

Progetto e Piani di Formazione

“Sviluppo di soluzioni a supporto della realizzazione di un’infrastruttura di calcolo Cloud-HPC e dell’implementazione una piattaforma software di servizi per la gestione di beni culturali digitali”

Progetto

Le attività di questi assegni di ricerca, della durata di 12 mesi, si focalizzeranno primariamente sulla progettazione di soluzioni tecnologiche da impiegare a supporto del piano operativo di sviluppo di una piattaforma di servizi per la catalogazione e l’elaborazione di beni culturali digitali (ISPC). La piattaforma software poggerà su un’infrastruttura di calcolo che sfrutta risorse computazionali di tipo Cloud e HPC. Verranno a tale scopo impiegate tecniche di virtualizzazione che garantiranno un accesso uniforme alle risorse fisiche eterogenee sottostanti. La piattaforma ISPC disporrà di un Data Lake per l’immagazzinamento strutturato ed efficiente di notevoli quantità (Big Data) di beni culturali nativamente digitali e digitalizzati, e fornirà all’utente finale strumenti per la navigazione e l’elaborazione dell’informazione che sfruttano tecnologie semantiche e AI-based. Da un punto di vista implementativo, la piattaforma adotterà il paradigma architetturale a microservizi, che garantisce flessibilità durante le fasi di sviluppo e conferisce scalabilità nella fase operativa.

Per le attività di progettazione di cui sopra, saranno considerate centrali tecnologie disponibili sul mercato e in letteratura relativamente a: a) virtualizzazione “pesante” e “leggera” di risorse computazionali (Hypervisor-based, containerization); b) gestione dello storage strutturata e non strutturata (SQL, NOSQL, Object storage, graph DB); c) soluzioni ML/DL-based per data-mining, text-mining, natural language processing (NLP) e object-detection in ambito computer vision; d) tool-chains per lo sviluppo e l’operatività di sistemi software secondo il paradigma DevOps.

Il progetto si focalizzerà su due obiettivi. Il primo consiste nella progettazione, realizzazione e test di un prototipo di Platform as a Service (PaaS), ispirata al paradigma architetturale “Data Fabric”, che offra all’attore utilizzatore strumenti visuali per l’elaborazione dei dati contenuti nella piattaforma ISPC e, al contempo, fornisca all’attore sviluppatore servizi per sviluppare nuove funzionalità da aggiungere alla piattaforma stessa. Per tali scopi, si ritiene che l’assegnista debba possedere conoscenze ed abilità nella gestione ed integrazione di dati eterogenei, interfacciamento ed integrazione di servizi, sviluppo di microservizi distribuiti in ambito Cloud, tecniche e tecnologie per l’implementazione di politiche di sicurezza a supporto dei dati e dell’accesso ai servizi. Il secondo obiettivo si concretizza nella progettazione, realizzazione e test di un prototipo di middleware che offra allo sviluppatore strumenti per implementare agevolmente applicazioni AI-based in grado di sfruttare la potenza di calcolo sia del Cloud che di infrastrutture HPC. Per tali scopi, si ritiene che l’assegnista debba possedere conoscenze ed abilità nello sviluppo di microservizi distribuiti in ambito Cloud, conoscenze di strumenti e librerie per l’implementazione di servizi AI-based, conoscenze di tecniche di sviluppo in ambienti di interactive programming, conoscenza di tecnologie e tecniche per l’implementazione di politiche di sicurezza a supporto dei dati e dell’accesso ai servizi.

Il programma di formazione dei due assegni prevede le seguenti attività:

- Ricerca sullo stato dell'arte: necessaria per la raccolta di tutte le informazioni necessarie e tool esistenti per la realizzazione di middleware per lo sviluppo di applicazioni AI-based su infrastrutture di calcolo tipo Cloud/HPC
- Progettazione di un middleware in grado di integrare risorse di tipo Cloud e risorse di tipo HPC in un "continuum" runtime, e offrire allo sviluppatore un toolkit per lo sviluppo di applicazioni ML/DL-based nel succitato continuum
- Sviluppo software e test sul campo di un prototipo software del middleware e di un insieme di applicazioni di esempio
- Ricerca sullo stato dell'arte: necessaria per la raccolta di tutte le informazioni necessarie e tools open source esistenti utili allo sviluppo di servizi a supporto della PaaS
- Progettazione di un'architettura di Data Fabric efficiente, performante e scalabile finalizzata alla gestione/elaborazione dei dati provenienti dalla piattaforma ISPC
- Sviluppo software e test sul campo di un prototipo software di architettura di di Data Fabric

Sono inoltre previste attività di gruppo orientate alla predisposizione di materiali e prototipi. Gli assegnisti matureranno inoltre esperienze significative relativamente a: redazione di documenti di carattere divulgativo e presentazione di risultati e prototipi alla comunità scientifica e ad aziende del territorio.

Piano di formazione assegno n. 1

Il piano di formazione associato all'assegnato di ricerca n. 1 mira a conseguire il primo obiettivo di progetto e prevede le seguenti attività organizzate nel periodo di un anno.

I semestre

- Studio dello stato dell'arte e degli strumenti open source a supporto della realizzazione del middleware.
- Collaborazione con i referenti della piattaforma ISPC per identificare uno scenario ed un set di casi d'uso su cui sarà necessario validare l'output di questo progetto. Occorrerà contestualmente definire i requisiti funzionali del middleware e i livelli di QoS che questo dovrà garantire.
- Definizione di un'architettura di middleware e del relativo modello a componenti. Definizione delle interfacce e delle specifiche di integrazione del middleware sia con i servizi della piattaforma ISPC che con l'infrastruttura HPC.

Il semestre

- Implementazione del middleware in modo incrementale, attraverso cicli di prototipazione e valutazione sperimentale.
- Integrazione del middleware con servizi dalla ISPC e con le interfacce messe a disposizione dell'infrastruttura HPC.
- Valutazione del risultato ottenuto in un ambiente di scala ridotta. In tale ambiente verranno riprodotti lo scenario e i casi d'uso definiti durante il primo semestre con i referenti della ISPC.
- Definizione delle linee guida di utilizzo del middleware a beneficio degli sviluppatori applicazioni AI-based

Piano di formazione assegno n. 2

Il piano di formazione associato all'assegno di ricerca n. 2 mira a conseguire il secondo obiettivo di progetto e prevede le seguenti attività organizzate nel periodo di un anno.

I semestre

- Studio dello stato dell'arte e degli strumenti open source a supporto dell'implementazione di piattaforme software ispirate ai paradigmi architetturali di tipo Data Fabric e Data Mesh.
- Collaborazione con i referenti della piattaforma ISPC per identificare uno scenario ed un set di casi d'uso su cui sarà necessario validare l'output di questo progetto. Occorrerà contestualmente definire i requisiti funzionali della PaaS e i livelli di QoS che questa dovrà garantire.
- Definizione di un'architettura di piattaforma PaaS e del modello a componenti che dovrà sostenere il paradigma di Data Fabric. Definizione delle interfacce e delle specifiche di integrazione della PaaS con i servizi e i dati forniti dalla ISPC.

II semestre

- Implementazione della piattaforma PaaS in modo incrementale, attraverso cicli di prototipazione e valutazione sperimentale.
- Integrazione della piattaforma PaaS con i servizi e i dati forniti dalla ISPC.
- Valutazione del risultato ottenuto in un ambiente di scala ridotta. In tale ambiente verranno riprodotti lo scenario e i casi d'uso definiti durante il primo semestre con i referenti della ISPC.
- Definizione delle linee guida di utilizzo della piattaforma PaaS, a beneficio degli utilizzatori, e delle linee guida di implementazione a beneficio degli sviluppatori di nuovi servizi