

## **Tecnologie del linguaggio umano agili e potenti per la difesa - Partecipazione a una sfida tecnologica**

### **Agile and robust human language technologies for defence - Participation to a technological challenge**

**TOPIC ID:** EDF-2023-LS-RA-CHALLENGE-DIGIT-HLTP

**Ente finanziatore:** Commissione europea, Fondo europeo per la Difesa

**Obiettivi ed impatto attesi:** Con la digitalizzazione del campo di battaglia, che porta a interfacce utente sempre più complesse e a volumi sempre maggiori di dati linguistici da elaborare, le tecnologie linguistiche come l'interazione multilingue scritta o parlata, la traduzione e il recupero delle informazioni sono necessarie in un numero crescente di sistemi di difesa, in particolare per il C4ISR e le operazioni congiunte multinazionali e/o di mantenimento della pace.

Queste tecnologie sono state oggetto di numerose ricerche per diversi decenni, che hanno portato a miglioramenti impressionanti per alcune applicazioni come gli assistenti vocali, i call center semi-automatici, i servizi di traduzione online, ecc. Tuttavia, anche se la maggior parte delle tecniche utilizzate nei sistemi attuali sono nate dalla ricerca nel campo della difesa, questi miglioramenti si basano anche sulla disponibilità di enormi quantità di dati, tipicamente da parte di attori internet con un'ampia base di utenti, e sono quindi avvenuti soprattutto nel settore civile. Nel settore della difesa, dove le quantità di dati che possono essere messe a disposizione degli sviluppatori sono molto più limitate per motivi di riservatezza, i miglioramenti sono stati anch'essi più limitati e sono ancora necessari progressi per soddisfare i requisiti della maggior parte delle applicazioni.

Per risolvere il problema della mancanza di disponibilità di dati operativi per gli sviluppatori di sistemi, una soluzione è quella di ricorrere a dati simili ma condivisibili. Tuttavia, ciò implica generalmente la creazione di dati ad hoc, che può essere relativamente costosa e non risolve completamente il problema. Una soluzione più adatta sarebbe che i sistemi potessero imparare direttamente dai dati degli utenti senza rivelare alcuna informazione riservata agli sviluppatori. Ciò non solo consentirebbe di utilizzare i dati del dominio finora inutilizzati e di rispondere alle principali preoccupazioni in materia di sicurezza e sovranità, ma porterebbe anche a un significativo progresso delle prestazioni grazie a un uso molto più efficiente dei dati esistenti. Inoltre, mentre esiste già un crescente numero di ricerche su approcci innovativi all'apprendimento automatico relativi a questo tema, come l'apprendimento semi-supervisionato e auto-supervisionato, l'apprendimento attivo, l'apprendimento di trasferimento e l'apprendimento frugale, incanalare questi sforzi attraverso un'organizzazione di sfide tecnologiche ha il potenziale per portare a una svolta. In questo contesto, il bando mira a creare non solo sistemi generici che offrano prestazioni migliori per un'ampia gamma di condizioni, ma anche sistemi che possano essere adattati dagli utenti per

offrire prestazioni migliori per applicazioni specifiche.

L'obiettivo generale del bando è quello di creare una libreria europea di componenti generici e adattivi di tecnologia del linguaggio umano che offrano prestazioni elevate per diverse applicazioni di difesa. In particolare, le tecnologie dovrebbero essere resistenti al rumore e alla qualità della comunicazione, coprire un'ampia gamma di lingue e dialetti, compresi quelli con scarse risorse, gestire un vocabolario specifico e offrire un'elaborazione più robusta delle informazioni semantiche di alto livello.

Ambito di applicazione:

Le proposte devono riguardare soluzioni tecnologiche per l'elaborazione di informazioni linguistiche nelle loro diverse forme, cioè parlate e scritte (scrittura a mano, documenti stampati o testi dattiloscritti), al fine di riconoscerle, comprenderle e tradurle. Queste soluzioni devono essere valutate nel quadro della sfida tecnologica organizzata nell'ambito di questo bando. Le proposte dovranno in particolare affrontare la questione dell'adattamento del sistema guidato dall'utente, ovvero la capacità dei sistemi di apprendere dalla supervisione degli utenti senza l'intervento degli sviluppatori e senza regredire in termini di prestazioni. Le tecnologie devono essere integrate in dimostratori con interfacce di facile utilizzo ed essere facilmente integrabili in altri sistemi di difesa.

Tipi di attività

La tabella seguente elenca i tipi di attività ammissibili per questo tema e se sono obbligatorie o facoltative (cfr. articolo 10, paragrafo 3, del regolamento FES):

Tipi di attività

(art. 10(3) Regolamento FES)

Ammissibile?

(a)

Attività che mirano a creare, sostenere e migliorare le conoscenze, i prodotti e le tecnologie, comprese le tecnologie dirompenti, che possono avere effetti significativi nel settore della difesa (generare conoscenza)

Sì (obbligatorio)

(b)

Attività volte ad aumentare l'interoperabilità e la resilienza, compresi la produzione e lo scambio sicuro di dati, a padroneggiare tecnologie critiche per la difesa, a rafforzare la sicurezza degli approvvigionamenti o a consentire l'efficace sfruttamento dei risultati per i prodotti e le tecnologie della difesa (integrazione delle conoscenze)

Sì (obbligatorio)

(c)

Studi, quali studi di fattibilità per esplorare la fattibilità di prodotti, tecnologie, processi, servizi e soluzioni nuove o aggiornate

Sì (facoltativo)

(d)

Progettazione di un prodotto della difesa, di un componente materiale o immateriale o di una tecnologia, nonché definizione delle specifiche tecniche su cui è stata sviluppata tale progettazione, compresi eventuali test parziali per la riduzione del rischio in un ambiente industriale o rappresentativo

Sì (facoltativo)

(e)

Prototipazione di sistema di un prodotto, componente materiale o immateriale o tecnologia della difesa

No

(f)

Test di un prodotto della difesa, di un componente materiale o immateriale o di una tecnologia

No

(g)

Qualificazione di un prodotto della difesa, di un componente materiale o immateriale o di una tecnologia

No

(h)

Certificazione di un prodotto della difesa, di un componente materiale o immateriale o di una tecnologia

No

(i)

Sviluppo di tecnologie o beni che aumentano l'efficienza nel ciclo di vita dei prodotti e delle tecnologie della difesa

No

Le proposte devono coprire almeno i seguenti compiti come parte delle attività obbligatorie:

Generazione di conoscenza:

ricerca sulle tecnologie del linguaggio umano, compresi gli approcci innovativi per l'adattamento dei sistemi guidato dall'utente;

partecipazione alle campagne di valutazione organizzate nell'ambito della sfida tecnologica, tra cui:

- scambio con altre parti interessate sui piani di valutazione;
- sottoporre i sistemi a misurazioni sperimentali delle prestazioni durante le campagne di test gestite dagli organizzatori della sfida;
- partecipazione a workshop di debriefing.

Integrazione delle conoscenze:

- integrazione dei moduli tecnologici in dimostratori che possano essere testati da utenti rappresentativi della difesa.

Le proposte devono includere descrizioni chiare dei criteri per valutare il completamento del pacchetto di lavoro. I criteri devono includere la partecipazione alle campagne di test organizzate nell'ambito della sfida tecnologica e la consegna delle descrizioni dei sistemi sottoposti ai test.

Requisiti funzionali

Le soluzioni proposte devono soddisfare i seguenti requisiti:

- I sistemi devono essere basati su componenti software che eseguono una serie di funzioni di elaborazione del linguaggio umano. Questi componenti devono essere integrati in dimostratori con un'interfaccia facile da usare. Dovrebbero consentire agli utenti di adattare i componenti utilizzando i propri dati, senza l'intervento degli sviluppatori del sistema.
- I sistemi dovrebbero essere in grado di funzionare localmente, senza connessione a una rete geografica, ad eccezione di funzioni specifiche per le quali ciò può essere debitamente giustificato ed è compatibile con le missioni operative (ad esempio, per ottenere prestazioni più elevate durante l'adattamento sotto la supervisione dell'utente).
- I sistemi devono essere ottimizzati in termini di memoria e di ingombro della CPU e, più in generale, devono disporre di risorse ragionevoli in termini di dimensioni dell'hardware, peso, prezzo e consumo energetico, in vista della loro potenziale integrazione in sistemi di difesa esistenti o futuri di maggiori dimensioni.
- I sistemi devono accettare come file di input informazioni linguistiche nelle loro diverse forme:
  - o parlato (in file audio o nel flusso audio di file video);
  - o documenti scritti a mano o stampati;
  - o testo.

Per ognuna di queste forme, i sistemi dovrebbero accettare un'ampia varietà di possibili input senza necessariamente avere informazioni sul tipo specifico di input. La variabilità può riguardare i parlanti o gli scrittori, lo stile di scrittura, il vocabolario, gli accenti, il rumore, le condizioni di registrazione o di scansione, i canali di trasmissione, ecc. I sistemi devono quindi essere indipendenti dal parlante/scrittore, indipendenti dal canale, resistenti a vari accenti, tipi di rumore, ecc. In particolare, devono essere resistenti alle condizioni che si verificano più frequentemente negli ambienti militari (ad esempio, forte rumore,

bassa qualità della comunicazione, parlato non madrelingua, ecc.)

L'ambito delle funzioni di elaborazione del linguaggio umano deve comprendere:

- identificazione della lingua dal parlato, dai documenti e dal testo;
- riconoscimento del parlato;
- riconoscimento di documenti scritti a mano e stampati;
- individuazione di parole chiave dal parlato e da documenti scritti a mano o stampati;
- traduzione dal parlato, da documenti scritti a mano o stampati e da testi;
- estrazione di informazioni semantiche di alto livello (comprese quelle specifiche del settore militare), come il riconoscimento di entità denominate ed eventi e l'estrazione di relazioni;
- recupero di informazioni multilingue;
- riassunto e visualizzazione automatica di (multi)documenti.

I componenti software corrispondenti alle funzioni di cui sopra devono coprire più lingue e dialetti, compresi quelli dell'UE e non UE. La traduzione deve coprire tutte le lingue ufficiali dell'UE come lingue di destinazione. Le proposte devono indicare l'elenco delle lingue e dei dialetti che si prevede di coprire per ciascuna funzione.

I componenti software devono offrire prestazioni all'avanguardia. Per ogni funzione di elaborazione del linguaggio umano di cui sopra e per ogni lingua o dialetto coperto, le proposte devono menzionare le prestazioni oggettivamente misurate (comprese le informazioni sui dati e le metriche utilizzate per le misurazioni e, se del caso, sulla campagna di valutazione nell'ambito della quale sono state effettuate le misurazioni) e i relativi riferimenti.

I componenti software corrispondenti alle funzioni coperte dalla sfida tecnologica organizzata nell'ambito del bando devono essere presentati per la valutazione. Qualsiasi differenza tra la versione valutata attraverso la sfida e la versione integrata nel dimostratore deve essere documentata. Le proposte devono descrivere come gli approcci e i sistemi proposti affronteranno i compiti delineati nel piano di valutazione preliminare (cfr. allegato 4). I componenti corrispondenti a funzioni non coperte dalla sfida possono essere adattati e migliorati durante l'esecuzione del progetto, se lo si ritiene opportuno, eventualmente utilizzando i dati della sfida.

I componenti software devono essere facilmente configurabili e integrabili in sistemi di difesa al di là dei dimostratori prodotti nell'ambito della sfida. Dovrebbero seguire il più possibile gli standard, le migliori pratiche e le linee guida pertinenti, comprese quelle elaborate a livello di sfida, in particolare per i formati di input e output.

I sistemi e le interfacce utente dovrebbero aiutare il più possibile gli utenti a capire come gli output sono derivati dagli input (AI spiegabile), e in particolare fornire collegamenti agli input. Ad esempio, le traduzioni dovrebbero essere accompagnate da collegamenti alla lingua di partenza a livello di termine o frase, e le visualizzazioni dovrebbero includere collegamenti agli input che supportano le informazioni visualizzate. Dovrebbero essere fornite anche le conoscenze utili a determinare l'output o che possono aiutare gli utenti a dare un senso agli input, come le voci di un dizionario bilingue per la traduzione.

I sistemi dovrebbero consentire agli utenti di adattarli utilizzando i propri dati, ad esempio fornendo lotti di dati grezzi o annotati, o fornendo interattivamente una supervisione.

Impatto previsto:

Il risultato dovrebbe contribuire a:

- un rafforzamento dell'EDTIB (European Defence Technological and Industrial Base) e una maggiore autonomia tecnologica per i sistemi HLT orientati alla difesa;
- un utilizzo più ampio, economico e semplice dei sistemi HLT per la difesa;
- sistemi di difesa potenziati in vari settori, in particolare C4ISR;
- una maggiore libertà d'azione dell'UE.

**Criteri di eleggibilità:** Per essere ammissibili, i richiedenti (beneficiari ed enti affiliati) devono:

- essere soggetti giuridici (enti pubblici o privati)
- essere stabiliti in uno dei Paesi ammissibili, vale a dire
- Stati membri dell'UE (compresi i Paesi e territori d'oltremare (PTOM))
- Paesi non UE: - Paesi SEE elencati ("Paesi associati al FES", cfr. elenco dei Paesi partecipanti)
- avere la struttura di gestione esecutiva stabilita nei Paesi ammissibili
- non devono essere soggetti al controllo di un Paese terzo non associato o di un'entità di un Paese terzo non associato (a meno che non siano in grado di fornire garanzie - cfr. Allegato 2 - approvate dallo Stato membro o dal Paese associato al FES in cui sono stabiliti)

I beneficiari e gli enti affiliati devono registrarsi nel Registro dei partecipanti - prima di presentare la proposta - e dovranno essere convalidati dal Servizio centrale di convalida (REA Validation). Per la convalida, sarà richiesto loro di caricare documenti che dimostrino lo status giuridico e l'origine. Altre entità possono partecipare in altri ruoli, come partner associati, subappaltatori, terzi che forniscono contributi in natura, ecc.

Si noti che, nel Fondo europeo per la Difesa, anche i subappaltatori coinvolti nell'azione e i partner associati devono soddisfare le condizioni di stabilimento e controllo sopra elencate. I partner associati che non sono stabiliti in uno dei Paesi ammissibili (o che sono soggetti al controllo di un Paese terzo non associato o di un'entità di un Paese terzo non associato) possono tuttavia partecipare in via eccezionale se sono soddisfatte alcune condizioni (non contravvenire agli interessi di sicurezza e difesa dell'UE e degli Stati membri; coerenza con gli obiettivi del FES; risultati non soggetti a controllo o restrizione da parte di Paesi terzi non associati o entità di Paesi terzi non associati; nessun accesso non autorizzato a informazioni classificate; nessun potenziale effetto negativo sulla sicurezza dell'approvvigionamento di fattori di produzione critici per il progetto), previo accordo dell'autorità concedente e senza alcun finanziamento nell'ambito della sovvenzione.

Composizione del consorzio

Le proposte devono essere presentate da almeno 3 candidati indipendenti (beneficiari; entità non affiliate) provenienti da 3 diversi paesi ammissibili.

Durata del progetto: 60 mesi I progetti di durata superiore possono essere accettati in casi debitamente giustificati. Le proroghe sono possibili, se debitamente giustificate e mediante un emendamento

I progetti che comportano informazioni classificate devono essere sottoposti a controlli di sicurezza per autorizzare il finanziamento e possono essere soggetti a norme di sicurezza specifiche (dettagliate in una lettera sugli aspetti di sicurezza (SAL) allegata alla convenzione di sovvenzione). I progetti in cui gli

Stati membri dei beneficiari partecipanti e le entità affiliate decidono di istituire un quadro di sicurezza specifico ai sensi dell'articolo 27, paragrafo 4, del regolamento FES saranno soggetti a tale quadro di sicurezza specifico e le nuove informazioni classificate (risultati) generate dal progetto saranno oggetto di approvazione da parte di tali Stati membri.

**Contributo finanziario:** Il budget previsto per il presente topic è di 18 milioni di euro

Il budget del progetto non può eccedere i 7 milioni di euro

La sovvenzione sarà una sovvenzione forfettaria. Ciò significa che rimborserà un importo fisso, basato su una somma forfettaria o un finanziamento non legato ai costi. L'importo sarà fissato dall'autorità che concede l'aiuto sulla base del bilancio stimato del progetto e dei tassi di finanziamento che dipendono dal tipo di attività e dai partecipanti.

I candidati devono disporre di risorse stabili e sufficienti per attuare con successo i progetti e contribuire con la loro parte. Le organizzazioni che partecipano a diversi progetti devono avere capacità sufficienti per attuare tutti questi progetti. La verifica della capacità finanziaria sarà effettuata sulla base dei documenti che vi verrà richiesto di caricare nel registro dei partecipanti durante la preparazione della sovvenzione (ad esempio, conto profitti e perdite e stato patrimoniale, piano aziendale, relazione di audit prodotta da un revisore esterno autorizzato, certificazione dei conti per l'ultimo esercizio finanziario chiuso, ecc.). L'analisi si baserà su indicatori finanziari neutri, ma terrà conto anche di altri aspetti, quali la dipendenza dai finanziamenti e dal disavanzo e dalle entrate dell'UE negli anni precedenti.

**Scadenza:** 22 Novembre 2023 17:00:00 Brussels time

**Ulteriori informazioni:**

[call-fiche\\_edf-2023-ls-ra-challenge\\_en.pdf \(europa.eu\)](#)

