



ALLEGATO 2

Area tematica di Ricerca

Terza Edizione - Bando pubblico per la selezione di proposte progettuali nell'ambito del PNRR, MISSIONE 4 "Istruzione e ricerca" – COMPONENTE 2 "Dalla ricerca all'impresa" – LINEA DI INVESTIMENTO 1.3 "Partenariati estesi alle università, ai centri di ricerca, alle aziende per il finanziamento di progetti di ricerca di base" – Programma di Ricerca e Innovazione "SERICS – GOVERNANCE E PROTEZIONE DEI DATI", PE00000014, SPOKE 10, CUP G43C22002580001, finanziato dall'Unione Europea – OPEN CALL INNOVATION - NextGenerationEU

A. Introduzione

I progressi nel campo delle Tecnologie dell'Informazione e della Comunicazione (ICT) e la disponibilità di una varietà di collezioni di dati offrono l'opportunità di accedere facilmente ed elaborare un'enorme quantità di informazioni, ottenendo così una descrizione estremamente dettagliata del comportamento delle persone e delle attività delle organizzazioni. La disponibilità di queste collezioni di dati, insieme alla grande capacità delle moderne infrastrutture di rete e di calcolo, offre chiaramente importanti vantaggi agli utenti, alle organizzazioni e alla società in generale. La raccolta, condivisione e analisi dei dati, con il contributo di soggetti diversi, è infatti un grande fattore abilitante per una società sempre più evoluta digitalmente.

Un chiaro ostacolo alla realizzazione di tale visione è rappresentato dalle preoccupazioni relative alla sicurezza e alla privacy. I dati possono essere sensibili o riservati e non possono essere condivisi apertamente. Più in generale, i dati possono essere soggetti a restrizioni di accesso e utilizzo (comprese quelle derivanti dalle normative sulla protezione dei dati) e la loro riservatezza, nonché la loro integrità, dovrebbero essere garantite anche quando i dati sono memorizzati o processati da terze parti. Senza una adeguata protezione dei dati esiste, infatti, un forte rischio di uso improprio, con minacce alla tutela della riservatezza e della privacy di cittadini, aziende e istituzioni. Per la piena realizzazione di una moderna società digitale, è quindi fondamentale fornire soluzioni efficaci per la governance e la protezione dei dati.

Spoke 10, organizzato in un singolo progetto (DGDP), risponde a questa sfida sviluppando soluzioni avanzate che consentono a diversi attori (ad es. individui, aziende e istituzioni) di mantenere il controllo sui propri dati in vari scenari di rilascio, condivisione e analisi. Le soluzioni permetteranno anche l'implementazione delle restrizioni dettate dalla legislazione e regolamentazione sulla protezione dei dati. Lo Spoke affronta il problema in modo olistico, esaminando diversi aspetti della governance e della protezione dei dati durante l'intero ciclo di vita dei dati e in vari scenari.

Il progetto DGDP mira allo sviluppo di nuove soluzioni per la gestione privata e sicura dei dati in scenari moderni ed emergenti, consentendo la condivisione dei dati e l'analisi collaborativa, permettendo ai proprietari dei dati di mantenerne il controllo. Il progetto è organizzato in quattro Work Package (WP), che affrontano diversi problemi di ricerca cruciali per la realizzazione di questo obiettivo.

Il presente documento è allegato al bando pubblico per la selezione di proposte progettuali per lo svolgimento della attività di ricerca industriale e sviluppo sperimentale riguardanti le tematiche oggetto della ricerca fondamentale condotta nel contesto di Spoke 10.



Dopo una sintesi degli aspetti principali del progetto, si passa alla descrizione in maggior dettaglio dell'area tematica, delle attività, della pianificazione e i dei deliverable da realizzarsi.

B. Struttura del Progetto DGDP in Work Package

Nell'ambito dello Spoke 10, il progetto DGDP è strutturato nei WP di seguito descritti.

a. WP1 - Data Governance Framework

WP1 si concentra sulla progettazione di un framework di governance per la gestione sicura e privata dei dati. Il framework includerà la definizione di soluzioni per la modellazione di dati e metadati per esprimere le proprietà che devono essere considerate per la regolamentazione dell'accesso e della condivisione dei dati, per le relative specifiche ontologiche e ragionamenti correlati, e per l'identificazione e l'analisi dei requisiti di protezione. L'obiettivo è creare un framework completo in grado di esprimere e ragionare su requisiti provenienti da diverse parti, inclusi quelli imposti da normative sulla privacy (come il GDPR) e quelli derivanti da bisogni delle organizzazioni e dagli individui. Le attività del WP1 sono organizzate nei seguenti task:

Task 1.1 - Modellazione di dati e metadati (UNIMI, UNIROMA1)

Task 1.2 - Ontologie e ragionamento sui metadati (UNIMI, UNICA, UNIROMA1)

Task 1.3 - Requisiti di protezione dei dati (UNIMI, LEONARDO)

b. WP2 - Politiche di protezione dei dati

WP2 si concentra sulla definizione di soluzioni per supportare la specifica e l'applicazione dei requisiti di protezione dei dati, attraverso linguaggi e modelli per esprimere e ragionare su tali requisiti e per supportare la valutazione e l'amministrazione delle politiche. L'obiettivo è quello di offrire un modello e un linguaggio semplici, per consentirne l'uso effettivo da parte di non esperti del settore, ma anche flessibili, estendibili ed espressivi, quindi in grado di rappresentare i diversi requisiti di protezione dei dati. Le attività del WP2 sono organizzate nei seguenti task:

Task 2.1 - Modello e linguaggio delle politiche (UNIMI, UNIFI, UNIROMA1)

Task 2.2 - Implementazione delle politiche (UNIMI, UNIFI, UNIROMA1)

Task 2.3 - Composizione, ragionamento e amministrazione delle politiche (UNIFI, UNIROMA1)

c. WP3 - Wrapping e sanitizzazione dei dati

WP3 si concentra sulla definizione di tecniche per implementare la sicurezza e la privacy durante tutto il ciclo di vita dei dati. Tali tecniche includono soluzioni per sanitizzare i dati e garantirne la privacy e la riservatezza in base a diverse metriche e rischi per la privacy ed inoltre anche tecniche di wrapping che consentono la protezione dei dati nella loro memorizzazione e analisi. Il WP3 si occupa anche del supporto della funzionalità di accesso e analisi sui dati protetti con considerazione di diversi scenari di condivisione dei dati e di nuove applicazioni di Intelligenza Artificiale (AI) e Machine Learning (ML). Il WP3 sviluppa inoltre soluzioni per consentire la valutazione delle garanzie di protezione e la verifica del corretto funzionamento del sistema. Le attività del WP3 sono organizzate nei seguenti task:

Task 3.1 - Metriche per la privacy e modelli di privacy (UNIMI)

Task 3.2 - Tecniche di sanitizzazione dei dati (UNIMI, UNICA, UNISA)

Task 3.3 - Wrapping dei dati e protezione dei dati per condivisione e analisi efficiente ed efficace (UNIMI, UNISA)



Task 3.4 - Interrogazione di, e computazione su, dati protetti (UNIMI, UNICA, UNISA)

Task 3.5 - Scenari applicativi di AI/ML (UNIMI, UNISA, LEONARDO)

Task 3.6 - Verifica e valutazione (UNIMI, UNISA, LEONARDO)

d. WP4 – Integrazione con tecnologie e modelli di gestione dei dati emergenti

WP4 si concentra sulle tecnologie emergenti di gestione dei dati, studiando l'applicazione della protezione dei dati a nuovi modelli e moderne piattaforme di gestione su larga scala, di scenari collaborativi distribuiti, garantendo l'applicazione efficiente ed efficace - e quindi la scalabilità e applicabilità - delle misure di protezione, con l'obiettivo di ridurre l'impatto sulla funzionalità di utilizzo dei dati. Le attività del WP4 sono organizzate nei seguenti task:

Task 4.1 - Nuovi modelli di dati (SMIMI)

Task 4.2 - Moderne piattaforme di gestione dei dati su larga scala (SMIMI)

Task 4.3 - Applicazione scalabile ed efficiente della protezione (SMIMI)

Task 4.4 - Scenari distribuiti e collaborativi (SMIMI)

C. Ambito e attività previste

L'ambito di riferimento del bando pubblico per la selezione di proposte progettuali a cui si riferisce il presente allegato riguarda lo svolgimento, nel contesto di riferimento dello Spoke 10 sopra introdotto, di attività mirate a:

- incrementare il Technology Readiness Level (TRL) delle attività di ricerca nell'ambito di riferimento;
- individuare, analizzare e ove fattibile realizzare applicazioni e strumenti (o parti e/o versioni di essi) volti a supportare le tematiche oggetto di ricerca nell'ambito di Spoke 10.

Saranno considerati e valutati i progetti che abbiano l'obiettivo di sviluppare idee, soluzioni, capacità avanzate o strumenti concreti inerenti a uno o più dei seguenti temi:

1. sviluppo di concetti, capacità avanzate e strumenti pratici per la data governance, quali la modellazione di dati e metadati, ontologie e ragionamento, definizione e gestione di requisiti di protezione dei dati.
2. definizione, implementazione e gestione di politiche di controllo degli accessi e di protezione dei dati.
3. definizione e gestione di metriche per la privacy, tecniche di sanitizzazione e wrapping dei dati, elaborazioni su dati protetti, aspetti di verifica e valutazione.
4. nuovi modelli dati, anche su larga scala e scenari distribuiti e collaborativi.

Saranno anche considerati eventuali ulteriori progetti purché afferenti ad ambiti di ricerca che rientrino nel contesto di Spoke 10.

Tra le attività finanziabili dal bando vi sono progetti pilota, dimostratori e/o esperimenti nell'ambito dei temi individuati.

Saranno particolarmente valorizzati i progetti che, raggiungendo un buon livello di TRL (almeno pari a quanto riportato nel paragrafo Deliverable), adottino un approccio integrato unendo tecnologie all'avanguardia o che



comunque afferiscano Emerging and Disruptive Technology (e.g., Intelligenza Artificiale) come definite da ACN, Agenzia Nazionale per la Cybersecurity, nell'ambito dell'Agenda di Ricerca e Innovazione per la cybersicurezza 2023-2026 e che rientrino nel contesto di Spoke 10.

D. Piano delle attività e milestone

Il piano di lavoro di Spoke 10 e del progetto DGDP è in linea con il piano generale del progetto SERICS riportato nell'Allegato C "cronoprogramma di attuazione e piano dei pagamenti" del Decreto Mur nr. 1556 del 11.10.2022 e successive modificazioni (Cronoprogramma).

Coerentemente con il Cronoprogramma, il progetto DGDP ha una durata prevista di 36 mesi come rappresentato in Figura 1, in cui sono riportate unicamente le attività e milestone maggiormente rilevanti per descrivere il contesto di riferimento.

Il progetto DGDP è stato avviato il 1° gennaio 2023 (evidenziato con T0 in Figura 1).

Si noti che la pianificazione riportata in Figura 1 non mostra il dettaglio dei Work Package in quanto sebbene esista una naturale interdipendenza tra di essi (ad esempio, i linguaggi/modelli delle politiche previste si baseranno anche sulla modellazione di dati e metadati e potranno ricorrere a soluzioni di wrapping e sanitizzazione per garantire la protezione), tutti i Work Package e i loro Task mirano allo sviluppo di nuove soluzioni di ricerca nel loro ambito e sono attivi in parallelo durante tutta la durata del progetto.

Milestone	Descrizione	T0	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36	
M1.0	Reclutamento Personale																																						
M11.0	Ricerca Fase 1																																						
M21.0	Ricerca Fase 2																																						

Figura 1. Cronoprogramma SERICS (estratto)

Le attività condotte dai Soggetti Beneficiari si collocheranno in una specifica milestone (M.31) del Cronoprogramma e la loro pianificazione dovrà attenersi al Cronoprogramma del progetto DGDP sopra descritto e alle disposizioni indicate nel bando a cui afferisce il presente allegato.

I Soggetti Beneficiari saranno tenuti a partecipare alle attività di coordinamento e controllo avanzamenti ed a effettuare la rendicontazione periodica delle spese ammissibili con la frequenza espressa nelle linee guida di rendicontazione di riferimento e comunque con cadenza almeno mensile non oltre il secondo giorno lavorativo successivo alla fine di ogni mese, sino alla completa rendicontazione di tutte le spese ammissibili di propria competenza. Contestualmente i Soggetti Beneficiari dovranno inoltre fornire una descrizione dell'avanzamento delle attività di ricerca condotte nel mese di riferimento che evidenzii le attività svolte nel mese, i risultati raggiunti, i deliverable realizzati (ove applicabile), secondo le modalità e standard definiti da UNIMI.

Le linee rosse verticali riportate nel Cronoprogramma di Figura 1 identificano i punti di controllo che separano i periodi di rendicontazione tecnico scientifica. Ulteriori punti di controllo sono da considerarsi con frequenza bimestrale ove non specificamente evidenziati da una linea rossa in Figura 1. In corrispondenza di ciascun punto di controllo, i soggetti che partecipano al progetto DGDP devono rendicontare i risultati tecnici e scientifici ottenuti nel periodo di riferimento. Allo stesso modo, i Soggetti Beneficiari dovranno fornire una relazione tecnica e scientifica per ciascuno dei punti di controllo che ricada all'interno del loro periodo progettuale, oltre che alla conclusione delle proprie attività, nei tempi e nelle modalità che saranno richieste da UNIMI. La



relazione tecnica e scientifica descriverà l'avanzamento delle attività nel periodo di rendicontazione, i risultati raggiunti, le attività e i deliverable in corso di realizzazione, secondo le modalità e standard definiti da UNIMI allegando alla relazione tutte le informazioni utili e le relative evidenze documentali.

I Soggetti Beneficiari si impegnano a partecipare con proprio personale alle attività di coordinamento e controllo degli avanzamenti e ad adeguare su richiesta di UNIMI la pianificazione delle proprie attività coerentemente ad eventuali modifiche che dovessero intercorrere riguardanti il Cronoprogramma.

E. Deliverable

Le attività svolte nell'ambito del presente bando produrranno soluzioni e relativi strumenti per la governance e la protezione dei dati.

I deliverable da realizzare includono le pubblicazioni scientifiche o altra documentazione che presenti i risultati della ricerca industriale e dello sviluppo sperimentale realizzati, eventuali prototipi, componenti software, dati o strumenti implementati in tale contesto.

Le pubblicazioni e altri deliverable realizzati dovranno necessariamente presentare il testo dell'acknowledgement SERICS e disclaimer seguenti: "This work was supported in part by project SERICS (PE00000014) under the NRRP MUR program funded by the EU - NGEU. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the Italian MUR. Neither the European Union nor the Italian MUR can be held responsible for them." e rispondere alle ulteriori specifiche presenti nel bando.

Spoke10 promuove il concetto di open science: i risultati finali dovranno essere sotto forma di documenti ove applicabile pubblicati in modalità Open Access, nel caso di software dovrà essere Open Source. La pubblicazione di eventuali dati dovrà seguire i principi FAIR (Findable, Accessibile, Interoperable, Reusable).

Per la disciplina della proprietà intellettuale si rimanda a quanto previsto dal bando.

Il livello target TRL (Technology Readiness Level) atteso è pari o superiore a 4.

I Soggetti Beneficiari realizzeranno inoltre i rendiconti tecnico scientifici corrispondenti ai punti di controllo (che ricadono all'interno del loro periodo progettuale) e contribuiranno alla realizzazione dei deliverable complessivi del Progetto DGDP integrando le informazioni afferenti ai risultati ottenuti nelle attività di propria competenza.