

Progetto e Piano di Formazione

Definizione e sviluppo di un supporto infrastrutturale data centric per la lavorazione, arricchimento e consolidamento di dati eterogenei

Progetto

Negli ultimi anni, il settore manifatturiero italiano si sta aprendo all'acquisizione di nuove competenze tecnologiche e ad una forte digitalizzazione industriale così come auspicato dalla quarta rivoluzione industriale (Industria 4.0 - I4.0). L'obiettivo principale della transizione digitale che sottende a I4.0 è la creazione di "smart factory" che sfruttano la digitalizzazione dei processi produttivi, e valorizzano i dati provenienti dalle linee di produzione, per conseguire elevati livelli di produttività e, al contempo, un aumento dell'efficienza dell'uso delle risorse a disposizione. A livello di filiera, lo sfruttamento pieno dei dati provenienti dai dipartimenti produttivi è limitato dalla difficoltà di gestione dall'enorme quantità di dati generata dai macchinari impiegati. A questa difficoltà si aggiunge anche l'eterogeneità dei dati prodotti dai molteplici dispositivi OEM. Per poter trattare una tale mole di dati eterogenei, occorre implementare un processo molto oneroso e con un alto livello di specializzazione, noto col nome di Data Preparation & Consolidation, che nello specifico consente di esplorare, combinare, ripulire, arricchire, filtrare, trasformare e validare i dati grezzi in set di dati curati e pronti per poter essere messi a disposizione di altri processi di elaborazione quali data discovery e business intelligence. L'adozione di una piattaforma avanzata di data preparation consentirebbe alle imprese di filiera di abbattere i tempi di preparazione dati, con conseguenti benefici in termini di aumento di produttività.

L'attività di questo assegno di ricerca si colloca all'interno del progetto PR-FESR **Data Transformation and Homogenization platform – DATA**. Le attività dell'assegnista riguarderanno lo studio, la progettazione e la prototipazione di una piattaforma software data centric per la Data Preparation & Consolidation di dati provenienti dal livello di Operation Technology (OT), shop floor, servendo applicazioni e servizi a livello IT. L'assegnista sarà impegnato nel design e nello sviluppo di una piattaforma Edge-Cloud in grado di sfruttare tecnologie per l'unificazione e l'arricchimento del dato, nonché tecniche di Machine Learning che consentiranno la creazione di nuovi modelli di analisi e il trasferimento dei modelli realizzati da una filiera produttiva ad un'altra. L'assegnista in particolare affronterà le seguenti tematiche:

- I. Analisi e valutazione dello stato dell'arte delle moderne e più recenti tecniche di manipolazione dati (es. data polishing, homogenization, aggregation, enrichment, ecc) con particolare focus sull'eterogeneità dei dati industriali;
- II. Analisi e valutazione dello stato dell'arte delle piattaforme Edge/Cloud, sia open source, che proprietarie, più utilizzate in ambienti industriali diversificati (es. EdgeX Foundry, Siemens Industrial Edge, Azure IoT, ecc) sulle quali progettare, implementare e installare le funzionalità di Data Curation;

- III. Analisi e valutazione della fattibilità tecnica dell'integrazione di queste tecnologie e funzionalità in ambienti industriali reali, nodi Edge/Cloud, e serviti da tecnologie di virtualizzazione a container (es Docker.);
- IV. Definizione e implementazione di una nuova architettura data centric compliant e capace di integrarsi con tutti o la maggior parte degli ambienti industriali studiati negli step precedenti e che sia capace integrare le funzionalità a microservizi.
- V. Identificazione di casi d'uso reali per la valutazione dell'efficacia dell'architettura, considerando come metriche principali le prestazioni, l'utilizzo di risorse e la possibilità di integrare logica intelligente nei cicli di controllo senza comprometterne le prestazioni.

Piano di Formazione

Il piano di formazione associato alle attività dell'assegno di ricerca prevede le seguenti attività organizzate nel periodo di un anno.

I semestre

Identificazione di casi d'uso rilevanti, con particolare focus sugli ambienti industriali e tenendo conto dei requisiti forniti dai partner del progetto DATHA.

Analisi dello stato dell'arte relativo alle tematiche di preparazione dati, consolidamento e interoperabilità di processi e dati.

Definizione dei requisiti tecnici e funzionali di data preparation & consolidation attraverso il coinvolgimento utenti finali e degli exploiters del progetto DATHA.

Progettazione e sviluppo di microservizi per la preparazione dei dati produttivi e validazione degli output provenienti dai microservizi per la preparazione dei dati

Integrazione dei microservizi con architetture Edge-Cloud per interoperabilità di processi e dati.

Analisi dello stato dell'arte di architetture Edge/Cloud innovative basate su microservizi altamente scalabili basate su tecnologie Docker/Kubernetes.

Il semestre

Analisi e individuazione delle attività di preparazione e pulizia dei dati maggiormente impattanti sul processo di preparazione dei dati.

Sviluppo di moduli intelligenti per il riconoscimento e trasformazione dei dati produttivi.

Testing e Validazione dei moduli intelligenti per la preparazione dei dati produttivi.

Progettazione e sviluppo di una piattaforma scalabile ed elastica per il providing, la gestione, e il supporto alle funzionalità di preparazione, pre-processing, e analytics del dato basato su tecniche di AI/ML.

Validazione delle performance della piattaforma in termini di aderenza del modello dei dati con quelli forniti dagli exploiter industriali del progetto DATHA, interoperabilità e latenza dei microservizi.

Integrazione del prototipo di piattaforma software sulla linea pilota di Bi-Rex per implementazione di use case aziendali in scenari di I4.0.