BIG DATA ANALYTICS

Dai dati alla conoscenza, dalla conoscenza alle decisioni



LEZIONE INTRODUTTIVA

Prof.ssa Romina Eramo
Dip. Scienze della Comunicazione - UNITE

Una giornata qualsiasi. Quanti dati produciamo?

Ogni interazione digitale lascia tracce che possono diventare oggetto di analisi.

01

Social media

Like, commenti, condivisioni, tempo di visualizzazione

02

Musica e video

Ascolti, ricerche, preferenze, raccomandazioni

03

Acquisti online

Ricerche, clic, carrelli, conversioni

04

Mobilità e servizi

Percorsi, prenotazioni, utilizzo di app

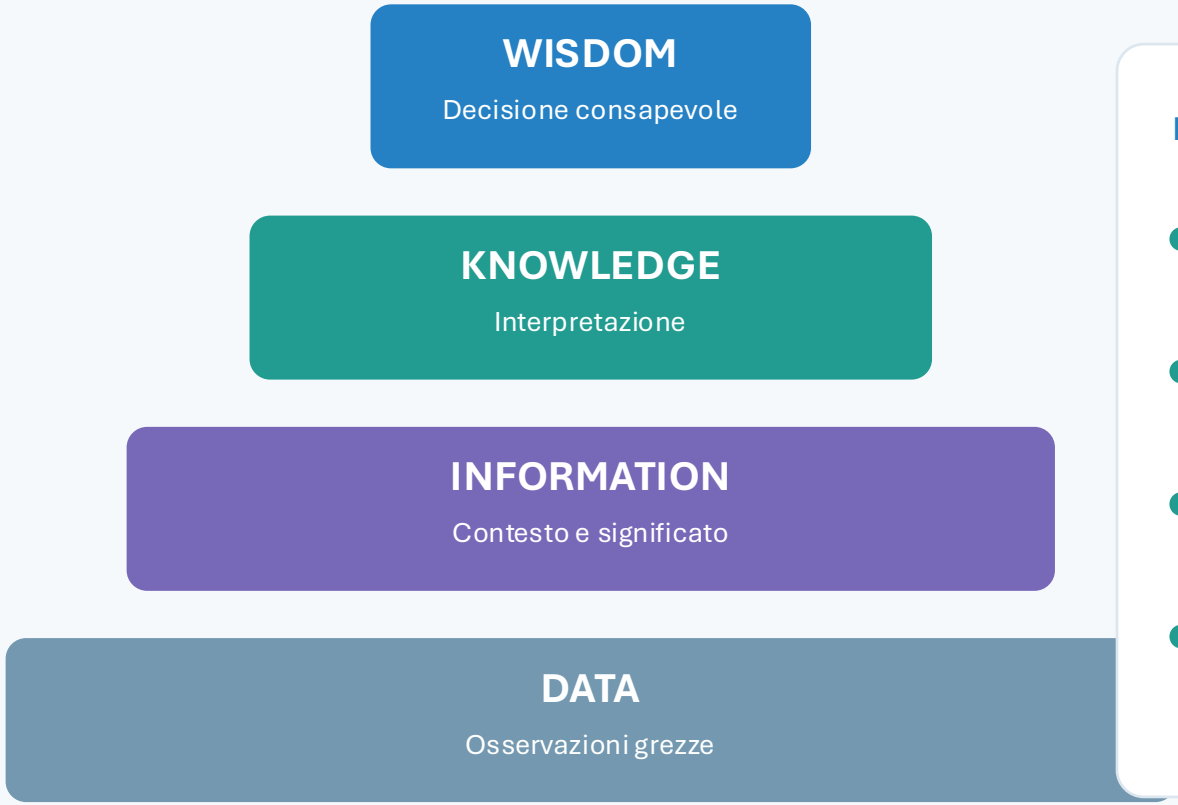
Dai dati alle domande di comunicazione

Il punto di partenza di un’analisi è sempre una domanda rilevante.

Perché un contenuto diventa virale?	Engagement e diffusione
Come cambia la reputazione di un brand?	Social listening
Una campagna ha raggiunto il pubblico giusto?	Performance e segmentazione
Che cosa interessa ai cittadini?	Comunicazione pubblica

Dal dato alla conoscenza: modello DIKW

Il dato, senza contesto e interpretazione, non basta per decidere.



ESEMPIO SOCIAL

- 12.000 visualizzazioni
- +40% rispetto alla media
- Il video sembra più efficace
- Testare più video nella campagna

Che cosa sono i Big Data?

Non si tratta soltanto di “tantissimi dati”.

Dati la cui scala, velocità, varietà o complessità richiede approcci e tecnologie adeguati per la gestione e l’analisi.

- La dimensione conta, ma non è l’unico aspetto.
- Le sfide riguardano raccolta, integrazione, qualità ed elaborazione.
- L’obiettivo finale è ricavare valore dai dati.

Le 5 V dei Big Data

Cinque prospettive per comprenderne le caratteristiche.

V — VOLUME

Quantità

Milioni di interazioni

V — VELOCITY

Velocità

Commenti in tempo reale

V — VARIETY

Varietà

Testi, immagini, video

V — VERACITY

Attendibilità

Bot, errori, rumore

V — VALUE

Valore

Insight per le decisioni

Tipologie e fonti dei dati

La varietà dei dati incide sugli strumenti di analisi.

STRUTTURATI

Tabelle, database

Es.: clic, impression, date

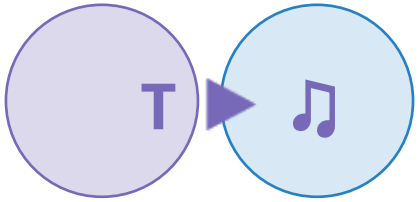
SEMI-STRUTTURATI

```
{
  "post": 1042,
  "likes": 280,
  "type": "video"
}
```

JSON, log, metadati

Es.: eventi di una app

NON STRUTTURATI



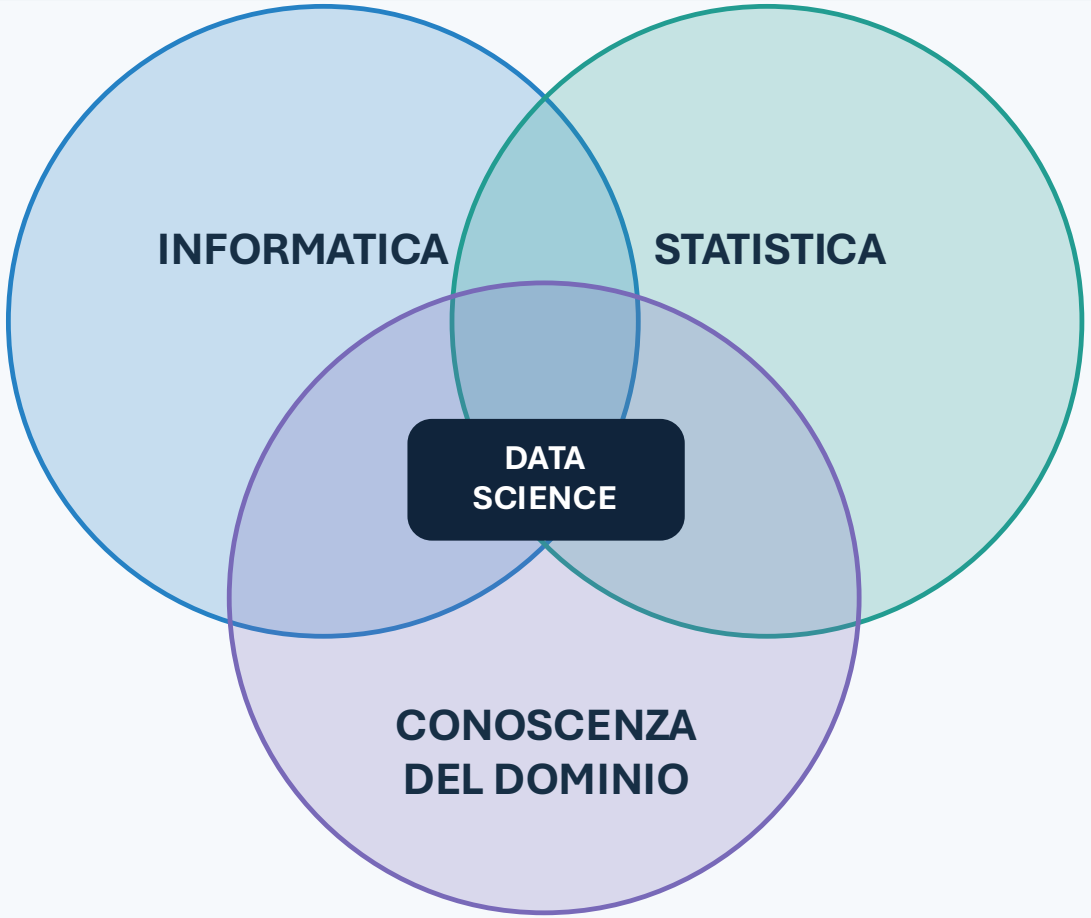
Testi, foto, audio, video

Es.: recensioni e post

Le fonti possono includere siti web, social media, piattaforme, sondaggi e open data.

Che cos’è la Data Science?

Un approccio interdisciplinare per estrarre conoscenza dai dati.



Il processo di Data Science

Un percorso iterativo: dai problemi ai risultati utilizzabili.

01 **Definire il problema** → 02 **Raccogliere i dati** → 03 **Preparare i dati**

04 **Analizzare e modellare** → 05 **Interpretare i risultati** → 06 **Comunicare e decidere**

↺ I risultati possono portare a ridefinire la domanda iniziale.

Le quattro tipologie di Analytics

Quattro domande complementari.

DESCRIPTIVE

Che cosa è successo?

Visualizzazioni, interazioni, copertura

DIAGNOSTIC

Perché è successo?

Cause possibili del calo di engagement

PREDICTIVE

Che cosa potrebbe succedere?

Probabilità di raggiungere un obiettivo

PRESCRIPTIVE

Che cosa conviene fare?

Scelte per migliorare una campagna

Big Data, Data Science e Big Data Analytics

Tre concetti connessi, ma non intercambiabili.

CONCETTO	FOCUS	DOMANDA GUIDA
BIG DATA	Caratteristiche dei dati e sfide tecnologiche	Con quali dati abbiamo a che fare?
DATA SCIENCE	Metodi e processi per ricavare conoscenza	Come possiamo ottenere conoscenza?
BIG DATA ANALYTICS	Tecniche per analizzare dati complessi o su larga scala	Che cosa possiamo scoprire?

La Data Science può operare anche su piccoli dataset. Non tutte le analisi richiedono tecnologie Big Data.

Caso di studio: una campagna Instagram

Scenario didattico: promuovere l’offerta formativa di un’università.

120.000
Visualizzazioni

8.400
Interazioni

2.100
Clic al sito

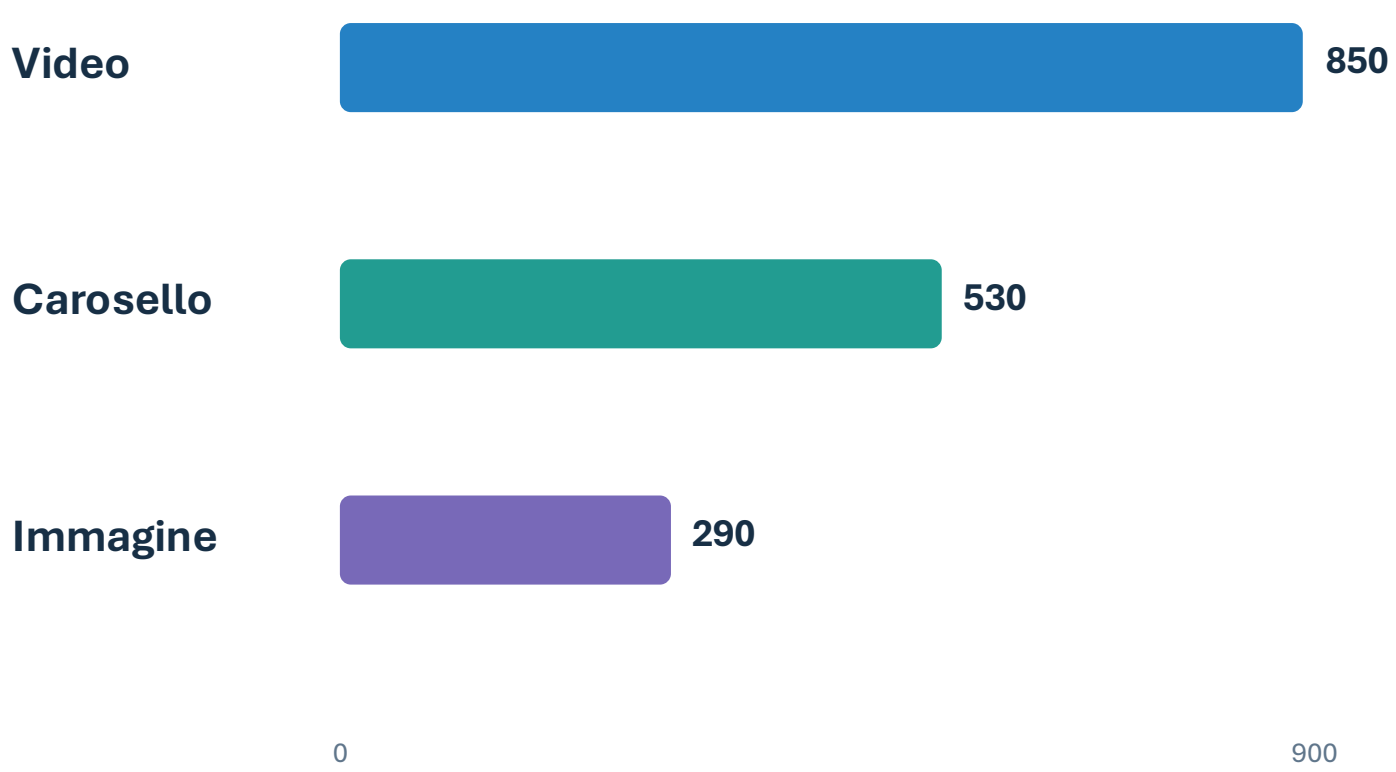
Domanda di analisi

Quali contenuti favoriscono davvero l’interesse verso i corsi di laurea?

Dati simulati per uso didattico.

Confrontiamo i formati dei contenuti

Interazioni medie per post — dati simulati.



Cosa possiamo concludere?

- I video mostrano più interazioni in media.
- Non sappiamo se il formato sia la causa.
- Servono altre variabili e confronti.

Grandi dati ≠ conoscenza affidabile

Qualità, rappresentatività ed etica sono centrali.

PRIVACY

Quali dati possiamo raccogliere e con quali basi giuridiche?

BIAS

Il campione rappresenta davvero il pubblico di interesse?

QUALITÀ

I dati contengono errori, duplicati o attività di bot?

INTERPRETAZIONE

Correlazione significa causalità? No.

Attività in aula: progettare un’analisi

Lavoro in piccoli gruppi · 10–15 minuti

Scenario: un’organizzazione vuole capire come viene percepita online.

1. Quali dati raccoglieresti?
2. Quale domanda vorresti risolvere?
3. Che tipo di Analytics useresti?
4. Quali limiti e rischi considereresti?

IL MESSAGGIO DA PORTARE A CASA

Il valore dei dati non risiede soltanto nella loro quantità.

Sta nella capacità di porre le domande giuste, analizzarli criticamente e trasformare i risultati in decisioni consapevoli.