



Unione europea  
Fondo sociale europeo

Regione Emilia-Romagna



# DATA LAB

## GUARDA AVANTI

Big Data, nuove competenze  
per nuove professioni.

(Progetto finanziato rivolto a laureati  
in tutte le aree disciplinari)





PAG 3 <b>IL PROGETTO</b>	PAG 5 <b>ANTICIPARE LA CRESCITA CON LE NUOVE COMPETENZE SUI BIG DATA</b>	PAG 8 <b>INTRO TO BIG DATA</b>	PAG 10 <b>BASI DI STATISTICA, MATEMATICA, GESTIONE DB</b>
PAG 12 <b>MARKETING &amp; SOCIAL MEDIA ANALYTICS</b>	PAG 14 <b>DATA JOURNALISM</b>	PAG 16 <b>MACHINE LEARNING &amp; VIDEO ANALYTICS</b>	PAG 18 <b>DATA PRIVACY, SECURITY AND ETHICS</b>
PAG 20 <b>4.0 OPERATIONS (IOT &amp; BIGDATA ANALYTICS)</b>	PAG 22 <b>TECNOLOGIE E SOFTWARE PER IL TRATTAMENTO BIG DATA</b>	PAG 25 <b>DEEP LEARNING - ARTIFICIAL INTELLIGENCE NATURAL LINGUISTIC PROCESSING</b>	
PAG 28 <b>BIG DATA &amp; STARTUP</b>		PAG 30 <b>PROJECT WORK SETTORIALE</b>	

## IL PROGETTO

L'innovazione tecnologica e la digitalizzazione dell'informazione sono sempre più pervasive nella vita delle persone, sia nella sfera privata, sia in quella sociale che lavorativa. Siamo ormai abituati a dispositivi più o meno intelligenti (e automatizzati) che memorizzano tutti i nostri dati, si sostituiscono alle attività lavorative umane e suggeriscono o prendono decisioni al posto nostro.

L'insieme di questi dati, che ha preso l'espressione di "Big Data", indica un insieme di informazioni talmente ampio da non essere archiviabile attraverso i comuni sistemi hardware e tipicamente generati da processi digitali eterogenei, con formati strutturati (database), ma molto più di frequente non strutturati (video, immagini, post sui social). Stiamo assistendo a una rivoluzione che non è solo tecnologica, ma anche antropologica. I cambiamenti che sono in atto nella società moderna non sono riconducibili solo alle "macchine", ma hanno a che fare con la persona e con il suo rapporto con le tecnologie.

La vera natura dei Big Data deriva, infatti, dall'esperienza quotidiana delle persone e dal funzionamento automatico dei sistemi digitali. Vivendo, lavorando, utilizzando "device", più o meno consapevolmente, emettiamo dati così come i dispositivi digitali tra loro interconnessi che costituiscono il framework nel quale ci muoviamo e ci appoggiamo quotidianamente, pensiamo all'uso quotidiano di motori di ricerca, social media, acquisti online, fruizione di servizi anche culturali online.

Tutto questo produce Big Data, un flusso continuo di segnali digitali provenienti dalle attività umane, private, sociali ed economiche. Una vastità di dati che apre spazi di osservazione e conoscenza nuovi, modificando la struttura dei processi decisionali individuali e collettivi, modificando l'economia e la società, rendendole più dinamiche e complesse.

Partendo da questa premessa, l'obiettivo del progetto è lo sviluppo di conoscenze e competenze che permettano alle persone con alto livello di istruzione in ogni area disciplinare di divenire soggetti capaci di dare senso, di creare nessi logici, elaborare giudizi di sintesi e trasformare i dati in informazioni ad alto valore aggiunto e diffonderlo all'interno delle organizzazioni in cui lavoreranno e vivranno. Il progetto si articola in 11 percorsi formativi modulari e gratuiti di seguito descritti.

## DESTINATARI

Il progetto, completamente finanziato dalla Regione Emilia Romagna e dal Fondo Sociale Europeo, si rivolge a laureati (da non più di 24 mesi) in ogni area disciplinare (umanistica, economico sociale, scientifica), residenti o domiciliati in Regione Emilia Romagna da data antecedente l'iscrizione ai percorsi formativi.

Ogni candidato potrà creare il proprio percorso formativo individuale, fino ad un massimo di 500 ore, scegliendo tra gli 11 progetti, i più idonei al suo profilo, e con il supporto del personale didattico.

## CRITERI DI SELEZIONE

L'iter di selezione prevede la verifica dei requisiti formali attraverso l'analisi delle schede di iscrizione e controlli a campione ed un colloquio di selezione orientativa durante il quale sarà creato con ogni partecipante un piano formativo individuale all'interno dell'Operazione (con possibilità di iscriversi a più progetti fino ad un massimo di 500 ore di formazione). Il piano formativo individuale sarà oggetto di verifica, ed eventuale ridefinizione, anche in itinere, per valutare i percorsi più idonei a ogni allievo nonché eventuali servizi per il lavoro attivabili al termine dell'Operazione.

## MODALITÀ DI PARTECIPAZIONE

Per iscriversi è necessario presentare domanda tramite format online dal sito [www.bigdata-lab.it](http://www.bigdata-lab.it).

Sono previste per ogni corso diverse edizioni sull'intero territorio regionale a partire da Novembre 2018 fino a Dicembre 2019. È possibile visitare il sito [www.bigdata-lab.it](http://www.bigdata-lab.it) o contattare i soggetti gestori del progetto.

## ENTI GESTORI

Il progetto è gestito in partnership da una compagine di enti di formazione: Formindustria, ASSOform Romagna, CIS, CISITA Parma, FAV, Forpin, Ifoa, Il Sestante Romagna, Nuova Didactica.

## UNIVERSITÀ PARTNERS

Hanno aderito le università: UniMoRe, Alma Mater Studiorum di Bologna, Università degli Studi di Ferrara, Università di Parma, che collaborano nella progettazione e docenza di alcuni percorsi formativi e partecipano al comitato tecnico scientifico.

## SOGGETTI PROMOTORI

Tetra Pak Packaging Solutions spa, Dedagroup - Public Services srl, Digital Transformation Institute, Crif spa, Var Group spa, Energy Way srl, The Hub Reggio Emilia, Tabulaex srl, Binella 175 srls, Smarten srl, Fondazione per le scienze religiose Giovanni XXIII, Cineca, Confindustria Emilia Romagna.

## CONTATTI













Per informazioni:

Tel. 800 036425 (Numero verde)

Mail. [bigdata@formindustria.it](mailto:bigdata@formindustria.it)

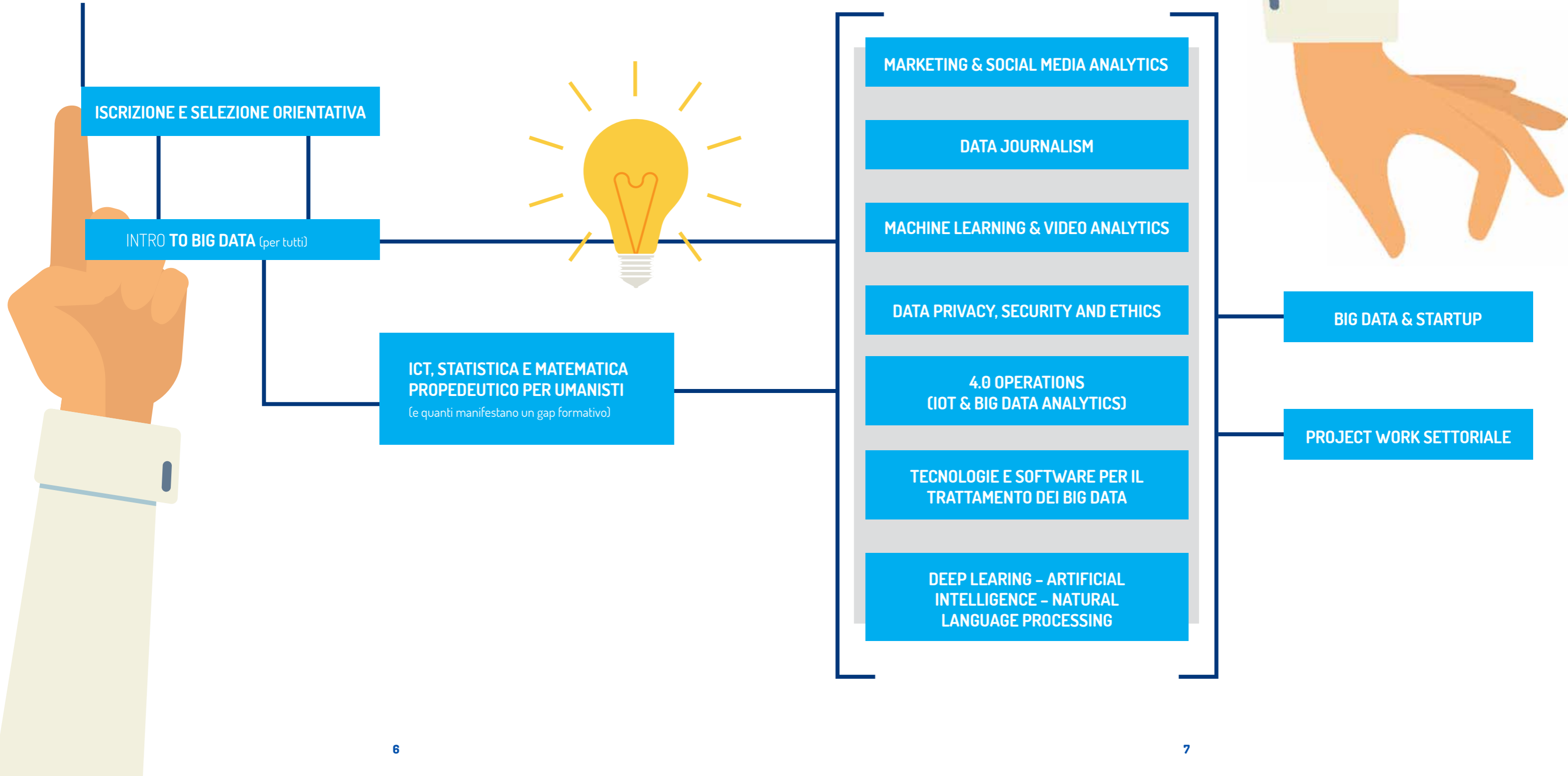
Sito. [www.bigdata-lab.it](http://www.bigdata-lab.it)

# ANTICIPARE LA CRESCITA CON LE NUOVE COMPETENZE SUI BIG DATA

	Indirizzo umanistico	Indirizzo economico giuridico	Indirizzo scientifico
1. Intro to Big Data			
2. Basi di statistica, matematica, gestione db			
3. Marketing & Social media analytics			
4. Data journalism			
5. Machine learning & video analytics			
6. Data privacy, security and ethics			
7. 4.0 Operations (Iot & Big Data analytics)			
8. Tecnologie e software per il trattamento dei Big Data			
9. Deep Learning - Artificial Intelligence - Natural language processing			
10. Big Data & Startup			
11. Project Work settoriale			

**Nota:** nella tabella sono indicati i percorsi consigliati, ma ogni partecipante potrà valutare soluzioni alternative, derivanti da conoscenze pregresse, da validare in sede di selezione orientativa.

# PERCORSI FORMATIVI



# INTRO TO BIG DATA

## DESCRIZIONE DEL PERCORSO

Il progetto INTRO TO BIG DATA è un percorso introduttivo e di allineamento previsto per tutti i destinatari, a prescindere dall'area formativa di provenienza. Il progetto introduce quindi al mondo dei BIG DATA, ovvero della possibilità, per effetto della digitalizzazione di processi e prodotti derivata dalle nuove applicazioni tecnologiche dell'industria 4.0 e del web, di disporre di enormi quantità di dati e, soprattutto, poter usare algoritmi e tecnologie in grado di estrarre ed utilizzare questi dati in tempi brevissimi e con costi ridotti rispetto al passato.



## OBIETTIVI

L'obiettivo del progetto è di creare una piena consapevolezza delle **potenzialità dei big data** e dei campi di applicazione, per poter costruire di conseguenza percorsi di specializzazione rispondenti alla domanda del sistema produttivo e coerenti con il proprio background formativo e universitario dei partecipanti.

## DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività. Le competenze necessarie in ingresso, oltre al requisito formale della laurea, sono trasversali: una forte attitudine alle tecnologie digitali, predisposizione al pensiero critico e capacità di autoaggiornamento.



## CONTENUTI DEL PERCORSO

- Introduzione e scenario di tipo Data Driven;
- le 4 tipologie di Data Analysis: descrittiva, predittiva, prescrittiva, automatizzata;
- Le professioni legate ai Big Data: ruoli e mansioni
- Le soft skills per l'Industria 4.0 ed il lavoro in team multidisciplinari
- Introduzione ai Big Data (concetti base e caratteristiche distintive)
- I principi base dei BIG DATA ed evoluzioni
- Introduzione alle tecnologie
- Introduzione all'intelligenza artificiale e al machine learning: famiglie di algoritmi e casi d'uso
- Elenco sorgenti big data: social, iot, log, geospatial, voce, images, web, ...
- Testimonianze aziendali con riferimento ai diversi ambiti di applicazione (energia, sanità, cultura, mobilità, finanza, servizi e innovazione sociale)



## COMPETENZE

La rivoluzione dei Big Data consente di creare per molteplici settori nuovi servizi e nuove opportunità di business: dall'automotive alla sanità, dal mondo educational alla telefonia, al settore bancario, ecc. L'applicazione delle potenzialità legate ai BIG DATA è pressoché sconfinata ed offre enormi opportunità di riduzione degli sprechi, miglioramento delle performance, nuovi servizi a valore aggiunto. Inoltre, è trasversale a tutti i processi aziendali perché le ricadute ed il valore sono di volta in volta strategici per la Ricerca & Sviluppo, il Marketing e le Vendite, il processo produttivo, l'amministrazione e finanza con il coinvolgimento della funzione IT a supporto. L'impatto delle nuove tecnologie riguarda tutti i settori e incide sulle stesse modalità di fruizione di servizi anche in ambiti che possono apparire distanti, come quello culturale e delle scienze sociali ed offre pertanto ricadute e conoscenze utili anche per i laureati di area umanistica.

## DESCRIZIONE DURATA

Il corso della durata di 80 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

## PRIMA CHIUSURA ISCRIZIONI: 30/11/2018

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna

# BASI DI STATISTICA, MATEMATICA, GESTIONE DB

## DESCRIZIONE DEL PERCORSO

“SAPERE” SIGNIFICA AVERE INFORMAZIONI UTILI PER PRENDERE DECISIONI.

Questa affermazione sembra apparentemente semplice, ma è molto insidiosa, perché i dati sono una cosa diversa dalle informazioni. I dati sono sequenze di lettere e numeri a cui attribuiamo un significato.

Questo significa che, prima di approcciare i big data disponibili, dobbiamo ripulire i database disponibili di tutti i dati senza significato, perché non identificativi di un processo. Il primo passo da fare, prima di iniziare una qualunque analisi, quindi, è il data cleaning, senza il quale saremo destinati a scoprire solo alla fine che ogni nostra valutazione può essersi basata anche su dati inaffidabili.

Una volta certi di avere dei dati solidi, è possibile cominciare a domandarsi: come trasformare dei dati in informazioni utili?

Per farlo è necessario, perciò, porsi queste domande:

- Cosa vogliamo sapere?
- Perché lo vogliamo sapere
- Cosa deve essere misurato?
- Come viene misurato?

Fare chiarezza su queste domande, significa individuare gli strumenti che permettono di estrarre le informazioni dai dati. Questi strumenti appartengono tutti ad una unica scienza: la statistica applicata nella gestione dei database.

## OBIETTIVI

L'obiettivo del progetto è allineare le conoscenze e le competenze dei laureati in discipline umanistiche ed economico giuridiche a quelle dei laureati in discipline scientifiche, per poter frequentare efficacemente, di conseguenza, percorsi di specializzazione rispondenti alla domanda del sistema produttivo.

Per i destinatari provenienti dall'area umanistica e delle scienze sociali si tratta di una opportunità fondamentale per acquisire competenze matematiche, statistiche e di utilizzo di data base applicati al mondo dei Big data, nonché degli ambiti di sviluppo professionali legati alle scienze umane. Per i destinatari provenienti dall'area giuridico-economica il progetto offrirà una prospettiva evolutiva rispetto a competenze, già acquisite attraverso il percorso universitario, ma ottimizzabili per ricondurle a professionalità innovative del paradigma data driven.

Per i destinatari provenienti dall'area scientifica, il progetto sarà l'occasione di ripercorrere tematiche e competenze note, ma in ottica di aggiornamento e acquisizione di nuovi approcci metodologici.

## DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività. Le competenze necessarie in ingresso, oltre al requisito formale della laurea, sono trasversali: una forte attitudine alle tecnologie digitali, predisposizione al pensiero critico e capacità di autoaggiornamento.

A questo progetto possono accedere laureati provenienti dalle tre aree disciplinari, umanistica, giuridico-economica e scientifica. Di fatto il progetto rappresenterà il primo tassello del percorso formativo individuale di ognuno, per allineare le conoscenze del contesto, in ottica interdisciplinare e si rivolge in particolare a quanti hanno gap sulla statistica e l'IT.

## CONTENUTI DEL PERCORSO

- Basi di ricerca operativa, complessità, alberi decisionali, regressione lineare e logistica
- Statistica Descrittiva
- Statistica analitica
- La presa di decisioni sulla base delle informazioni raccolte: variabili dipendenti (cause) e variabili indipendenti (effetti)
- Basi di programmazione a oggetti (classi, interfacce, ...)
- Cenni di Programmazione Java e Python
- Introduzione ai database relazionali, SQL
- Introduzione ai db NoSQL



## COMPETENZE

Al termine del percorso i partecipanti acquisiranno le conoscenze e competenze utili per destreggiarsi nel mondo dell'analisi statistica, ovvero far dialogare le informazioni individuando le correlazioni esistenti fra di loro al fine di prendere delle decisioni. Gli allievi sapranno, quindi che prima di qualunque analisi, in qualsiasi contesto professionale, il Data Cleaning è strategico per non incorrere in dati inaffidabili e quindi completamente inutili per una valutazione; alla fine del percorso i partecipanti dovranno altresì essere in grado di riconoscere e selezionare il linguaggio o il database più corretto per ognuno dei casi di possibile utilizzo dei Big Data, per poter comunicare in modo efficace all'interno di un team.

## DESCRIZIONE DURATA

Il corso della durata di 80 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

## PRIMA CHIUSURA ISCRIZIONI: 30/11/2018

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna

# MARKETING & SOCIAL MEDIA ANALYTICS

## DESCRIZIONE DEL PERCORSO

Sembra che le aziende del nostro paese abbiano ormai compreso come l'analisi dei Big Data rappresenti una fonte di vantaggio competitivo e un fondamentale strumento di evoluzione dei loro modelli di business. Lo testimonia una recente ricerca di Microsoft-Ipsos Mori da cui emerge che il 66% delle PMI italiane possiede competenze e strumenti per gestire le informazioni di business e il 50% prevede di investire in data analytics.

Tuttavia, spesso sembra esserci una impasse su come mettere a frutto le informazioni preziosissime fornite dai Big Data, una volta raccolte e analizzate. Le imprese italiane hanno quindi compreso l'importanza di estrarre insight dai dati, ma sono ancora lontane da strategie di business "data driven".

La specializzazione del Progetto contribuisce a sviluppare (e diffondere) un approccio strategico Data Driven, ampliando l'orizzonte del valore e di senso del mercato, modificando il modello di business (e la mentalità) attraverso strategie di marketing "ad personam" (customer centered business).

Applicazioni virtuose di data driven marketing strategy sono ad esempio quelle che nel retail permettono – a partire dall'analisi di vendite e altri dati in store – di ricavare actionable insights, che vengono poi utilizzate per modificare la strategia quasi in tempo reale, oppure quelle che permettono l'utilizzo di software tools finalizzati al monitoraggio dei prezzi dei beni di consumo venduti online, con la conseguente possibilità di reindirizzare la pricing strategy. Adottare strategie data driven significa quindi cambiare il modo in cui le aziende si organizzano internamente: incoraggiare le imprese a mantenere un ciclo continuo di interazione con i propri clienti, continuare ad imparare ed evolversi seguendo le esigenze degli utenti.



## OBIETTIVI

L'obiettivo del progetto è quello di introdurre i benefici dell'utilizzo dei Big Data in contesto marketing, analisi di mercato e lead generation.

Alla fine del corso, lo studente sarà in grado di comprendere e utilizzare i principali canali di marketing digitale in modo da poter pianificare obiettivi, modelli e kpi. Inoltre lo studente sarà in grado di collegarsi alle varie sorgenti provenienti da social media e digital analytics in modo da sfruttarne i dati per il monitoraggio delle performance e le informazioni nascoste all'interno.

## DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività. Le competenze necessarie in ingresso, oltre al requisito formale della laurea, sono trasversali: una forte attitudine alle tecnologie digitali, predisposizione al pensiero critico e capacità di autoaggiornamento.

A questo progetto possono accedere prioritariamente laureati provenienti dalle aree disciplinari umanistica e giuridico-economica permettendo di mantenere una continuità formativa rispetto al percorso di studi e di specializzare le competenze innovative legate all'applicazione dei big data nel marketing.

I potenziali destinatari provenienti dall'area scientifica non sono esclusi. Verrà valutata nella fase di selezione orientativa una possibile traiettoria e attitudine in questo ambito su richiesta degli interessati.



## CONTENUTI DEL PERCORSO

- Marketing 4.0
- Data Driven Strategy
- Modelli: rappresentazione della realtà per decidere runtime
- Metriche: individuazione delle metriche e degli obiettivi
- Social media Analytics e Grafi
- Ciclo: Discovery – Listening – Planning – Execution – Measurement
- Piattaforme di marketing e customer engagement
- Big Data e Social Media Analysis: delimitazione propedeutica del campo d'indagine
- Fare analisi dei Social: elementi metodologici di Social Media Listening e Social Media Analysis
- Data Visualization e Social Media Intelligence: come estrarre utili insight dai Big Data della Internet of People
- Principali applicazioni di Big Data Analytics per i dati originati da Social Analytics e web
- Marketing automation ed esempi e esercizi di machine learning applicato



## COMPETENZE

La figura formata dal progetto acquisirà metodologie di gestione agile e snella, tecniche di innovazione utilizzate da aziende tecnologiche a crescita elevata replicabili in qualsiasi organizzazione, approcci pragmatici per ottimizzare il marketing in un ambiente frammentato e in continua evoluzione.

In uno scenario "liquido", in rapido e continuo mutamento le aziende esprimono il fabbisogno di figure reattive e curiose, in grado di progettare e gestire campagne di marketing "adattive", capaci di rispondere a scenari in costante evoluzione.

Il Progetto ha lo scopo di fornire agli studenti gli strumenti e le conoscenze necessarie ad analizzare dati su larga scala provenienti da Online Social Networks. All'interno del percorso vengono presentati i software di analisi necessari, vengono poi illustrati i risultati ottenuti dall'applicazione ai vari social network.

Nel corso delle ore di laboratorio, gli studenti applicano i concetti appresi in aula su dataset reali di Online Social Networks. Gli studenti apprendono così l'utilizzo di strumenti software di riferimento per l'analisi di dati provenienti da Online Social Networks.

## DESCRIZIONE DURATA

Il corso della durata di 80 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

## PRIMA CHIUSURA ISCRIZIONI: 31/12/2018

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna

# DATA JOURNALISM

## DESCRIZIONE DEL PERCORSO

È possibile definire il giornalismo dei dati come una forma di investigazione che mira a sviluppare storie a partire dai dati, una forma speciale di interpretazione di materiale di ricerca, basata su pratiche statistiche e una forma specifica di presentazione dei dati che vuole rendere graficamente visibili i messaggi, spesso con l'ausilio di un'applicazione Web. Inoltre, alle volte la pubblicazione dei dati grezzi e delle loro fonti in termini di open data è riconosciuta come una componente fondamentale del data journalism.

Il profilo professionale del DATA JOURNALIST ha differenze significative rispetto al giornalismo tradizionale, compresa l'importanza attribuita alla visualizzazione, il tasso più basso di selezione giornalistica e una maggiore trasparenza nell'esposizione dei risultati delle indagini. Nel caso delle applicazioni interattive, inoltre, la selezione e l'interpretazione dei dati viene spesso lasciata ai lettori che, da soli, possono interrogare l'applicazione e sviluppare un proprio mindset sui fatti, cliccando, ad esempio, sulle mappe. In questo senso, il data journalist è meno gatekeeper di un giornalista tradizionalmente inteso.

Un altro punto di distanza è dato dalla pubblicazione dei risultati di ricerca, i dataset originali, aspetto che caratterizza esclusivamente il giornalismo fatto con i dati. Il background dei data journalist, inoltre, necessita anche di un forte senso dei numeri e di una certa competenza con i metodi delle scienze sociali e della statistica.

(European Journalist Observatory 2015)



## OBIETTIVI

Il Progetto ha l'obiettivo di fornire competenze teoriche e capacità pratiche nell'ambito della realizzazione di prodotti giornalistici data-driven, ovvero di articoli, inchieste, comunicati stampa incentrati prioritariamente sull'analisi qualitativa di dati e informazioni statistiche, da realizzare attraverso un processo di valorizzazione delle fonti informative disponibili (offline e online, a livello nazionale ed internazionale).

## DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività. Le competenze necessarie in ingresso, oltre al requisito formale della laurea, sono trasversali: una forte attitudine alle tecnologie digitali, predisposizione al pensiero critico e capacità di autoaggiornamento.

A questo progetto possono accedere laureati provenienti dall'area umanistica e delle scienze sociali permettendo di mantenere una continuità formativa rispetto al percorso di studi e di specializzare le competenze innovative legate all'applicazione dei big data alla argomentazione giornalistica.

I potenziali destinatari provenienti dalle altre aree non sono esclusi. Verrà valutata nella fase di selezione orientativa una possibile traiettoria e attitudine per questa specializzazione su richiesta degli interessati.

## CONTENUTI DEL PERCORSO

- Il Workflow del Data Journalism: ricerca dei dati, verifica e organizzazione, analisi, incrocio con altri dati, visualizzazioni e Storytelling. Gli step indispensabili nell'attività del Data Journalist
- Data Driven Journalism: formare, coerentemente con le esigenze del lavoro del giornalista, la capacità di navigare negli Open & Big Data per produrre articoli e reportage
- Data Visualization: Come presentare nel modo più efficace per la comprensione del lettore i dati raccolti
- Infografica: Come costruire degli schemi grafici basati sui dati per facilitare (o sostituire) l'articolo testuale nel veicolare i contenuti soprattutto sui nuovi canali di fruizione digitali
- Database Journalism, formare sulle tecniche disponibili per costruire un sistema di gestione delle informazioni in cui le informazioni sono organizzate in un database (al contrario di una struttura organizzativa tradizionale incentrata sulla storia) e che quindi facilitano successive ricerche e approfondimenti



## COMPETENZE

Il principale risultato atteso è quello di formare gli operatori del mondo dell'informazione su come agire nel nuovo contesto Digitale: sia nel reperimento di informazioni su cui basare il proprio lavoro, sia nel come poi veicolare i contenuti e renderli accattivanti per i nuovi mezzi di fruizione.

Il data journalism specialist è una figura professionale che potrà trovare sbocchi lavorativi in redazioni di testate, web e tradizionali, in agenzie di comunicazione, in uffici stampa. Autonomamente il data journalism specialist potrà proporsi sia nel ruolo di redattore free lance, sia di consulente di comunicazione, contribuendo in particolare a valorizzare i contenuti di blog e di siti di informazione tematici e corporate, a qualificare l'informazione e a favorire l'engagement attraverso i social network.

## DESCRIZIONE DURATA

Il corso della durata di 80 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

## PRIMA CHIUSURA ISCRIZIONI: 31/01/2019

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna



# MACHINE LEARNING & VIDEO ANALYTICS

## DESCRIZIONE DEL PERCORSO

Il progetto MACHINE LEARNING & VIDEO ANALYTICS è relativo ai macrotemi dell'A.I.- Artificial Intelligence, con specifico riferimento ad aspetti inerenti l'Intelligent Data Processing.

Il Machine Learning, in italiano Apprendimento Automatico, indica gli algoritmi di programmazione delle macchine attraverso cui si insegna ai computer e ai robot a fare azioni ed attività in modo naturale, come gli esseri umani o gli animali, ovvero imparando dall'esperienza. Usando metodi matematico-computazionali gli algoritmi di ML insegnano ad apprendere informazioni direttamente dai dati, senza modelli matematici ed equazioni predeterminate e migliorano le loro prestazioni in modo "adattivo" mano a mano che gli "esempi" da cui apprendere aumentano.

Il Video Analytics aiuta le organizzazioni a sviluppare funzionalità di protezione, intelligence e investigative complete utilizzando i video.

È possibile utilizzare ricerche avanzate, oscuramento e analytics di riconoscimento facciale per individuare immagini rilevanti e informazioni critiche tra diversi file video provenienti da diversi tipi di videocamera. Sono supportati sia da una selezione di videocamere che trasmettono in diretta che dall'acquisizione di video preregistrati da videocamere fisse e in movimento.

Attraverso software di analisi video è possibile estrarre informazioni dai video acquisiti per elaborare insight e schemi.



## OBIETTIVI

L'obiettivo del corso è quello di introdurre gli studenti alle tematiche relative al Machine Learning ed al Video Analytics attraverso inquadramenti teorici, casi applicativi, utilizzo di metodologie e tecnologie.

Le competenze acquisite al termine del corso seguono un semplice filo logico: lo studente sarà in grado di individuare da un dataset generico come poterne trarre valore tramite l'applicazione di algoritmi di Machine Learning, per poi procedere al deploy dell'algoritmo e alla relativa visualizzazione dei risultati.

## DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività. Le competenze necessarie in ingresso, oltre al requisito formale della laurea, sono trasversali: una forte attitudine alle tecnologie digitali, predisposizione al pensiero critico e capacità di autoaggiornamento.

A questo progetto possono accedere prevalentemente laureati provenienti dall'area scientifica.

## CONTENUTI DEL PERCORSO

- Introduzione al machine learning: introduzione e casi d'uso
- Teoria degli algoritmi di Machine Learning (con apprendimento supervisionato, non supervisionato, con rinforzo, semi-supervisionato)
- Framework open source come Scikit Learn e relativi esempi di Dataset: caricamento dati e training, parametri del modello, ...
- Processo di implementazione algoritmi di ML: definizione del problema, raccolta dei dati, data cleaning, costruzione del modello, cross validation, valutazione dei risultati, visualizzazione dei risultati
- Dai Big Data alla visualizzazione
- Concetti introduttivi alla Video Analytics, machine learning come riduttore della complessità e dimensione di un dataset
- Best Practices di Video Analytics: Scelta software di cattura ed elaborazione video
- Intelligent Video Analytics: Trasformare i video in insight. Trovare immagini rilevanti tra svariati tipi di file video



## COMPETENZE

I contesti lavorativi in cui ad oggi spendere le conoscenze e competenze ricavabili con questo corso sono dunque legati agli ambiti applicativi succitati nell'introduzione al progetto sia per il Machine Learning che nello step di visualizzazione dei dati in esito: quindi innanzitutto quelli della Ricerca e Sviluppo ma anche del Marketing e Vendite e Delivery di Servizi evoluti e smart sia per il mondo B2B che B2C.

La spendibilità delle conoscenze acquisite è dimostrata da numerosi studi previsionali sulle professioni del futuro; molte organizzazioni investono nell'acquisizione di masse di dati che permettono migliore conoscenza delle abitudini di consumo e acquisto della clientela e l'uso di questo patrimonio richiede conoscenze specialistiche che non sono più solo informatiche.

## DESCRIZIONE DURATA

Il corso della durata di 80 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

## CHIUSURA ISCRIZIONI Entro il 28/02/2019

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna

# DATA PRIVACY, SECURITY AND ETHICS

## DESCRIZIONE DEL PERCORSO

Con il ruolo crescente del mondo Digitale nella vita delle persone (accesso a servizi privati e pubblici, condivisione delle proprie opinioni/ gusti/abitudini sui canali social,...) sta diventando sempre più urgente tutelare la privacy di chi vi accede, permettendo allo stesso tempo ai fornitori di servizi e ai decisori pubblici di avere accesso a quelle informazioni che sono necessarie per operare, magari aggiungendo valore ed efficientando processi, prodotti e servizi pubblici e privati nei più svariati ambiti.

La rivoluzione BIG DATA non investe solo la funzione IT, ma mette in atto una rivoluzione di impatto antropologico, modificando non solo tecnicamente tutte le attività umane, ma penetrando a livello più profondo nelle modalità di relazione tra l'uomo e i suoi artefatti (si parla di homo+machina sapiens). Da qui la necessità di analizzare il fenomeno e comprenderlo in modo interdisciplinare per valutarne e governare tutti i risvolti cogliendo opportunità e mitigando, in chiave preventiva, le minacce. Sono sotto gli occhi di tutti i rischi legati all'esposizione mediatica ed alla circolazione dei nostri dati sensibili (dai dati anagrafici a quelli sanitari o bancari), in ultimo l'attuale caso "Cambridge Analytica" che ha investito addirittura Facebook e fatto emergere una sottovalutazione del tema. Si gioca una grossa partita sia in termini di rischi di truffe o reati, che di limitazione della libertà personale e pervasività delle aziende che accentrano e gestiscono i nostri dati nella sfera privata, con tentazioni o derive manipolatorie che attentano ai processi democratici. Il tema della sicurezza informatica e del rispetto della privacy sono, quindi, altrettanto centrali e imprescindibili per chi si occupa di trattamento e valorizzazione dei dati a tutti i livelli, dal data entry al decision making.



## OBIETTIVI

Il progetto offre un approfondimento verticale sui temi della protezione e privacy (da un punto di vista sia legale che tecnologico) e sui risvolti etici derivanti dalle trasformazioni digitali. Ci si pone l'obiettivo di formare un profilo professionale attuale, in grado di inserirsi nella nuova traiettoria di sviluppo professionale accelerata dalla nuova normativa europea in materia di privacy, che prevede una figura professionale specifica: il Data Protection Officer.

## DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività. Le competenze necessarie in ingresso, oltre al requisito formale della laurea, sono trasversali: una forte attitudine alle tecnologie digitali, predisposizione al pensiero critico e capacità di autoaggiornamento. A questo progetto possono accedere laureati provenienti dalle tre aree disciplinari, umanistica, giuridico-economica e scientifica. Per chi proviene dalle prime due aree è necessario aver partecipato ai progetti 1 e 2 per l'acquisizione dei contenuti di contesto e delle conoscenze di base relative ai DB.

## CONTENUTI DEL PERCORSO

### CYBERSECURITY

Politiche di sicurezza informatica: risorse da proteggere, valutazione costo e rischio; La qualità e affidabilità dei sistemi informativi; Software antivirus; sicurezza di archivi e DB: backup e recovery; Come evitare lo Spam per l'invio di e-mail; Sicurezza sul web: CMS e plugin di terze parti come rischi per la sicurezza dei siti; Crittografia e gestione dei certificati; Protezione da attacchi per le applicazioni web: tipologie e motivazione degli attacchi; Digital forensics.

### NORMATIVA IN MATERIA DI PRIVACY

Ruoli e responsabilità nel trattamento del dato personale; Studio del ciclo di vita del dato Nuovo regolamento europeo 2016/679 sulla privacy: nuovo approccio tra accountability e integrazione con i sistemi di gestione; I sistemi di gestione ed il risk based thinking; Il D.Lgs. 231/01 e l'adozione del modello organizzativo Discussione su casi di studio/best practices

### ETICA DEI BIG DATA:

ToS: terms of services; Strumenti legali e tecnologici per difendere la privacy: Ubiquitous Commons; L'uomo e la macchina: Artificial moral agent



## COMPETENZE

Al termine del progetto i partecipanti saranno in grado di:

- comprendere le normative sulla data privacy (in particolare GDPR) e le implicazioni sull'operatività delle organizzazioni pubbliche e private che gestiscono dati;
- comprendere le esigenze informative delle aziende e quali precauzioni vadano messe in atto per rimanere compliant con le normative;
- collaborare all'implementazione di strumenti per assicurare la compliance, in termini procedurali e tecnologici;
- riconoscere rischi etico-legali nel trattamento dei dati e contribuire ad una corretta governance.

## DESCRIZIONE DURATA

Il corso della durata di 80 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

## PRIMA CHIUSURA ISCRIZIONI: 31/12/2018

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna

## 4.0 OPERATIONS (IOT&BIGDATA ANALYTICS)

### DESCRIZIONE DEL PERCORSO

Il progetto 4.0 OPERATIONS (IOT&BIGDATA ANALYTICS) è un percorso specialistico rivolto prevalentemente a laureati di area scientifica.

Nel mondo industriale la diffusione dell'IIoT e dell'Industrial IIoT permette di riorganizzare e rivedere radicalmente la produzione in forma integrata con la progettazione, l'organizzazione del lavoro, il controllo di prodotto, il marketing e le vendite, la relazione con i clienti e la successiva manutenzione. Tutti questi componenti possono essere gestiti in Real Time facendo leva sull'intelligenza ambientale (ad esempio quella della building automation e del Facility Management) connessa con l'intelligenza che anima i prodotti stessi. Con L'Industria 4.0 cambia completamente la pianificazione del ciclo di vita e il modo in cui l'azienda gestisce i prodotti, li segue e li controlla anche quando i processi di produzione si svolgono su più imprese diverse, in contesti diversi, con partner diversi in zone geografiche diverse.

L'Industria 4.0 e la Smart Manufacturing hanno le proprie basi nella connessione e nella integrazione e permettono di avere una visione di insieme di tutte le fasi del ciclo di vita del prodotto, anche quando esce dall'azienda di produzione ed entra nel circuito della distribuzione per entrare nelle case o nelle fabbriche dei clienti, producendo dati in grado di fornire valore e nuovi servizi.



### OBIETTIVI

Questo progetto ha come scopo quello di preparare i destinatari alle implicazioni dei prevedibili impatti che l'Internet Of Things (IIoT) avrà sulle esigenze di analisi dati delle aziende. Infatti, la crescente presenza di oggetti intelligenti (es. sensori e attuatori) in grado di comunicare tra loro e con gli strumenti di controllo, apre scenari del tutto nuovi per la comprensione e per il presidio dei processi aziendali. La quarta rivoluzione industriale, nota come Industry 4.0., è certamente spinta dalla crescita di nuove tecnologie fortemente guidate dal digitale (Additive Manufacturing, Advanced Robotics,...) ma ha al centro la comunicazione e la "collaborazione digitale" tra aziende clienti e fornitori, processi aziendali e apparati industriali, in cui IIoT è uno degli elementi più rilevanti. Questa comunicazione continua tra apparati (sensori, attuatori e controller) riveste sia una valenza di tipo tattico (il controller elabora input di sensori per agire attraverso attuatori) che una valenza "strategica" nel momento in cui una attenta analisi dei dati scambiati può permettere di capire meglio i processi di business e fare delle azioni di miglioramento continuo. Ma può anche fare comprendere meglio il comportamento di clienti e fornitori per modificare di conseguenza la propria strategia. Questo implica la necessità di dominare gli aspetti relativi ai dati: storicizzazione dei dati rilevanti, protezione della riservatezza dei dati "critici", assicurazione della disponibilità dei dati che saranno il cuore della fabbrica del domani. Questo modulo è quindi particolarmente indicato per soggetti con competenze sui processi industriali.

### DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in regione Emilia Romagna in data antecedente l'iscrizione alle attività.

Il progetto 4.0 OPERATIONS (IOT&BIGDATA ANALYTICS) è particolarmente rivolto a soggetti con formazione di tipo tecnico-scientifico, ad esempio alcune branche dell'ingegneria (Gestionale, Informatica, Meccanica). Alcune di queste competenze saranno fornite nel progetto introduttivo di allineamento della competenze, INTRO TO BIG DATA e al Progetto 2, BASI DI STATISTICA, MATEMATICA, GESTIONE DB, ciononostante per il forte focus sui processi industriali, è un progetto destinato a profili con già una esposizione a queste tematiche.

In particolare le competenze raccomandate in ingresso sono:

- un livello base di Business Process Analysis
- un livello base di statistica
- un livello base di programmazione in ambiente Python

### CONTENUTI DEL PERCORSO

#### PRINCIPI BASE DI INDUSTRIA 4.0 E TECNOLOGIE IN AMBITO.

- Tecniche di analisi e modellazione dei processi aziendali
- Internet of Things (Introduzione)
- Esigenze HW e SW di un progetto IIoT (Arduino)
- Comprendere come vengono gestiti i dati in un rete IIoT
- Disegno di soluzioni software per Big Data e IIoT
- Costruzione di script Python per gestire file dati di grandi dimensioni, raccogliere dati di sensori e interagire con il mondo reale tramite attuatori e altri dispositivi di output
- Laboratorio di sviluppo di soluzioni IIoT e Big Data Analytics
- Costruire strumenti "Decision Support Systems" per guidare i processi della fabbrica
- Laboratorio di disegno di strumenti di analisi e simulazione basati su Big Data
- Testimonianza CISCO

### COMPETENZE

L'obiettivo del percorso è offrire una comprensione completa dell'Internet of Things (IIoT). Lo scopo è sviluppare abilità fondamentali utilizzando attività di laboratorio pratiche che stimolano gli studenti ad applicare la risoluzione creativa dei problemi e la prototipazione rapida nel campo interdisciplinare dell'elettronica, del networking, della sicurezza e dell'analisi dei dati. L'approccio incentrato sulla persona si traduce in un professionista in grado di ideare, progettare, prototipare e presentare una soluzione IIoT per esigenze aziendali già presenti o per lo sviluppo di nuovi prodotti.

### DESCRIZIONE DURATA

Il corso della durata di 80 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

### PRIMA CHIUSURA ISCRIZIONI: 31/12/2018

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

### RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna

# TECNOLOGIE E SOFTWARE PER IL TRATTAMENTO DEI BIG DATA

## DESCRIZIONE DEL PERCORSO

Il valore di un dato si moltiplica del 40% quando correlato e combinato con altri dati. Per sfruttare appieno la mole di dati, la loro grande varietà e variabilità sono indispensabili architetture di gestione dei big data flessibili, scalabili, sicure ed efficienti: da qui la necessità di un Hub di dati aziendali (EDH) per gestire, trasformare, pulire, integrare i dati a disposizione. In questo modo è possibile rafforzare il coinvolgimento del cliente, ottimizzare l'offerta e migliorare le previsioni e il merchandising. Dunque, un hub aziendale che offra due vantaggi principali: la governance multi-tenancy e la sicurezza dei dati. Questo sistema è ovviamente applicabile a ogni tipologia di settore economico.

Per portare solo alcuni esempi:

1. per il settore viaggi: la customer journey è fondamentale per poter dare un servizio migliore ad esempio da parte della compagnie aeree che hanno la necessità di conoscere e monitorare le scelte degli utenti, la soddisfazione dei servizi erogati, i flussi di informazioni, la logistica, ecc.
2. per il settore manifatturiero: l'ottimizzazione della supply chain e del controllo di tutta la filiera di distribuzione dei prodotti è fondamentale tanto quanto il monitoraggio della logistica e distribuzione al fine di migliorare la vendita del prodotto finale
3. per il settore smart city ed energy: la costruzione delle abitazioni moderne dovrebbe passare per l'implementazione e lo sviluppo della domotica e dell'IoT, l'analisi delle esigenze prioritarie degli utenti, la gestione della mobilità cittadina e delle emissioni inquinanti possibili solo tramite la conoscenza e gestione dei flussi di traffico e delle abitudini dei cittadini di un dato territorio.



## OBIETTIVI

Uno dei principi chiave per operare con i Big Data è lo stoccaggio di tutti i dati originali, indipendentemente da quando questi saranno utilizzati. Quindi col tempo gli archivi possono assumere dimensioni anche incredibilmente elevate. Anche se nulla impedisce di realizzare l'archiviazione dei dati tramite un classico database relazionale, spesso questa scelta porta a investire risorse economiche importanti sia in termini computazionali, sia di storage. Questi e altri motivi portano alcuni colossi dell'innovazione, tra cui Google e Facebook, ad adottare strumenti diversi dagli RDBMS per gestire e i loro Dataset: tra le tecnologie Open Source create per questo scopo una delle più diffuse e utilizzate è Apache Hadoop. Il progetto ha come obiettivo quello di permettere ai partecipanti di comprendere le architetture esistenti per il trattamento di Big Data e per la memorizzazione in db noSQL, in modo da essere in grado di utilizzare le principali funzionalità di questi strumenti.

## DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività.

Il progetto TECNOLOGIE E SOFTWARE PER IL TRATTAMENTO BIG DATA) è particolarmente rivolto a soggetti con formazione di tipo tecnico-scientifico, in particolare in ambito informatico (Ingegneria informatica/telecomunicazioni) Alcune di queste competenze saranno fornite nel progetto introduttivo di allineamento della competenze, INTRO TO BIG DATA e al Progetto 2, BASI DI STATISTICA, MATEMATICA, GESTIONE DB, ciononostante per il forte focus sui processi industriali, è un progetto destinato a profili con già una esposizione a queste tematiche.

In particolare le competenze raccomandate in ingresso sono:

- un livello base di gestione server e cloud
- un livello base di gestione Data Base (SQL, noSQL)
- un livello base di programmazione Java



## CONTENUTI DEL PERCORSO

- Overview della definizione di soluzioni architetturale per il trattamento di Big Data Hadoop common (strato software comune con funzioni di supporto)
- Hadoop Distributed File System (HDFS - derivato da Google's GFS)
- Hadoop tools, ecosistema e distribuzioni
- Comprendere MapReduce
- Utilizzo di tecnologie open source per trasferire i dati processati da HDFS in un database SQL e viceversa (ETL)
- Differenze tra SQL e linguaggi usati per i BIG DATA
- Operazioni CRUD e aggregazioni avanzate
- Elaborazione dati ed esportazione verso NOSQL
- Utilizzo di MongoDB Connector for Hadoop
- L'importanza di saper comunicare i dati: strumenti di "Data Visualization"
- Algoritmi di data mining (per operazioni di classificazione, regressione, clusterizzazione)
- Algoritmi per la market basket analysis
- Algoritmi per trovare item simili in larghe quantità di dati
- Librerie Python per la gestione e la visualizzazione dei dati (ad es. Pandas, Scikit-Learn, Matplotlib, Seaborn)
- Apache Spark per l'analisi dei dati: RDD, Structured queries, e use case di esempio)



## COMPETENZE

Il percorso è verticale rispetto al principale framework dei Big Data: Apache Hadoop, che supporta applicazioni distribuite con elevato accesso ai dati sotto una licenza libera, in particolare consente di elaborare in modo efficiente grandi set di dati. La gestione dei dati può comportare query su dati strutturati all'interno di database di grandi dimensioni, oppure ricerche su file system distribuiti od infine operazioni di manipolazione, visualizzazione, trasferimento e cancellazione di file e directory distribuite su più server. L'analisi computazionale comporta lo sviluppo di metodi ed algoritmi scalabili per la gestione e l'analisi di Big Data. L'analisi dei dati e la modellazione può comportare la riduzione dimensionale, la suddivisione (clustering), la classificazione (ranking), la previsione e la possibilità di future estrazioni.

La Visualizzazione può coinvolgere modalità grafiche che forniscono informazioni su grandi quantità di dati, ricchi di informazioni visive con riepiloghi dei risultati, dove l'analisi visuale talvolta è il momento stesso in cui i dati vengono valutati. Apache Hadoop è un framework che consente l'elaborazione distribuita di grandi insiemi di dati attraverso cluster di servers, oppure sui servizi di cloud computing. È stato progettato per scalare da singolo server a migliaia di macchine distribuite, ognuna delle quali offre capacità di calcolo e di immagazzinamento.

Per la gestione dei Big Data offre alcune componenti molto utilizzate: HDFS e MAP Reduce. Apache Hadoop è un ambiente che supporta applicazioni distribuite con elevato accesso ai dati sotto una licenza libera; permette alle applicazioni di lavorare con migliaia di nodi e petabyte di dati. Hadoop è stato ispirato dalla MapReduce di Google e dal Google File System. Un grande numero di librerie software utilizzano il framework Hadoop per gestire e analizzare i dati. Nella fattispecie, trovano una larga diffusione in questo momento, in caso di dataset di estensione limitata librerie Python per la gestione e la predizione dei dati quali ad esempio Pandas e Scikit Learn e, in caso di dati con maggiore estensione, applicazioni costruite con il framework Apache Spark.

### DESCRIZIONE DURATA

Il corso della durata di 156 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

### PRIMA CHIUSURA ISCRIZIONI: 30/03/2019

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna

# DEEP LEARNING - ARTIFICIAL INTELLIGENCE NATURAL LINGUISTIC PROCESSING

## DESCRIZIONE DEL PERCORSO

Il progetto DEEP LEARNING, ARTIFICIAL INTELLIGENCE, NATURAL LINGUISTIC PROCESSING è relativo alle nuove applicazioni tipiche della AI-Artificial Intelligence che vanno sotto il nome di Deep Learning e di una sua importante applicazione detta Natural Language Processing ovvero il trattamento informatico (computer processing) del linguaggio naturale, per qualsiasi scopo. È pertanto un percorso formativo centrale dentro il progetto perché risponde all'esigenza di avvicinarsi ad un settore di grande sviluppo tra le alte tecnologie; i contenuti si occupano di prodotti che oggi vediamo nella vita di tutti i giorni, come ad esempio Siri nell'iPhone o i traduttori alla portata di ogni utente di Internet, o i prodotti di analisi del Sentiment e delle opinioni dei clienti nei social media, tutti basati sul Natural Language Processing.



## OBIETTIVI

L'obiettivo del corso è quello di introdurre gli studenti alle tematiche relative all'AI ed in particolare al Deep Learning e al Natural Language Processing, attraverso inquadramenti teorici, casi applicativi, utilizzo di metodologie e tecnologie. Al termine del percorso i partecipanti avranno compreso cosa si intende per Deep Learning e Natural Language Processing, come si è arrivati allo stato attuale dell'arte delle conoscenze e della ricerca, quali sono le tecniche che si adottano ad oggi e quali sono i principali ambiti di applicazione. In particolare quindi per il Deep Learning sapranno distinguere e prepararsi a lavorare sulle Neural Networks (DNN; RNN, CNN...) con applicazioni nei settori delle natural language processing, - automatic speech recognition, image recognition, visual art processing, customer relationship management (CRM), mobile advertising, bioinformatics/medical informatics. Relativamente ad NLP analogamente alla luce della conoscenza di analisi lessicali, grammaticali, sintattiche, semantiche sapranno in futuro iniziare percorsi lavorativi per la realizzazione di Bot, chatbot, virtual assistant (testuali, vocali), Sentiment analysis, Traduzioni, Estrazione di informazioni e ricerca.

## DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività.

Il progetto DEEP LEARNING, ARTIFICIAL INTELLIGENCE, NATURAL LINGUISTIC PROCESSING è un percorso specialistico previsto per i laureati dell'area scientifica. Questa non è una disciplina aperta soltanto agli informatici: poiché ha a che vedere con la raccolta, la collocazione e l'analisi dei dati, spazia in numerosi ambiti, e matematica, statistica, programmazione possono giocare tutte una parte molto importante. Le competenze necessarie in ingresso, oltre al requisito formale della laurea, sono trasversali: possono accedere fisica, informatica, statistica classica, bio-informatica e ingegneria chimica, con una forte attitudine alle tecnologie digitali, predisposizione al pensiero critico e capacità di autoaggiornamento.

Chiaramente è sempre possibile che, eccezionalmente, un candidato con background umanistico o giuridico economico frequenti il progetto, se ha già acquisito competenze di tipo matematico, informatico e statistico di base ed in fase di consulenza orientativa manifesta forte motivazione e successo formativo nei progetti propedeutici.

I prerequisiti comunque ideali per affrontare questo corso sono:

- esperienza di programmazione (es. linguaggio Python);
- conoscenza dei concetti fondamentali del Machine Learning (come quelli conseguibili frequentando il progetto 5 di questa operazione);
- Conoscenza delle basi della Statistica dell'Analisi matematica e dell'Algebra lineare.

## CONTENUTI DEL PERCORSO

### INTRODUZIONE A DEEP LEARNING:

- Introduzione alle reti neurali e al layering di algoritmi di machine learning: reti neurali e deep learning concetti base, funzioni di costo
- Algoritmi e use case di Supervised, Unsupervised, Reinforcement Learning applicati al Deep Learning: costruzione di sistemi di deep learning basati su layer di ML
- Sviluppo con framework Tensorflow:
- Introduzione al framework ed esercitazioni
- Applicazioni nel contesto "traditional data" (tabelle, log, ...):
- Sviluppo su dati tabellari, analisi di serie temporali, regressioni
- Applicazioni in contesto non strutturato
  - Computer vision
  - Speech recognition
  - Natural Language Processing
- Altri tool di rapid prototyping: intro ai framework di sviluppo come RapidMiner, Azure ML

### APPROFONDIMENTI SU DEEP LEARNING:

- ANN: Artificial Neural Networks
- DNN: Deep Neural Networks
- RNN Recurrent Neural Networks
- CNN: Convolutional Neural Networks

### LE PRINCIPALI APPLICAZIONI DEL DEEP LEARNING:

- Natural language processing
- Automatic speech recognition
- Image recognition
- Visual art processing
- Customer relationship management

## CONTENUTI DEL PERCORSO

### LE PRINCIPALI APPLICAZIONI DEL DEEP LEARNING:

- Mobile advertising
- Bioinformatics / medical informatics

### LE CARATTERISTICHE PRINCIPALI DEL NATURAL LANGUAGE PROCESSING E LE APPLICAZIONI:

- Analisi lessicale
- Analisi grammaticale
- Analisi sintattica
- Analisi semantica
- Bot, chatbot, virtual assistant: testuali, vocali
- Sentiment analysis
- Traduzioni
- Estrazione di informazioni e ricerca

## COMPETENZE

I contesti lavorativi in cui ad oggi spendere le conoscenze e competenze ricavabili con questo corso sono legati agli ambiti applicativi succitati sia per il Deep Learning che nello specifico per il NLP: quindi innanzitutto quelli della Ricerca e Sviluppo ma anche del Marketing e Vendite e Delivery di Servizi evoluti e smart sia per il mondo B2B che B2C. La spendibilità delle conoscenze acquisite è dimostrata da numerosi studi previsionali sulle professioni del futuro; molte organizzazioni investono nell'acquisizione di masse di dati che permettono migliore conoscenza delle abitudini di consumo e acquisto della clientela e l'uso di questo patrimonio richiede conoscenze specialistiche che non sono più solo informatiche. Secondo Stack Overflow, che gestisce una ampia comunità online di softwaristi (oltre 64.000), gli esperti di machine learning e di NLP sono tra le figure più ricercate dal mercato e il percorso qui proposto si limita ad anticipare questo fabbisogno ormai più che emergente sul mercato italiano. Con una percentuale del 14,3, gli specialisti in machine learning occupano le posizioni più in alto tra gli sviluppatori che affermano di cercare un nuovo posto di lavoro. Al secondo posto seguono i data scientist con il 13,2 per cento. Le aziende cresciute su Internet, ed oggi anche le start up innovative, raccolgono enormi quantità di informazioni e costituiscono la naturale destinazione di lavoratori di questo tipo.

### DESCRIZIONE DURATA

Il corso della durata di 156 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

### PRIMA CHIUSURA ISCRIZIONI: 31/12/2018

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna

# BIG DATA & START UP

## DESCRIZIONE DEL PERCORSO

Il mondo dei "Big Data" è uno degli ambiti su cui maggiormente si sta concentrando il focus delle startup: sia come ambito di attività principale (es. fornitura di servizi di analisi traffico canali social) sia come competenza necessaria per abilitare altri servizi (es. analisi del comportamento dei propri clienti, matching ideale tra domanda e offerta di servizi, ecc.). Agire all'interno di una startup presenta dinamiche completamente differenti dall'agire all'interno di aziende ben consolidate, con processi di business definiti e chiara governance. Si deve essere in grado di rivestire più ruoli, essendo generalmente realtà ancora di ridotte dimensioni, e si deve essere in grado di dominare aspetti di gestione societaria: raccolta capitali con presentazione ad investitori dei piani e dei risultati raggiunti, aspetti di gestione del personale, presentazione di domande per finanziamenti, ecc.



## OBIETTIVI

Il progetto ha come obiettivo quello di fornire agli studenti una buona comprensione dei passi da compiere per sviluppare una startup nel mondo dei Big Data. Saranno analizzati in dettaglio tutti gli aspetti del fare impresa, dalla creazione di un business plan, alla raccolta del capitale, alla gestione dei rapporti tra soci, alla conduzione dell'azienda.

Le conoscenze qui apprese saranno utili sia per chi intenda sviluppare una nuova startup, sia per chi abbia l'ambizione di collocarsi all'interno di startup già esistenti, che richiedono personale in grado di adattarsi velocemente al modo di lavorare di una startup. Inoltre il progetto avrà la funzione di creare professionisti in grado un domani di essere attori attivi nel mondo della open innovation, in cui sempre più le grandi aziende ricercano persone in grado di identificare e stringere relazioni con startup per portare l'innovazione all'interno delle aziende in modo complementare ai dipartimenti di Ricerca e Sviluppo interni.

## DESTINATARI E REQUISITI D'ACCESSO

12 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività.

Il progetto BIG DATA & STARTUP è indicato per laureati di ogni area disciplinare. L'obiettivo del progetto è di creare una piena consapevolezza di cosa significa costituire e lavorare in una startup con i Big Data come elemento distintivo, valorizzando per ognuno il background di provenienza.

Per tutti i partecipanti il progetto sarà una occasione di spaziare al di fuori della propria zona di confidenza e maturare competenze utili nel dialogare con gli altri attori della startup, clienti e fornitori.



## CONTENUTI DEL PERCORSO

- Orientamento all'imprenditorialità: entrepreneurial mindset
- Competenze imprenditoriali:
  - Strumenti/metodologie per prendere decisioni, concetto di rischio
  - Come si scrive e si valuta un piano economico finanziario
  - Come si ottimizzando i rapporti con le banche
- Elementi di web marketing, comunicazione, promozione, utilizzo dei social media
- Introduzione alle forme giuridiche e identificazione di quelle più opportune
- Introduzione alla normativa fiscale (incl. Agevolazioni per startup, ricerca...)
- Aspetti di contrattualistica e contabilità
- Validazione del modello di business con metodologie quali il Business Model Canvas
- Come identificare e inserire in azienda le risorse umane
- Come motivare il personale, team coaching
- Implicazioni del GDPR sulla conduzione della startup
- Open Innovation nelle aziende
- Testimonianze di aziende startup nel campo dei Big data



## COMPETENZE

Le principali attività economiche delle startup in regione sono legate all'ICT, dalla produzione di software alla consulenza informatica, ma anche hardware, prodotti di elettronica, ottica, apparecchi elettromedicali.

Questo progetto ha, quindi, come scopo quello di preparare i destinatari all'essere imprenditori nel campo dei Big Data, e non ha una particolare focalizzazione su uno specifico insieme di studenti e mercati; sono competenze che ogni socio di startup dovrebbe comprendere e dominare per portare avanti la sua impresa in modo consapevole attraverso:

1. Competenze e conoscenze relative alla promozione della business idea alla base della startup
2. Competenze e conoscenze relative alla trasformazione della business idea in modello imprenditoriale di business
3. Competenze e conoscenze relative alla gestione imprenditoriale della startup

## DESCRIZIONE DURATA

Il corso della durata di 80 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

## PRIMA CHIUSURA ISCRIZIONI: 30/03/2019

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna

# PROJECT WORK SETTORIALE

## DESCRIZIONE DEL PERCORSO

Da quanto emerge da diverse ricerche degli ultimi anni (AlmaLaurea, Camere di Commercio, Associazione Treelle, Fondazione Rocca) le imprese italiane, nella valutazione relativa all'inserimento di candidati, si affidano a vari elementi: le conoscenze pratiche che il candidato ha del settore in cui andrà ad operare, la conoscenza delle lingue e, soprattutto, il possesso delle soft skills, cioè competenze trasversali (ad es. la capacità di adattarsi, il problem solving, il team working, l'analisi e la sintesi, l'autonomia di giudizio, la comunicazione efficace ecc...). Le competenze trasversali finiscono per essere spesso decisive per l'inserimento in azienda e nei percorsi di carriera. Questi riscontri sono stati di recente corroborati dai contenuti della "New Skills Agenda", con cui a giugno 2016 la commissione europea ha delineato le competenze strategiche per favorire la crescita e l'innovazione ed evitare l'esclusione socio-lavorativa. Indica le soft skills come leva strategica sia per colmare i gap di circa 70 ml di cittadini europei con insufficienti competenze trasversali che per qualificare a tutto tondo la domanda di lavoratori altamente specializzati in grado di entrare da protagonisti nella rivoluzione 4.0.

Il project work nasce proprio per allenare gli studenti a sperimentare e sviluppare le soft skills. Attraverso il project work i partecipanti non solo potranno lavorare in team work, ma dovranno cimentarsi con le modalità e i flussi di comunicazione organizzati tipici del mondo del lavoro, mettendosi alla prova per trovare soluzioni reali a concreti casi aziendali, in maniera collaborativa. La metodologia, ampiamente utilizzata nei percorsi formativi per dare concretezza alle acquisizioni teoriche applicandole su casi concreti mutuati da contesti reali, viene qui ulteriormente allargata, anche nella durata, con l'obiettivo di dare una contestualizzazione all'impiego dei big data nei diversi domini e settori possibili.



## OBIETTIVI

Il progetto "Project work settoriale" rappresenta la fase conclusiva dell'Operazione a cui potranno accedere destinatari provenienti dalle tre aree disciplinari universitarie, dopo aver frequentato almeno uno o più progetti specialistici. L'obiettivo del progetto è di creare un vero e proprio spazio di applicazione della teoria acquisita nei progetti precedenti ad un caso concreto mutuato da aziende e dati reali. Si pone idealmente come un ponte verso le imprese per creare occasioni durante le quali i partecipanti possano acquisire consapevolezza dei loro possibili percorsi di inserimento professionale e si possa avere anche un match tra domanda e offerta. Ognuno dei partecipanti, in base all'area di provenienza (umanistica, giuridica-economica o scientifica) ed al percorso fatto all'interno dell'Operazione, potrà scegliere di iscriversi ad un project work valutando insieme al team didattico quello più idoneo. Ovviamente, dato l'ulteriore specializzazione prevista dall'applicazione a casi pratici con le variabili e la complessità del caso, si auspica e si stima che i partecipanti che provengono dalle aree umanistica e giuridico economica, frequentino almeno i primi due progetti (salvo casi di competenze pregresse per il progetto 2) ed almeno un progetto specialistico. In questo modo, scegliendo poi il project work più idoneo, l'esperienza potrà avere davvero un valore aggiunto ed avviare percorsi professionali concreti.

## DESTINATARI E REQUISITI D'ACCESSO

6 Giovani in possesso di una laurea triennale o magistrale o una laurea a ciclo unico, conseguita da non più di 24 mesi, residenti o domiciliati in Regione Emilia Romagna in data antecedente l'iscrizione alle attività.

Il progetto "Project work settoriale" rappresenta la fase conclusiva dell'Operazione a cui potranno accedere destinatari provenienti da tutti i tre percorsi universitari, dopo aver frequentato uno o più progetti.

Ognuno dei partecipanti, in base all'area di provenienza (umanistica, giuridica-economica o scientifica) ed al percorso fatto all'interno dell'Operazione potrà scegliere di iscriversi ad un project work valutando insieme al team didattico quello più idoneo.



## CONTENUTI DEL PERCORSO

Il progetto della durata di 80 ore è articolato in:

- 20 ore di aula durante le quali il tema verrà presentato con i necessari approfondimenti in termini teorici e di contesto;
- 60 ore di lavoro individuale o in gruppo, a distanza o in presenza, per sviluppare il project work, curando anche la presentazione del lavoro e la reportistica;
- presentazione finale alla presenza di imprese, relatori, team didattico con modalità tradizionale o in forma di hackaton.

Temi e ambiti applicativi saranno definiti in collaborazione tra team di esperti di contenuti e didattici e discussi e/o validati dal CTS. Una parte fondamentale del project work è la presentazione finale, davanti a gruppi più o meno allargati. In questo caso si prevede l'organizzazione di un momento finale di presentazione con relatori esterni o interni all'operazione provenienti da enti e soggetti pubblici o privati coinvolti e interessati ai BIG DATA.



## COMPETENZE

Il project work è una metodologia didattica ormai consolidata, soprattutto nell'ambito della formazione professionalizzante e specialistica, che si ispira al principio generale del learning by doing, apprendere facendo. Il coordinatore dell'attività, da solo o supportato da esperti, guida nello sviluppo del progetto relativo a contesti reali, indicando obiettivi e possibili soluzioni in caso di stallo. I risultati del pw vanno elaborati, discussi e analizzati in modo collegiale per verificarne la validità e la fattibilità e procedere nei diversi steps ognuno con il proprio eventuale ruolo. Il progetto, una volta terminato, viene presentato davanti ad un collegio formato da team didattici, docenti interni ed esterni, rappresentanti di aziende ed enti, che valutano l'operato dei gruppi.

## DESCRIZIONE DURATA

Il corso della durata di 80 ore si svolgerà indicativamente con una media di 32 ore settimanali in orario diurno

## PRIMA CHIUSURA ISCRIZIONI: 30/03/2019

La data si riferisce alla scadenza della prima edizione del corso. Sono previste varie edizioni e per le scadenze successive si rimanda al sito [www.bigdata-lab.it](http://www.bigdata-lab.it)

## RIFERIMENTI

Operazione Rif. PA 2018-9495/RER approvata con DGR n° 633 del 2 maggio 2018 e co-finanziata dal Fondo Sociale Europeo PO 2014-2020 Regione Emilia-Romagna





# DATA LAB

## GUARDA AVANTI

**Big Data**, nuove competenze  
per nuove professioni



Per informazioni:

Tel. 800 036425 (Numero verde)

Mail. [bigdata@formindustria.it](mailto:bigdata@formindustria.it)

Sito. [www.bigdata-lab.it](http://www.bigdata-lab.it)

