

# **DOTTORATO DI RICERCA IN “BIG DATA ED INTELLIGENZA ARTIFICIALE”**

## **Caratteristiche del Corso**

**Ciclo:** XXXIX

**Coordinatore:** Prof.ssa Barbara Martini – SSD ING-INF/05

**Sede Amministrativa:** Piazza Mattei, n° 10 – Roma (RM) 00186

**Data presunta di inizio del corso:** 01 dicembre 2023

**Durata:** 3 anni

**Aree CUN:** 06 – Scienze mediche, 08 – Ingegneria civile e Architettura, 09 – Ingegneria industriale e dell’informazione, 12 – Scienze giuridiche, 13 – Scienze economiche e statistiche, 14 – Scienze politiche e sociali.

### **Curricula:**

- Big data management per la transizione digitale
- Intelligenza artificiale per industria 5.0 ed economia circolare

### **Presentazione del progetto:**

Il corso di dottorato “Big Data e Intelligenza artificiale” forma ricercatori con una preparazione scientifica sia nell’ambito Data Science, indirizzata all’approfondimento di aspetti metodologici, quali l’analisi statistica e computazionale dei dati, la gestione dei big data e gli algoritmi di machine learning, sia orientata ad una conoscenza teorica ed applicativa dell’Intelligenza Artificiale, con particolare riferimento all’informatica e alla matematica applicata, in grado di abbinare una forte capacità analitica alla capacità di sviluppo algoritmico.

I Big Data sono ovunque e influenzano le scelte strategiche dell’impresa. Il termine Big Data si riferisce letteralmente alla grande quantità di dati ed informazioni che vengono acquisite e gestite quotidianamente da società o enti. Più che l’entità di questi dati, però, ciò che attira l’attenzione è il loro utilizzo ovvero come essi possano essere analizzati in modo da estrapolare informazioni importanti per tutte quelle aziende che si occupano di statistiche ed analisi di mercato.

I Big Data non avrebbero alcun valore se non fosse possibile analizzarli ed estrapolare informazioni: le macchine rispetto agli esseri umani hanno la possibilità di analizzare una grande quantità di informazioni in tempi brevissimi. In altri termini, il semplice fatto di avere accesso a grandi serie di dati non è sufficiente per produrre un risultato. La convergenza tra Big Data e Intelligenza Artificiale è pertanto inevitabile anche per rendere più efficienti e intelligenti i processi aziendali.

Il processo produttivo dei Big data è ancora da conoscere compiutamente e non riguarda solo

i confini disciplinari tra statistica, informatica, scienze sociali ed economiche per la creazione di una data science interdisciplinare. Serve una appropriata statistical literacy che deve estendersi anche ai dati provenienti dai nuovi strumenti di comunicazione per una corretta interpretazione dei dati. L'analisi dei Big data deve infatti tenere conto della qualità dei dati per non incorrere in rappresentazioni errate e falsate della realtà.

La disponibilità di grandi quantità di dati favorisce la convergenza di discipline diverse per lo sviluppo di modelli ed algoritmi in grado di spiegare più a fondo la complessità dei fenomeni sociali, economici, biologici, naturali, tecnologici e culturali. In particolare, l'utilizzo dei Big Data insieme ad algoritmi di Intelligenza Artificiale consente di esaminare i dati al fine fornire analisi concrete e riutilizzabili, individuare correlazioni nascoste e trend ancora sconosciuti e, in ultima analisi, prendere decisioni più ponderate ed efficaci, soprattutto nell'ambito dei processi produttivi e dell'automazione industriale anche in chiave di cybersecurity.

Sul solco della generazione precedente, l'Industria 5.0 intende cogliere le opportunità legate allo sviluppo dell'Intelligenza Artificiale non tanto (e non solo) per migliorare l'efficienza e la flessibilità del processo produttivo ma, andando oltre, per sostenere l'industria nel miglioramento delle condizioni di lavoro per l'uomo e, in generale, nel suo servizio a lungo termine all'umanità e al pianeta.

In questa ottica di valorizzazione dell'umano e dell'ambiente, Industria 5.0 con Intelligenza Artificiale (in combinazione con IoT, Cloud Computing e Reti ad alta velocità) giocano un ruolo chiave per garantire adeguati circuiti di feedback tra la trasformazione industriale e una rivalutazione del capitale, compresi i flussi di capitale umano e sociale, in altre parole, per lo sviluppo di una Economia Circolare come modello economico rigenerativo.

In questo contesto, il programma di studio fornirà competenze da applicare a diversi contesti, con l'obiettivo di creare esperti che contribuiscano all'avanzamento di nuovi saperi scientifici in grado non solo di elaborare metodi e modelli di interpretazione dei dati e di maneggiare le più moderne tecnologie in ambito IoT, Cloud e Reti ad alta velocità, ma anche di fornire risposte a nuovi campi di frontiera di ricerca della scienza e soluzioni di sistema a molte delle sfide globali, in primis quello di sviluppo sostenibile dell'economia.

I corsi sono progettati per fornire una solida preparazione multidisciplinare e interdisciplinare nelle diverse aree di ricerca che fanno parte della Data Science e dell'Intelligenza Artificiale, inclusi aspetti statistici, tecnologici e di cybersecurity. La forte impronta interdisciplinare è funzionale a un posizionamento nel mondo del lavoro estremamente versatile che risponda al meglio alle sfide odierne. L'attenzione agli aspetti applicativi e la stretta interazione con il territorio e il mondo produttivo consentiranno di mettere a frutto le conoscenze acquisite in aziende, enti di ricerca e pubbliche amministrazioni, rispondendo alla loro necessità di innovare e di trovare soluzioni di ampio respiro a problemi sempre più complessi e, spesso, anche con diversi portatori di interesse coinvolti.

### **Obiettivi del corso:**

L'obiettivo del corso è quello di creare figure di elevata qualificazione, in grado di esercitare attività di ricerca specifica e qualificata e/o di perseguire progetti di ricerca applicata e di innovazione tecnologica nel campo dei Big Data e dell'Intelligenza Artificiale nell'ottica sia della Transizione digitale che dell'Industria 5.0 ed Economia Circolare. La formazione di elevato livello e ampio respiro consentirà una più facile collocazione dei futuri dottori di

ricerca nelle imprese che basano i loro processi produttivi sull'innovazione e sul digitale.

Il corso intende colmare un deficit di competenze negli ambiti Big Data e Intelligenza artificiale soprattutto all'interno del mondo aziendale dove si osserva una carenza di talenti e innovatori. Per colmare questo deficit, il corso formerà una classe d'eccellenza di ricercatori in grado di integrare conoscenze in ambito statistico, tecnologico ed economico e di trovare dei punti di contatto per formulare soluzioni di ampio respiro e rispondere più efficacemente alle sfide odierne legate alla complessità e alla presenza di diversi interessi in gioco delle parti interessate.

In particolare, gli obiettivi possono essere sintetizzati come segue:

- abbinare una formazione nelle diverse aree della Data Science e Intelligenza Artificiale con una preparazione multidisciplinare estesa alle materie tecnologiche ed economiche;
- favorire un approccio interdisciplinare allo studio e alla ricerca per consentire analisi di sistemi complessi in un ambito caratterizzato da forti interazioni tra aspetti tecnici, economici e sociali connessi all'utilizzo dei dati;
- garantire forte caratterizzazione industriale sul percorso formativo del dottorando, anche attraverso periodi di permanenza all'estero in centri di ricerca e realtà industriali di estremo rilievo nel panorama nazionale e internazionale;
- fornire competenze trasversali per sviluppare autonomamente un'attività di ricerca e orientare la ricerca tenendo conto dello stato dell'arte, valutando le alternative metodologie nonché le prospettive e le ricadute sul benessere dell'uomo e della società.

I curricula previsti sono 2 con i seguenti obiettivi:

- Big Data Management per la transizione digitale: l'obiettivo è focalizzato sull'intera gestione del ciclo di vita del dato, dalla raccolta alla sua conservazione, passando per l'analisi, richiede l'impiego di nuove tecnologie innovative che abilitino l'estrazione di valore dai Big Data. I Big Data applicati al tema dell'innovazione e della trasformazione digitale, hanno aperto nuovi scenari e interessanti opportunità per le aziende.
- Intelligenza artificiale per Industria 5.0 ed Economia circolare: l'obiettivo è sfruttare l'Intelligenza Artificiale in contesti industriali per favorire nuovi processi produttivi in cui anche le tecnologie (IoT, robot, cybersicurezza, cloud e reti ad alta velocità) sono utilizzate non solo per perseguire una maggiore efficienza, ma soprattutto migliori condizioni di vita dell'uomo e dell'ambiente anche attraverso un modello rigenerativo dell'economia.

In ottica di interdisciplinarietà, i dottorati su due curricula potranno eventualmente svolgersi in sinergia qualora siano finanziati dalla stessa azienda e indirizzino lo stesso tema di ricerca trattandolo da due prospettive differenti, ovvero statistico e ingegneristico-tecnologico. Il perseguimento di questi obiettivi sarà continuamente monitorato attraverso una attenta supervisione del lavoro dei dottorandi e supportato attraverso attività di formazione e appropriate strutture e attrezzature. Il programma di formazione prevede corsi per acquisire sia competenze tecniche nell'ambito del curriculum di riferimento che competenze trasversali. Il programma dei corsi prevede anche attività di laboratorio in didattica interattiva. Le verifiche finali quando previste si svolgeranno in presenza mediante test a risposte aperte. Il programma di studio prevede anche un periodo di permanenza all'estero e di esperienza in

ambito industriale. Gli studenti potranno anche usufruire di appositi spazi in sede per favorire attività di ricerca di gruppo e di cross-fertilizzazione e per permettere loro di essere seguiti al meglio dai tutori di riferimento (un docente universitario e un professionista aziendale). L'attività di ricerca dei dottorandi sarà sottoposta a continua verifica sia interna con report a fine di ogni anno e con una presentazione a metà percorso al collegio docenti sullo stato di avanzamento della ricerca, che esterna favorendo la presentazione dei lavori di ricerca a conferenze e workshop, nonché la pubblicazione degli stessi su riviste specializzate. Inoltre gli studenti potranno contare su utilizzo di attrezzature come i laboratori virtuali di Infocamere o attrezzature delle aziende con cui Universitas Mercatorum ha stipulato specifiche convenzioni. Infine per le sperimentazioni su larga scala gli studenti saranno guidati all'uso delle e-infrastrutture europee per la ricerca.

### **Sbocchi occupazionali e professionali previsti**

I dottori di ricerca in Data Science e Intelligenza Artificiale è il vero interprete della data economy e del data-driven business e grazie alla diffusione di tecnologie come Analytics e Machine Learning aiuta la direzione aziendale a prendere decisioni mirate. Saranno in grado di coniugare le skill tecniche trasversali a le soft skill come la curiosità, la comunicazione chiara, capacità di presentazione dei propri risultati, di problem solving e di lavoro di squadra per condividere le proprie tesi, ascoltare le idee altrui e cooperare per giungere ad una soluzione comune. Potranno anche ricoprire funzioni chiave, come ad esempio quello di elaborare report che aiutino la direzione aziendale a definire le strategie aziendali.

I dottori di ricerca potranno intraprendere un percorso accademico di ricerca ma al tempo stesso potranno lavorare e mettere a frutto il proprio know-how all'interno di aziende ICT, aziende in diversi ambiti (ad esempio logistica, manifatturiero, farmaceutico), laboratori di ricerca e sviluppo, centri di ricerca a livello nazionale o internazionale, che si occupano della progettazione, sviluppo e gestione di sistemi intelligenti complessi nonché della gestione di grandi moli di dati.

Nel campo della governance territoriale, il dottore di ricerca in Big Data e Intelligenza Artificiale sarà in grado di armonizzare tra loro fonti e archivi amministrativi diversi puntando a costruire un unico registro interconnesso tra luoghi, individui, imprese, utile per comprendere cosa accade intorno a un determinato fenomeno in termini di localizzazione, strutture, network e scambio di informazioni.

Data scientist: Il profilo dello scienziato dei dati è piuttosto tecnico e richiede di avere conoscenza di modelli matematico-statistici e algoritmi e delle librerie software necessari per implementarli e avere competenze di business intelligence, di semantica, di ontologie per la gestione delle informazioni, di metodi e tecnologie per la gestione di analisi data-driven innovative.

Il Data Scientist si occupa principalmente della raccolta dati suddivisi in: 1) dati strutturati: organizzati ed elaborati per categorie da appositi software, si tratta di dati relativi a servizi e prodotti elettronici, di informazioni presenti su internet e sui social media; 2) dati non strutturati: si tratta di materiale come recensioni, immagini, testi, audio, video e messaggi. In particolare il Data Scientist si occupa di:

- Integrazione di archivi e diverse fonti di Big Data
- Verificare l'attendibilità dei dati

- Programmazione degli algoritmi e analisi statistiche avanzate
- Traduzione di dati in risultati validi per l'azienda.

Data Engineer: È la figura centrale nella gestione della Data Pipeline, l'infrastruttura che dal luogo in cui i dati vengono raccolti li trasporta agli strumenti di front-end. Ha il compito di costruire, migliorare e progettare interi sistemi di controllo e di analisi dei dati, garantendone la corretta composizione e la qualità delle fonti dalle quali provengono. Il compito del Data Engineer è quello di fornire al Data Scientist in maniera tempestiva i dati in formati utilizzabili per le analisi. Una caratteristica fondamentale del data engineer è la capacità di riuscire a gestire e controllare adeguatamente il machine learning, elemento rilevante nell'impiego della raccolta e della condivisione dei dati. Il data engineer si occupa quindi dell'aspetto costitutivo e della composizione del processo di analisi dati, è esperto nelle warehousing solution e lavora in contatto con molte altre figure fondamentali per la gestione di tali informazioni.

3. Data Science Manager: Ha la responsabilità di aiutare l'organizzazione a sfruttare i dati per guidare i processi aziendali operando come figura di raccordo tra il team di Data Scientist e Data Engineer e il team di Management aziendale. Da una parte ha il compito di guidare le attività di produzione, analisi e gestione dei dati promuovendo il miglioramento continuo dei processi metodologici e operativi del team. Dall'altra ha il compito di fornire indicazioni affinché la direzione possa prendere decisioni informate nelle varie aree di sviluppo rilevanti per l'organizzazione sulla base di analisi e predizioni sui dati. Ne consegue che per questo ruolo, il Data Science Manager deve sia conoscere tecniche e tool di analisi, gestione e l'interpretazione dei dati sia avere solide soft skill di leadership, comunicazione e project management. Collabora con varie unità dell'organizzazione per (i) comprendere gli obiettivi dello sviluppo di progetti nuovi ed esistenti; (ii) identificare le opportunità per estrarre valore dai set di dati e/o fornire approfondimenti quantitativi e qualitativi; (iii) produrre soluzioni innovative e risolvere problemi analitici mirati, ad esempio migliorare un prodotto per rispondere o prevedere mutevoli esigenze del clienti, intervenire sul processo di produzione per migliorare le condizioni di lavoro o prevenirne i guasti.

### **Coerenza con gli obiettivi del PNRR**

Le competenze digitali, all'interno della "Strategia Nazionale per le Competenze Digitali" costituiscono un asse strategico fondamentale per la crescita sociale ed economica, a patto che si realizzino tre condizioni di evoluzione sostenibile: la popolazione acquisisce sempre più consapevolezza digitale; le pubbliche amministrazioni e le imprese sono indotte a un miglioramento dei servizi forniti; il sistema educativo si organizza per coprire le esigenze di sviluppo delle competenze digitali. All'interno del PNRR, nella Missione 1 si distinguono gli interventi per il comparto pubblico con un investimento per una task force di circa 1.000 professionisti a supporto delle amministrazioni per digitalizzare processi e un programma di re-skill con corsi on line su trasformazione digitale, transizione green, innovazione sociale, change management. Per le aziende sono previsti investimenti nella "Transizione 4.0", per sostenere le imprese che promuovono l'innovazione e la digitalizzazione attraverso crediti di imposta per beni capitali, ricerca e sviluppo e attività di formazione. Riguardo agli investimenti effettuati dalle imprese nei diversi ambiti della trasformazione digitale, rilevante risulta l'incremento degli investimenti in tecnologie, tra cui cloud, mobile, big data analytics, cyber security, IoT e software per l'acquisizione e la gestione di dati a supporto delle decisioni,

della progettazione e ingegnerizzazione dei prodotti/servizi. Per implementare gli investimenti in trasformazione digitale più innovativi, nei prossimi anni potranno essere strategiche professioni emergenti come il cloud computing specialist, big data specialist, data scientist, l'esperto in IoT, lo specialista nell'IA e il robotics specialist. Alla ricerca e formazione di queste ultime figure potrà dare un ulteriore slancio il "Programma strategico Intelligenza Artificiale 2022-2024".

In tale direzione, il dottorato "Big Data e Intelligenza Artificiale" intende colmare un deficit di competenze negli ambiti Big Data e Intelligenza Artificiale, sia sotto l'aspetto statistico che ingegneristico, per promuovere una cultura di innovazione che può essere benefica sia per il comparto pubblico che all'interno del mondo aziendale.

La Missione 4 del Piano Nazionale di Ripresa e Resilienza dedicata a Istruzione e Ricerca evidenzia un basso livello di spesa in R&S. Nel 2018, in Italia, la spesa in ricerca e sviluppo nel settore pubblico ammontava all'1,4% del PIL contro una media OCSE del 2,4% e, nel settore privato, allo 0,9% contro l'1,7% della media OCSE. Ciò comporta anche un basso numero di ricercatori e perdita di talenti. Circa la riforma e il potenziamento dei dottorati, il Piano prevede delle semplificazioni normative che possano agevolare il coinvolgimento di imprese e centri di ricerca per costruire percorsi di dottorato non finalizzati necessariamente alla carriera accademica. Negli ultimi 10 anni in Italia si è assistito a una riduzione dei dottorati pari al 40%; inoltre nel nostro Paese soltanto una persona su mille tra i 25 e i 34 anni completa ogni anno un corso di dottorato rispetto a una media UE di 1,5 e quasi il 20% di chi consegue ogni anno il titolo di dottore di ricerca si trasferisce all'estero. Diversamente, "chi rimane in Italia soffre di un profondo disallineamento tra l'alto livello di competenze avanzate che possiede e il basso contenuto professionale che trova sul lavoro" (dati Istat).

L'obiettivo del PNRR è di ridurre il divario sia numerico che anagrafico con gli altri Paesi europei, nell'aumento numerico delle borse di studio per dottorato, nella valorizzazione di competenze con dottorati che rispondano ai fabbisogni di innovazione delle imprese e promuovano l'assunzione dei ricercatori dalle imprese e nel finanziamento di cicli di dottorato destinati "all'efficientamento della gestione e dello sviluppo dell'enorme patrimonio culturale del Paese".

In tale ottica, il dottorato "Big Data e Intelligenza Artificiale" nell'ambito della promozione degli "ecosistemi dell'innovazione" si pone l'obiettivo di implementare la contaminazione e collaborazione tra Università e imprese favorendo la diffusione di modelli innovativi per la ricerca avanzata e tecnologica. Tale scelta va nella direzione in particolare della seconda componente della Missione 4, denominata "Dalla Ricerca all'impresa", in cui si descrive come gli obiettivi di collaborazione tra Università e impresa dovranno essere raggiunti grazie alla regia di MUR e MISE che, nell'ambito delle priorità del Piano Nazionale della Ricerca 2021-2027 e di Horizon Europe, dovranno favorire la mobilità tra Università, infrastrutture di ricerca e aziende, con la consapevolezza di dover migliorare la sinergia tra pubblico e privato nell'interesse strategico del Paese, attraverso la creazione di partnership di rilievo nazionale o con una vocazione territoriale e una valorizzazione della professionalità dei ricercatori con alte specializzazioni in discipline STEM.

## Curriculum dottorali afferenti al Corso di dottorato

### BIG DATA MANAGEMENT PER LA TRANSIZIONE DIGITALE

Il termine Data Management è estremamente ampio e comprende non solo aspetti tecnologici, ma anche aspetti organizzativi, di processo e di compliance. Il Data Management consente di valorizzare un asset strategico, il capitale dei dati, ponendolo a servizio degli obiettivi di sviluppo dell'organizzazione.

Per cogliere nuove opportunità di business in ambiti come Industria 5.0, Economia e Marketing, Smart City e Smart Health servono infrastrutture, tool e competenze di Data Science.

Le aziende data-driven sono quelle che considerano la gestione dei dati (data management) non come un fattore tecnico, ma come un pilastro strategico del business. Essere data-driven significa farsi guidare dai numeri, avere un approccio basato sui dati, per prendere decisioni informate, basate su fatti oggettivi e non su sensazioni personali. La trasformazione in data-driven company non può dunque avvenire con la sola tecnologia, ma con un percorso di change management in grado di portare la cultura del dato a tutti i livelli aziendali.

Non solo ma rappresentare la complessità dei fenomeni è anche la nuova sfida per la governance di un territorio. Ogni un evento può essere misurato rispetto a ciò che c'è intorno all'evento stesso, legando fonti eterogenee, diverse una dall'altra. Non si guarda a ciò che avviene solo in un determinato spazio, con una precisa localizzazione, ma il concetto di luogo diventa elemento di mediazione tra vari livelli informativi e ciò permette di governare i fenomeni a livello di spazio urbano.

Per supportare qualsiasi processo di data driven è necessario:

- sviluppare una forte Data Literacy diffusa in azienda, intesa come capacità di analizzare e interpretare i dati al fine di ricavare informazioni utilizzabili per prendere decisioni;
- formare un core team, formato da data scientist e data engineer, con competenze trasversali in ambito analytics, business e IT, che costituirà di fatto il principale centro di competenza sull'analisi dei dati in azienda.

### INTELLIGENZA ARTIFICIALE PER INDUSTRIA 5.0 ED ECONOMIA CIRCOLARE

Il curriculum vuole cogliere le opportunità legate allo sviluppo dell'Intelligenza Artificiale nei contesti industriali al fine di perseguire nuovo concetto di società in cui non solo Intelligenza Artificiale, ma anche tecnologie, quali IoT, robot, cybersicurezza, cloud e reti ad alta velocità, sono utilizzate attivamente non solo per perseguire un vantaggio economico, ma soprattutto per il miglioramento delle condizioni di vita dell'uomo e dell'ambiente.

Il curriculum mira a formare esperti che contribuiscano all'avanzamento dei saperi scientifici attraverso l'utilizzo integrato di metodologie di ricerca tecnologica, da una parte, e statistica ed economico-sociale dall'altra. Lo studente viene introdotto in un percorso formativo che permetta di acquisire una conoscenza integrata e una visione "complessa" dell'ecosistema delle applicazioni, tecnologie e soluzioni metodologiche, in grado di affrontare i problemi con un approccio sistemico e multi-disciplinare. In particolare, intende formare esperti in Data Science e Intelligenza Artificiale e loro applicazioni pratiche nei settori della Industria 5.0 e dell'economia avendo in mente i principi di Economia Circolare come modello economico rigenerativo.

La formazione combinata nella scienza dei dati e dei modelli di apprendimento automatico, da una parte, e delle tecnologie avanzate, dall'altra, sarà centrale per permettere di ideare e realizzare tale modello tramite (i) raccolta e distribuzione di informazioni dal processo produttivo e dalla filiera, (ii) metodi e tecniche per la gestione ed analisi dei dati e modelli di apprendimento automatico (iii) formazione di adeguati circuiti di feedback tra la trasformazione industriale e processi decisionali lungo la filiera, incluse forme di incentivo. In questo ambito, verranno anche considerati elementi di vulnerabilità dei dati e dei sistemi di elaborazione e comunicazione per prevenire e fronteggiare attacchi cyber e preservare la privacy.