenuare a riscurilor de deturnare și utilizare abuzivă a tehnologiei IA civile de către agenți iresponsabili. Fiecare atelier va fi organizat cu o universitate din UE și cu o universitate de înalt nivel dintr-o altă regiune a lumii, conectând astfel întotdeauna un set divers de participanți din UE cu un set divers din exteriorul UE. Atelierele vor acoperi apoi America Latină și zona Caraibilor, America de Nord, Africa și Asia și zona Pacificului. Acest lucru va permite participarea

studentilor (la nivel de masterat și doctorat) din întreaga lume, inclusiv din sudul global. Atelierele ar urma să se desfășoare în principal în limba engleză, dar, acolo unde este posibil, participanților li se va oferi posibilitatea de a se implica în activități bazate pe grupuri lingvistice alternative; și

- (b) un atelier internațional privind consolidarea capacităților durabile (1) – atelierul se va baza pe lecțiile învățate din atelierele regionale și va facilita schimbul de informații și de experiență între formatori și studenții selectați din universitățile implicate în proiect. În cadrul atelierului va fi discutat modul în care se pot perfecționa activitățile și instrumentele elaborate pe parcursul proiectului și cum pot fi acestea diseminate dincolo de grupul de universități participante. În cadrul acestuia se va discuta, de asemenea, despre modalitățile de sprijinire a implicării studenților în IA responsabilă, care abordează riscurile de deturnare și de utilizare abuzivă pentru pacea și securitatea internațională, odată ce aceștia au intrat pe piața forței de muncă.

Rețelele și prezența SIPRI și UNODA în Africa, Asia-Pacific, Europa, America de Nord și de Sud vor fi utilizate pentru a facilita și a sprijini aspecte ale activităților, după caz.

3.2.3. Rezultatele preconizate ale proiectului

Se preconizează că proiectul va crea modele de consolidare a capacităților și de activități de implicare pe care formatorii și persoanele care elaborează programe academice le-ar putea reproduce pentru a sensibiliza viitorii profesioniști din domeniul IA cu privire la problemele legate de deturnarea și utilizarea abuzivă a IA civile de către agenți iresponsabili și la modul în care acești profesioniști pot contribui la atenuarea acestor probleme prin procese de inovare responsabile. După finalizarea activităților proiectului, se așteaptă ca participanții (formatori, dar și studenți STIM) să poată utiliza și promova instrumente, metode și concepte de inovare responsabile pentru a identifica și a atenua riscurile de deturnare și utilizare abuzivă în dezvoltarea și difuzarea cercetării și inovării civile în domeniul IA.

Se preconizează, de asemenea, că activitățile proiectului vor genera o rețea de formatori, persoane care elaborează programe de învățământ și studenți care nu numai că vor promova activitățile proiectului în cadrul formării și al comunităților profesionale în domeniul IA (de exemplu, în timpul conferințelor Computational Intelligence Society a IEEE), ci vor fi, de asemenea, disponibili pentru a contribui cu capacitate tehnică la procesele de guvernare internațională conduse de state (de exemplu, procesul Convenției privind anumite tipuri de arme convenționale referitor la tehnologiile emergente în domeniul sistemelor de arme letale autonome).

Valoarea pe termen scurt și lung a acestor activități va fi demonstrată prin anchete prealabile și ulterioare activității.

3.3. Proiectul 3 – Facilitarea dezvoltării, a diseminării și a impactului durabile pe termen lung ale inovării responsabile în domeniul IA pentru pace și securitate

3.3.1. Scopul proiectului

Scopul proiectului 3 este de a facilita dezvoltarea, diseminarea și impactul durabile pe termen lung ale inovării responsabile în domeniul IA ca mijloc de atenuare a riscurilor pe care deturnarea și utilizarea abuzivă a cercetării și inovării civile în domeniul IA le pot prezenta pentru pace și securitate. Scopul său este de a realiza acest lucru prin mese rotunde cu industria IA, prin dialoguri multipartite, prin crearea unui raport public și prin activități de diseminare specifice. Proiectul urmărește să se asigure că activitatea generată, în special activitățile de formare, de consolidare a capacităților și de implicare, ajunge la comunitatea IA în general și are un impact asupra acesteia, la toate nivelurile (de la studenți la ingineri și alți profesioniști din domeniul IA) și dincolo de frontierele geografice, sectoriale și de altă natură. Pentru a spori posibilitatea de a avea un impact amplu și profund, este esențial să se coopereze cu organizațiile profesionale din acest spațiu, cum ar fi IEEE, și să se desfășoare angajamente multidimensionale la nivelul mediului academic, al industriei și al altor compartimente. Astfel de eforturi vor oferi reprezentanților interesați din diferite comunități IA posibilitatea de a-și asuma responsabilitatea față de problemă și de a-și exprima propriile opinii cu privire la modul în care eforturile de atenuare a riscurilor pot fi realizate și promovate în mod durabil în interiorul comunității globale IA și pe tot cuprinsul acesteia. De asemenea, pentru valoarea pe termen lung a proiectului pentru state, organizații interguvernamentale și alte entități, este important ca profesioniștii din domeniul IA să poată învăța de la experți guvernamentali implicați în atenuarea riscurilor în contextul dezarmării, al controlului armamentelor și al neproliferării și să colaboreze cu aceștia. De asemenea, este esențial pentru durabilitate să se asigure că cunoștințele generate prin activitățile de colaborare sunt analizate, consolidate și diseminate în mod corespunzător.

3.3.2. Descrierea proiectului

Acest proiect constă în următoarele componente principale:

- (a) dialoguri multipartite privind „inovarea responsabilă în domeniul IA pentru pace și securitate” (până la 9) – această serie de reuniuni de dialog în format virtual ar reuni experți din mediul academic, cercetare, sectorul privat și cel al controlului armelor tradiționale din UE și din afara acesteia pentru a discuta despre:
- i. tendințele tehnologice care pot genera riscuri de difuzare, deturnare și utilizare abuzivă, cu impact asupra păcii și securității internaționale;
 - ii. modul de implicare în atenuarea riscurilor prin procese, metode și mijloace de inovare responsabile, precum și oportunități și provocări pentru dialog și schimb de cunoștințe între comunitățile de părți interesate, inclusiv cele care își desfășoară activitatea în alte sectoare, cum ar fi industria biologică și chimică; și
 - iii. valoarea potențială, scopul și formatul unei rețele autonome de experți și de activități de dialog. Grupul de experți se va reuni de mai multe ori pe an și va acționa în vederea organizării a două evenimente publice pentru întreaga comunitate.

Dintre cele nouă reuniuni virtuale, două sunt menite să fie deschise publicului, pentru a facilita o consultare mai amplă.

(b) mese rotunde cu sectorul privat (până la 6) – această serie de mese rotunde virtuale va iniția un dialog cu actorii care lucrează cu procese responsabile de inovare în domeniul IA în sectorul privat (de exemplu, Parteneriatul privind IA) cu privire la modul în care acestea pot contribui la reducerea la minimum a riscurilor de deturnare și utilizare abuzivă a tehnologiilor IA civile de către agenți iresponsabili și va explora posibile stimulente în cadrul dezvoltării sectorului privat în acest sens. Subiectele vor include:

- i. relevanța contextului internațional în materie de securitate și dezarmare pentru sectorul privat;
- ii. cadrul (cadrele) juridic(e) în care se înscriu dezvoltarea, implementarea și exploatarea IA în întreaga lume;
- iii. modul de valorificare a mecanismelor de evaluare a riscurilor și a altor măsuri care fac parte din sau ar putea fi integrate în procesele de inovare responsabile și în programele de asigurare a conformității la nivelul întreprinderii; și
- iv. învățămintele care trebuie desprinse din alte sectoare industriale, procese și cadre legate de controlul armamentelor (de exemplu, industria biologică și industria chimică).

(c) un raport privind perspectivele comunității IA în ceea ce privește controlul armamentelor și atenuarea riscurilor în domeniul IA, care vizează comunitatea IA și cea care susține controlul armamentelor (1) – elaborarea acestui raport va include și va consolida constatările și recomandările proiectului într-un singur document de referință care vizează atât comunitatea de experți civili axată pe IA responsabilă, cât și comunitatea care susține controlul armamentelor. Raportul ar urma să examineze modul în care pot fi identificate, evaluate și abordate riscurile pentru pacea și securitatea internațională asociate cu deturnarea și utilizarea abuzivă a cercetării și inovării civile în domeniul IA.

(d) evenimente de diseminare care vizează consultarea și colaborarea cu comunitățile de IA și de control al armamentelor (de stabilitate) – echipa va căuta oportunități de a comunica activitatea și constatările și recomandările sale pe tot parcursul proiectului. Formatul evenimentelor și conținutul prezentărilor vor fi adaptate la nevoile grupurilor-țintă. Acestea pot include reuniuni ale CONOP, ale Adunării Alianței europene în domeniul inteligenței artificiale; ale grupului de experți guvernamentali privind tehnologiile emergente în domeniul sistemelor de arme letale autonome; ale grupului de lucru interagenției privind IA (IAWG-AI); ale inițiativei UIT privind „IA în slujba binelui”; ale Dialogului anual privind inovarea al UNIDIR; și ale Institutului Inginerilor Electrotehniști și Electroniști. Echipa va încerca, de asemenea, să colaboreze bilateral cu părțile interesate relevante din cadrul guvernelor, al mediului academic și al sectorului privat.

3.3.3. Rezultatele preconizate ale proiectului

Se preconizează că acest proiect va pune bazele dezvoltării, diseminării și impactului durabile ale inovării responsabile a proceselor de IA care abordează riscurile de răspândire, deturnare și utilizare abuzivă și implicațiile acestora pentru pace și securitate dincolo de durată imediată a deciziei Consiliului.

Se preconizează că dialogul multipartit va oferi un model de schimb de informații și de colaborare privind atenuarea riscurilor nu numai în cadrul comunității globale a IA, ci și între comunitatea de experți civili axată pe IA responsabilă și comunitățile de dezarmare, control al armamentelor și neproliferare. Un astfel de model ar putea fi utilizat pentru a familiariza factorii de decizie cu principalele progrese tehnologice și științifice relevante pentru inovarea responsabilă a IA și, de asemenea, pentru a familiariza publicul tehnic cu mediul în care factorii de decizie sunt implicați în prezent. Se prevede că proiectul va facilita relațiile durabile și implicarea între actorii interesați în cadrul acestor diferite comunități și între acestea. Se preconizează că astfel de efecte de rețea eterogene vor permite dezvoltarea într-o mai mare măsură și promovarea pe scară largă a inovării responsabile a IA pentru pace și securitate după momentul încheierii proiectului.

Se prevede că dialogul cu sectorul privat va permite o implicare mai mare și mai profundă a sectorului privat al IA în identificarea, prevenirea și atenuarea riscurilor în materie de pace și securitate care decurg din deturnarea și utilizarea abuzivă a cercetării și inovării civile în domeniul IA. Se preconizează că proiectul va oferi actorilor-cheie din procesele din sectorul privat o mai bună înțelegere și asumare a problemelor pe care urmărește să le abordeze. În plus, scopul proiectului este de a facilita adoptarea și punerea în aplicare (pe scară mai largă) a proceselor, metodelor și mijloacelor de inovare responsabile în cadrul mecanismelor și procedurilor existente de gestionare a riscurilor corporative.

Se prevede, de asemenea, că dialogul multipartit și masa rotundă cu sectorul privat vor genera informații cu privire la o serie de aspecte importante, inclusiv a) modul în care metodele și mijloacele de inovare responsabile pot fi îmbunătățite și utilizate într-o mai mare măsură pentru a identifica, a preveni și a atenua riscurile prezentate de deturnarea și utilizarea abuzivă a cercetării și inovării civile în domeniul IA; b) modul în care cercetarea și inovarea în domeniul IA pot fi utilizate în mod pozitiv pentru a sprijini obiectivele în materie de pace și securitate (de exemplu, aplicațiile de alertă timpurie în caz de conflict și pentru asistență umanitară); și c) modalitățile de facilitare a unui nivel mai susținut de dialog și schimb de informații între eforturile de atenuare a riscurilor întreprinse în cadrul comunității de experți civili axate pe IA responsabilă (și anume, diverse inițiative conduse de IEEE) și cele aflate deja în desfășurare în cadrul comunității de dezarmare, control al armamentelor și neproliferare la nivel interguvernamental.

Raportul și activitățile de diseminare vor analiza, consolida și disemina informațiile generate prin proiectele 1, 2 și 3, sprijinind astfel promovarea constatărilor și recomandărilor activităților proiectelor în cadrul comunității mondiale de IA, precum și în cadrul comunității decidenților. De asemenea, se preconizează că acestea vor contribui la asigurarea sustenabilității impactului după momentul încheierii proiectelor.

4. DURATA

Durata totală estimată a punerii în aplicare a proiectelor este de 36 de luni.

(d)