



1506  
UNIVERSITÀ  
DEGLI STUDI  
DI URBINO  
CARLO BO



Finanziato  
dall'Unione europea  
NextGenerationEU

## PROGETTO DI RICERCA FINANZIATO AI SENSI DEL DM 351 DEL 9 APRILE 2022

**Missione 4, Componente 1- Investimento 4.1** “*Estensione del numero di dottorati di ricerca e dottorati innovativi per la pubblica amministrazione e il patrimonio culturale*”.

M4C1 - Inv. 4.1 – n. 1. borsa Dottorati PNRR

### **Titolo del progetto formativo della borsa di dottorato**

Progettazione e Sviluppo di un Sistema Intelligente per l'Eleborazione e l'Analisi di Dati Citometrici. / *Design and Development of an Intelligent System for Flow Cytometry Data Processing and Analysis.*

**Docente referente per il progetto: Prof. Stefano Papa**

### **Descrizione del progetto formativo della borsa di dottorato**

- Obiettivi del progetto di ricerca;
- Coerenza del corso di dottorato con gli ambiti tematici del PNRR
- Coerenza del progetto proposto con la Misura scelta (1. Dottorati dedicati alle transizioni digitali e ambientali; 2. Dottorati PNRR; 3. Dottorati per la PA; 4) Dottorati per il patrimonio culturale) e relativa tematica
- Impatto della ricerca proposta in relazione a uno o più dei seguenti fattori: (i) miglioramento della sostenibilità ambientale; (ii) accelerazione di processi di trasformazione digitale; (iii) promozione dell'inclusione sociale;
- Breve descrizione dell'attività formativa e di ricerca che il dottorando dovrà svolgere in coerenza con il progetto di ricerca proposto anche in considerazione dell'eventuale percorso congiunto che il dottorando beneficiario della borsa svolgerà presso il soggetto (impresa/ente) coinvolto nel percorso dottorale.
- Attività di disseminazione e comunicazione dei risultati nell'ottica di una valorizzazione dei risultati della ricerca e della tutela della proprietà intellettuale che assicuri accesso aperto al pubblico ai risultati della ricerca e ai relativi dati nel minor tempo e con il minor numero di limitazioni possibile, secondo i principi “Open science” e “Fair data”.
- Garantire il rispetto dei principi orizzontali del PNRR.

#### **Descrizione del progetto di ricerca, in lingua italiana:**

Nel campo della Citometria, la condivisione delle informazioni sotto forma di file particolarmente rappresentativi di una patologia conosciuta è un campo di estrema importanza e, purtroppo, rappresentato, allo stato attuale, solamente dal progetto ESCCABase. Lo sviluppo continuativo di questa piattaforma, costituita da un sistema client-server dedicato, è quindi di particolare importanza, allo scopo di realizzare un vero e proprio atlante interattivo di riferimento per i professionisti nel campo della Citometria. Alcune implementazioni previste nel corso di questo progetto comprendono: la possibilità di modificare i parametri di trasformazione matematica, la possibilità di alterare i parametri di compensazione, l'aggiunta di nuove



raccomandazioni grafiche degli eventi, la possibilità di importare alcuni formati non conformi al formato FCS, la possibilità di alterare le decadi delle scale iperlogaritmiche ed altre implementazioni essenziali per completare le funzioni della piattaforma.

A queste funzionalità base, in questo progetto si propone di aggiungere una componente software intelligente, sviluppata allo scopo di effettuare *gating* automatico e diagnosi di patologie ematiche. Tale componente sfrutterà algoritmi e modelli sviluppati nell'ambito dell'intelligenza artificiale (AI). Negli ultimi anni, il numero di lavori scientifici che riportano applicazioni dell'AI in ambito medico sono cresciuti esponenzialmente, dimostrando che l'AI può essere un valido strumento per lo sviluppo di sistemi a supporto del medico [1]. Tra gli altri, gli algoritmi di machine learning (ML) sono quelli che vedono il più ampio utilizzo, essendo progettati per eseguire previsioni altamente accurate sul fenotipo, sulla diagnosi o sulla prognosi di un soggetto, senza essere esplicitamente programmati ma apprendendo dai dati un modello del sistema. Nel contesto della citometria, la maggior parte del lavoro disponibile in letteratura si concentra sull'applicazione di tecniche di deep learning ai dati di immagini di citometria a flusso, mentre l'applicazione del ML ai dati relativi ai marcatori fisici e molecolari estratti è ancora agli albori, principalmente limitata dalla necessità di generalizzare i risultati su dataset eterogenei e di fornire una corretta interpretazione ai risultati delle analisi [2]. In questo progetto proponiamo di applicare diversi algoritmi di ML su dati citometrici per automatizzare l'estrazione e l'identificazione di sottopopolazioni cellulari, migliorando così la velocità e l'accuratezza del *gating* manuale, che è fortemente limitato dall'impossibilità di gestire manualmente dati multidimensionali, e affetto dall'errore umano. I modelli di ML, allenati e testati su un dataset eterogeneo, saranno combinati con sistemi basati su regole, incorporando le conoscenze esistenti nel campo e superando i limiti della maggior parte degli approcci disponibili in letteratura che, lavorando solo su dati grezzi, non riescono ancora a fornire buone prestazioni e spiegazioni della classificazione prevista. I dati utilizzati nel progetto sono provenienti dal laboratorio analisi dell'Ospedale di Urbino, dal Centro di Citometria dell'Università di Urbino e dal Laboratory of Cytomics, Centro de Investigación Príncipe Felipe, Valencia. L'obiettivo finale è quello di rendere disponibile il sistema così sviluppato e validato, come strumento integrato nella piattaforma ESCCABase.

La proposta si inserisce nel dottorato in "Biomolecular and Health Sciences" dell'Università degli Studi di Urbino Carlo Bo, i cui contenuti rientrano nelle discipline delle scienze della vita, con particolare fuoco sulla definizione di metodologie innovative e tecnologiche in linea con gli obiettivi PNRR. La ricerca presentata in questo progetto si inserisce in questo contesto, proponendo l'applicazione di modelli e algoritmi sviluppati nell'ambito della computer-science a problematiche di natura prima di tutto biologica, nell'automatizzare l'analisi di dati di popolazione cellulare, ma anche clinica, laddove il progetto sviluppa un sistema a supporto della diagnosi. In quanto tale risulta essere un progetto altamente interdisciplinare a forte vocazione scientifico-tecnologica.

Verrà promossa l'adesione a reti nazionali ISCCA (Italian Society for Cytometric Cell Analysis) e internazionali ESCCA (European Society for Clinical Cell Analysis) e ISAC (International Society for Advancement of Cytometry) e l'inter-settorialità grazie al contributo sia dei membri del Collegio, sia di colleghi dei settori scientifici dell'Area 9, sia alle collaborazioni internazionali.

Il progetto è in piena coerenza con i pilastri del Next Generation EU e soddisfa largamente gli obiettivi trasversali del PNRR nel definire metodologie e strumenti per la transizione digitale e il trasferimento tecnologico, e della sesta missione del PNRR relativa alla Salute, promuovendo l'utilizzo di tecnologie innovative in medicina.

Le azioni formative riguardanti il dottorato di ricerca consistono prima di tutto in attività didattiche formative svolte presso l'Università degli Studi di Urbino Carlo Bo. Tali attività consistono in seminari e corsi obbligatori previsti dal programma di dottorato quali biostatistica, bibliometria, proprietà intellettuali/innovazione industriale/trasferimento dell'innovazione; qualità della ricerca, etica della ricerca, informatica e corsi di miglioramento della lingua inglese. Il dottorando potrà inoltre scegliere qualche insegnamento erogato presso strutture e scuole dell'Ateneo, i cui contenuti siano coerenti con il progetto e le finalità del dottorato. E' inoltre prevista la partecipazione a congressi e conferenze nazionali e internazionali, momenti ritenuti cruciali per la formazione del dottorando e per la diffusione dei risultati della ricerca.

Il/la dottorando/a opererà prevalentemente presso il Dipartimento di Scienze Biomolecolari, che mette a disposizione le strumentazioni e le competenze biologiche e cliniche necessarie alla ricerca proposta, ma



svolgerà la sua attività di ricerca anche in stretta collaborazione con i ricercatori del Dipartimento di Scienze Pure e Applicate che metteranno a disposizione le conoscenze tecniche e teoriche necessarie alla comprensione e utilizzo degli modelli, degli algoritmi e degli strumenti informatici alla base del progetto. Sono inoltre previsti 6 mesi di ricerca presso il Laboratory of Cytomics, Centro de Investigación Príncipe Felipe, Valencia, nei quali il dottorando avrà modo di sperimentare gli strumenti sviluppati su dati raccolti con metodologie diverse al fine di validare lo strumento su un dataset quanto più eterogeneo possibile.

Verrà favorita la valorizzazione dei risultati della ricerca e garantita la tutela della proprietà intellettuale, secondo i principi "Open science" e "FAIR Data", attraverso pubblicazioni scientifiche Open Access e attraverso la pubblicazione del software Open Source con licenza GNU GPLv3.

[1] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," Nature Medicine, vol. 25, no. 1, pp. 44–56, 2019.

[2] Z. Hu, S. Bhattacharya and A. J. Butte, "Application of Machine Learning for Cytometry Data," Frontiers in Immunology, vo. 12, art. 787574, 2022.

## PERIODO IN AZIENDA

Come richiesto dal DM 351/2022 e, nello specifico, dalla misura scelta, il progetto di ricerca sarà svolto in collaborazione con il seguente soggetto di cui in convenzione:

Ragione sociale: Prodotti Ganni srl  
Sede legale: Via Quintiliano 30 - 20138 - Milano  
Rappresentante legale: Dott. Luigi Bondurri

L'ente sopra citato ospiterà il dottorando beneficiario della borsa finanziata sulle risorse del DM 351/2022 periodo di n. \_\_\_6\_\_\_ mesi (min 6 - max 12) nel corso del dottorato.

## PERIODO ALL'ESTERO:

Il progetto di ricerca prevede inoltre un periodo all'estero di n°\_6\_mesi (min 6 - max 18) presso la seguente istituzione:

Prof. Jose-Enrique O'Connor Laboratory of Cytomics, Centro de Investigación Príncipe Felipe, Valencia, Spain

Si dichiara inoltre che il presente progetto è conforme al principio "di non arrecare un danno significativo" (DHS) ai sensi dell'art. 17 del regolamento (UE) 2020/852 in coerenza con gli orientamenti tecnici predisposti dalla Commissione Europea (Comunicazione della Commissione Europea 2021/C58/01) e garantisce il rispetto dei principi orizzontali del PNRR (contributo all'obiettivo climatico e digitale c.d. tagging, il principio della parità di genere e l'obbligo di protezione e valorizzazione dei giovani).

Urbino 16 giugno 2022

FIRMA DEL COORDINATORE  
Prof. Marco Bruno Luigi Rocchi



1506  
UNIVERSITÀ  
DEGLI STUDI  
DI URBINO  
CARLO BO



Finanziato  
dall'Unione europea  
NextGenerationEU

FIRMA DEL DOCENTE PROPONENTE/RESPONSABILE SCIENTIFICO

Prof. Stefano Papa