

*Gradiremmo conoscere se ritiene utile o, viceversa, inutile creare una nuova classe di laurea magistrale destinata a contenere corsi di studio riguardanti i vari aspetti riguardanti la Data Science.*

Premetto che non ho un'opinione precisa a riguardo, la cosa certa è che è necessario intervenire per permettere la formazione dei data scientist. Infatti non è più possibile evitare di dare seguito a quanto scritto nel rapporto sui Big Data redatto dal gruppo di lavoro MIUR e reso pubblico nel luglio 2016, ovvero la necessità di:

- inserire moduli di **data science** in tutti i corsi di studio
- realizzare percorsi di laurea magistrale focalizzati che potrebbero in molti casi essere inseriti in una varietà di classi disciplinari distinte a seconda dello specifico settore nel quale il **data scientist** svolgerà la sua attività (es: settore biomedico, scienze sociali e umane)
- realizzare che la collocazione di diverse classi disciplinari a seconda dello specifico settore è utile ad **evitare ostacoli** in ingresso e **barriere disciplinari** per il prosieguo della vita accademica/professionale degli studenti.

Un intervento è decisamente necessario, ma non mi è chiaro se sia più utile intervenire sulle singole classi (in particolare di laurea magistrale), per le quali si verificasse questa l'esigenza, aggiungendo un ambito specifico denominato "Data Science" che contenga i SSD propri di questa specializzazione (ad es. INF/01, MAT/08, MAT/09, ING-INF/05, SECS-S/01, SECS-S/02) oppure istituire una nuova classe di LM che affianchi ad un ambito "Data Science" ambiti disciplinari quali la Fisica, la Psicologia, l'Ingegneria, l'Economia ecc.

Le due scelte comportano sia rischi che benefici, nel primo caso infatti sarebbe automaticamente garantito che non ci sarebbero ostacoli per il prosieguo della carriera in ambito accademico dove un certo numero di queste figure con preparazione sia in ambito Data Science che in ambito Disciplinare sono decisamente necessarie, d'altro canto questa scelta rischia di frammentare in tanti rivoli disciplinari indipendenti le conoscenze e le applicazioni di Data Science sfavorendo il confronto interdisciplinare.

Per contro, nel secondo caso ci si troverebbe a gestire corsi di laurea magistrale con molti curricula (se penso alla situazione del mio Ateneo, già dal prossimo anno avremo due LM distinte e due curricula dedicati in altre due LM) con le conseguenti difficoltà e le possibili ricadute sui nuovi laureati che potrebbero essere sfavoriti dal titolo di Data Scientist al momento di tentare una carriera in ambito accademico. Ovviamente la situazione potrebbe essere evitata istituendo corsi di LM non affini nella medesima classe, ricadendo così de facto nel primo caso. Va in ogni caso sottolineato che presso gli Atenei stranieri esistono Dipartimenti o Centri interdisciplinari dedicati al Data Science.

A questo punto, riguardo alle domande da voi poste, di seguito trovate le mie risposte

✓ *Quali dovrebbero essere gli obiettivi culturali di questa classe e i contenuti disciplinari e le competenze trasversali indispensabili che dovrebbero essere fornite da tutti i corsi della classe?*

- da un lato mettere lo studente in condizione di affrontare le sfide emergenti nella ricerca in uno specifico ambito disciplinare data la necessità di combinare le competenze del ricercatore tradizionale con quelle del "data scientist";
- dall'altro costruire un profilo accademico e professionale particolarmente adeguato alla modellizzazione e analisi di Big Data in ambiti professionali esterni, applicativi e aziendali.

A tale scopo, i percorsi formativi dovrebbero prevedere un'offerta che permetta di approfondire temi di Contenuto Disciplinare, Statistica, Programmazione e gestione di grandi data-set, trattando questi argomenti non separatamente e indipendentemente, ma in modo coordinato, organico e interdisciplinare.

Nel dettaglio, i corsi, oltre alla parte strettamente disciplinare, dovrebbero provvedere:

- Conoscenze avanzate di machine learning, in particolare di programmazione e delle reti neurali;
- Corsi di data mining, analisi e gestione di grandi data-set;
- Corsi avanzati di statistica specifici per ciascuna disciplina;

- Conoscenza delle strutture hardware e corsi di High Performance Computing;
- Esperienza diretta delle tecniche informatiche di calcolo, con applicazioni pratiche ed esercitazioni di laboratorio in ambiti di rilevanza scientifica;

✓ *Quali potrebbero essere dei naturali sbocchi professionali, o sbocchi verso il proseguimento degli studi, coerenti con gli obiettivi della classe?*

- Qualunque azienda manifatturiera che dall'analisi di grandi moli di dati complessi voglia ottenere dei feed back per ottimizzare la produzione e i servizi verso il cliente
- Amministrazioni pubbliche
- Banche e finanza
- Enti pubblici deputati alla gestione delle politiche per l'ambiente
- Ospedali
- Uffici statistici
- Centri di calcolo
- Ambito accademico/enti di ricerca

✓ *È necessario introdurre altri elementi (per esempio presenza obbligatoria di tirocini o stage, attività laboratoriali, competenze linguistiche, eccetera) indispensabili per il raggiungimento degli obiettivi della classe?*

Ritengo sia necessario introdurre elementi complementari quali

- Tirocinio obbligatorio (in azienda o ente partner)
- Abilità comunicative
- Etica e privacy nella gestione dei dati
- Capacità di interpretare i dati ottenuti
- Conoscenze linguistiche
- Capacità di vedere uno stesso problema da punti di vista diversi
- Attività che insegnino allo studente come affrontare con il medesimo approccio metodologico problemi complessi in ambiti tra di loro diversi riconoscendo le analogie e gli aspetti comuni all'interno delle differenti tematiche