



DIPARTIMENTO  
**MATEMATICA**

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Esame di laurea - 20 Settembre 2024

# **DATA VISUALIZATION E INFOGRAFICHE: STUDIO E APPLICAZIONE ATTRAVERSO D3.JS**

Laureando: Jessica Carretta  
Relatore: Prof. Lamberto Ballan

# L'AZIENDA: ZUCCHETTI S.P.A.



Fondata nel **1978**



Presente in **15** paesi,  
in Italia con **185** sedi



Oltre **9000** dipendenti



Oltre **700'000** clienti

Avendo come **mission**:

*‘supportare le aziende nel migliorare la loro competitività ed efficienza operativa attraverso soluzioni innovative e di qualità.’*



# LA PROPOSTA DI STAGE

- » **Esplorare il campo della data visualization e delle infografiche** per rappresentare efficacemente i dati
- » Concretizzare quanto studiato **creando un'infografica web**

## Obiettivi legati alla ricerca

- Comprendere i principi fondamentali della Data Visualization
- Selezionare il grafico più adatto per tipo di struttura dati e obiettivo della visualizzazione

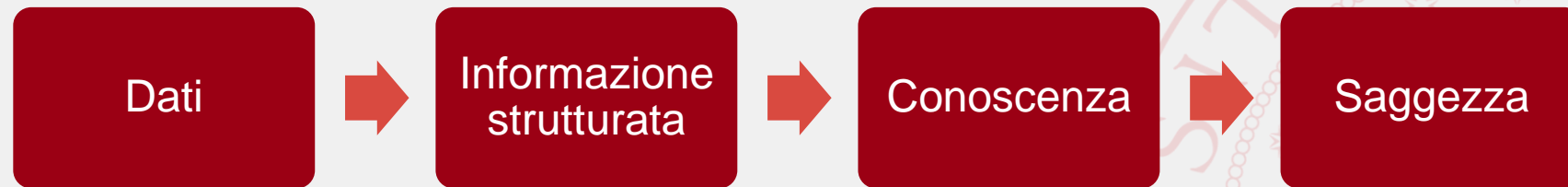
## Obiettivi legati alla pratica

- Realizzare grafici informativi e interattivi utilizzando D3.js
- Realizzare un'infografica web funzionante in HTML, CSS e JS che visualizzi i risultati degli algoritmi di Intelligenza Artificiale e Machine Learning
- Creare un template dell'infografica
- Individuare dei metodi di validazione dell'infografica\*

*\*Facoltativamente, implementare tali metodi di validazione, e.g. sviluppando infografiche alternative o implementando un chatbot con LLM che interagisca con l'utente sull'infografica.*

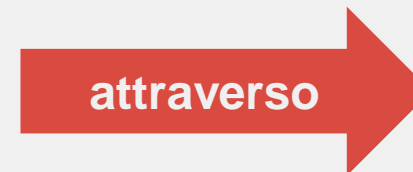
# LA DATA VISUALIZATION

Insieme dei meccanismi di traduzione da dati a rappresentazione grafica aventi lo scopo di **amplificare la cognizione dei dati** stessi



La data visualization permette di:

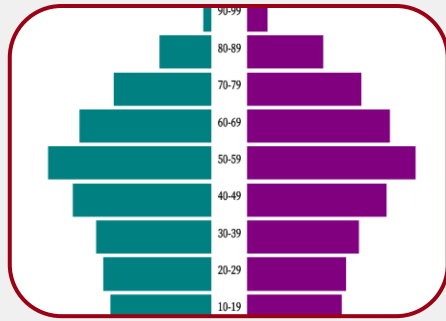
- » **Rendere i dati più comprensibili e memorabili**
- » **Facilitare nuove scoperte e individuare trend**
- » **Visualizzare velocemente relazioni e regolarità nei dati**
- » **Rendere più consapevole la presa di decisioni**



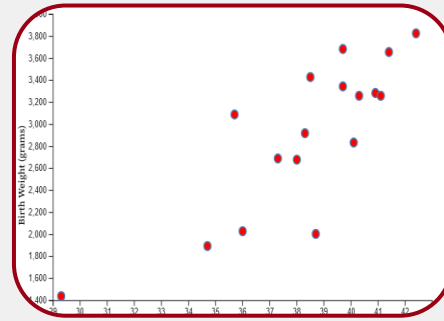
**Grafici, diagrammi e mappe**

# METODI DI CLASSIFICAZIONE DEI GRAFICI

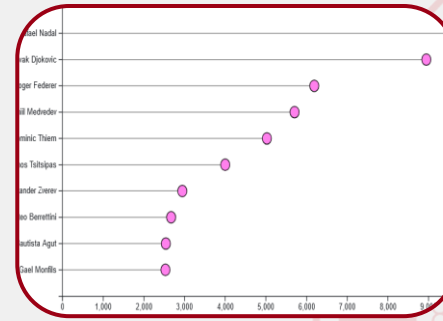
» Classificazione per **obiettivo della visualizzazione**



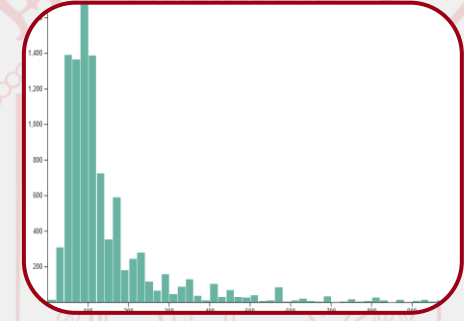
Divergenza



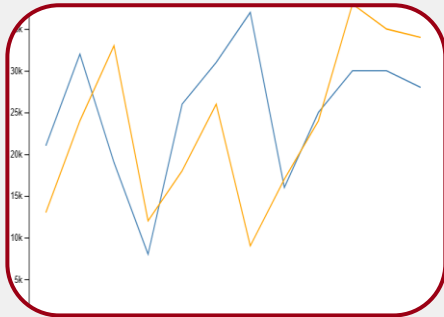
Correlazione



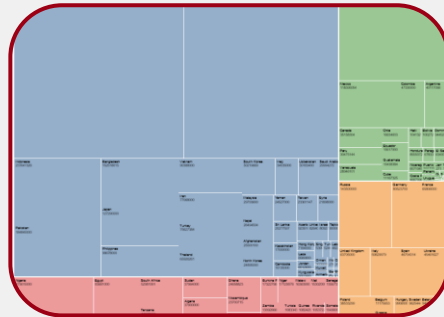
Ranking



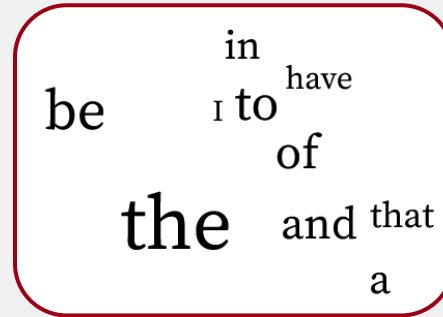
Distribuzione



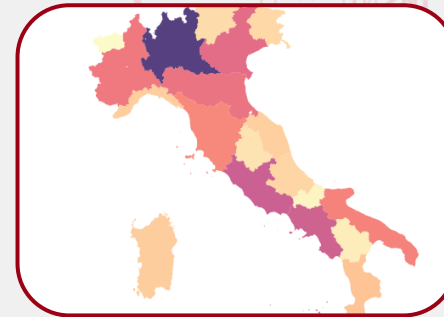
Cambiamento  
nel tempo



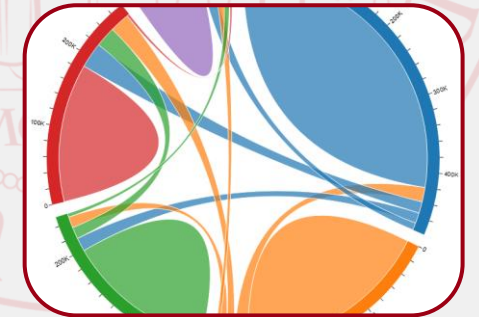
Composizione



Grandezze



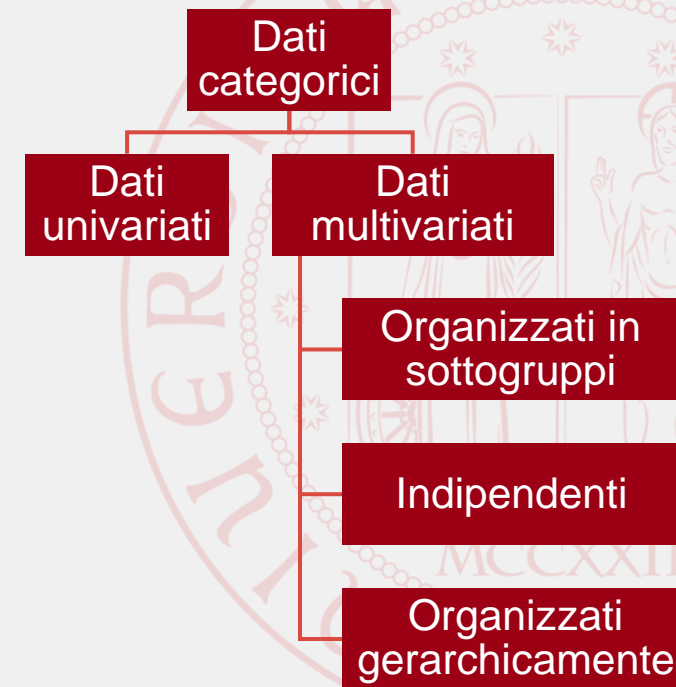
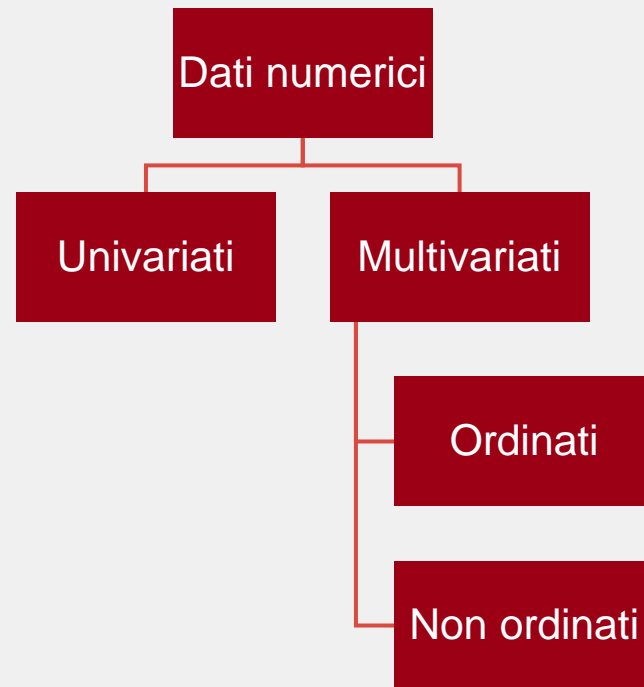
Spazi



Flussi

# METODI DI CLASSIFICAZIONE DEI GRAFICI

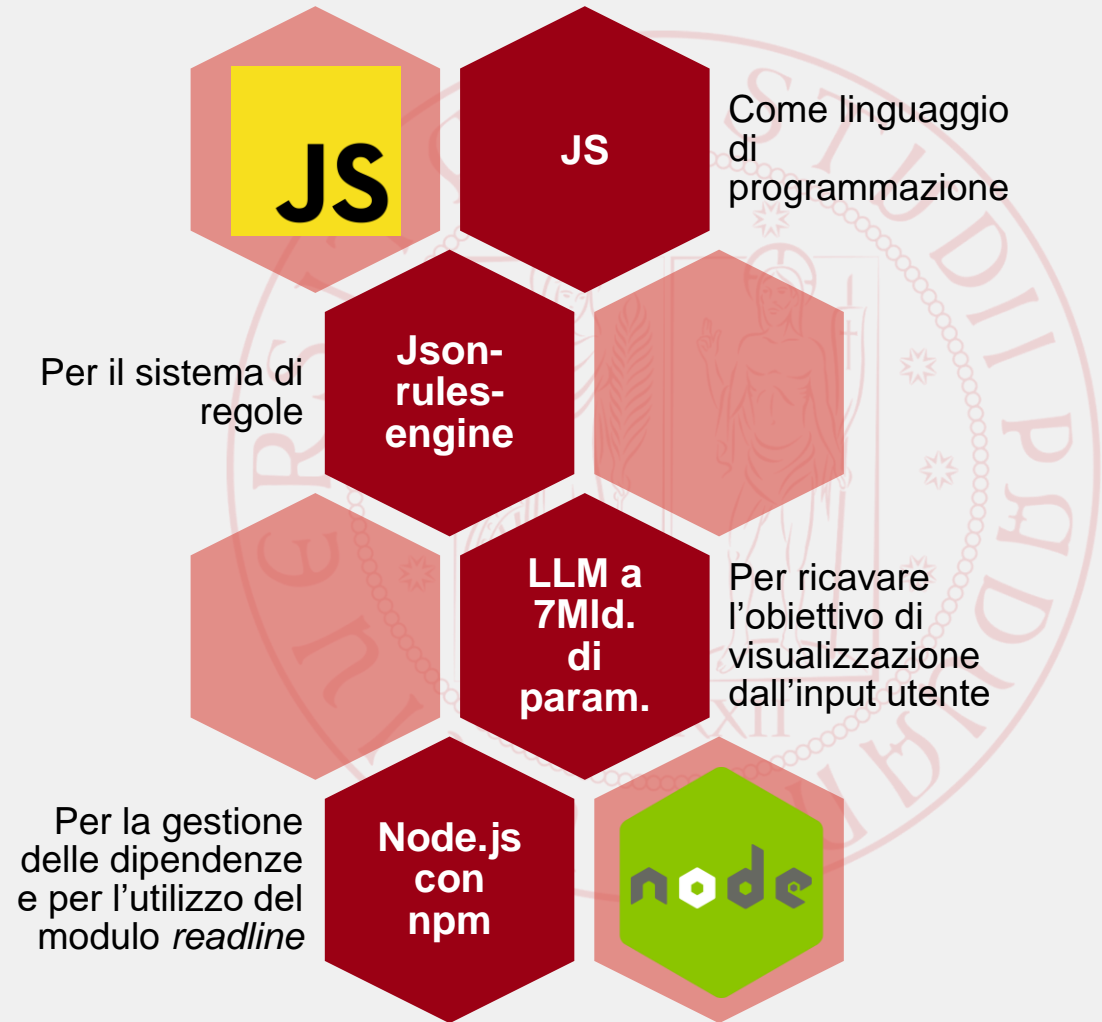
» Classificazione per **tipo di dato**



Con **dati misti con variabili categoriche organizzate in sottogruppi o gerarchicamente** si ha un'ulteriore distinzione in base al **numero di occorrenze** (singola o multiple)

# CHART-CHOOSER

- » **Strumento** prototipale **per la scelta del grafico** più adatto al specifico caso
- » Si compone di tre parti:
  - **Richiede all'utente** le informazioni sul **tipo di dato**
  - Identifica l'**obiettivo della visualizzazione** attraverso **LLM** (Large Language Model)
  - **Combina** i metodi di **classificazione** visti attraverso un **sistema di regole** a priorità





# LE INFOGRAFICHE

**Forma di racconto** che espone graficamente i punti salienti individuati in uno o più dataset a seconda anche dei destinatari della comunicazione

Le infografiche permettono di:

- » **Semplificare** notevolmente **la comprensione dei dati**
- » **Dare** una **visione d'insieme** del processo-oggetto analizzato
- » **Stimolare la curiosità** per ulteriori ricerche

attraverso

**Molteplici tecniche di comunicazione visiva**  
(e.g. grafici, illustrazioni)



# CLASSIFICAZIONE DELLE INFOGRAFICHE

Classificazione in base **all'output visivo**:

- » **Infografiche statiche** (e.g. infografiche dei giornali)
- » **Infografiche animate** (e.g. presentazioni multimediali)
- » **Infografiche interattive** (e.g. infografiche web)

Classificazione in base **all'obiettivo di visualizzazione**:



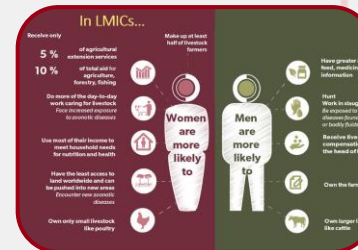
Infografiche  
basate su  
processi  
ordinati



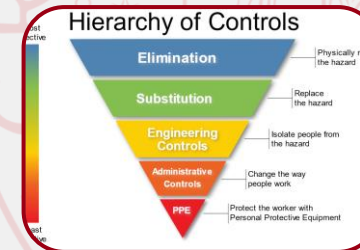
Infografiche a  
mo' di lista



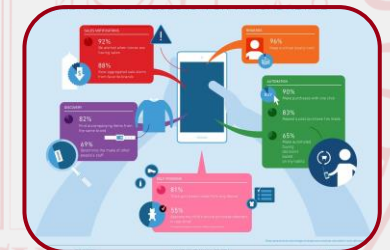
Infografiche  
basate sul  
tempo



Infografiche  
basate sul  
confronto



Infografiche  
basate su una  
gerarchia



Infografiche  
basate su  
informazioni  
correlate poste  
allo stesso  
livello

# INFOGRAPHIC-HELPER

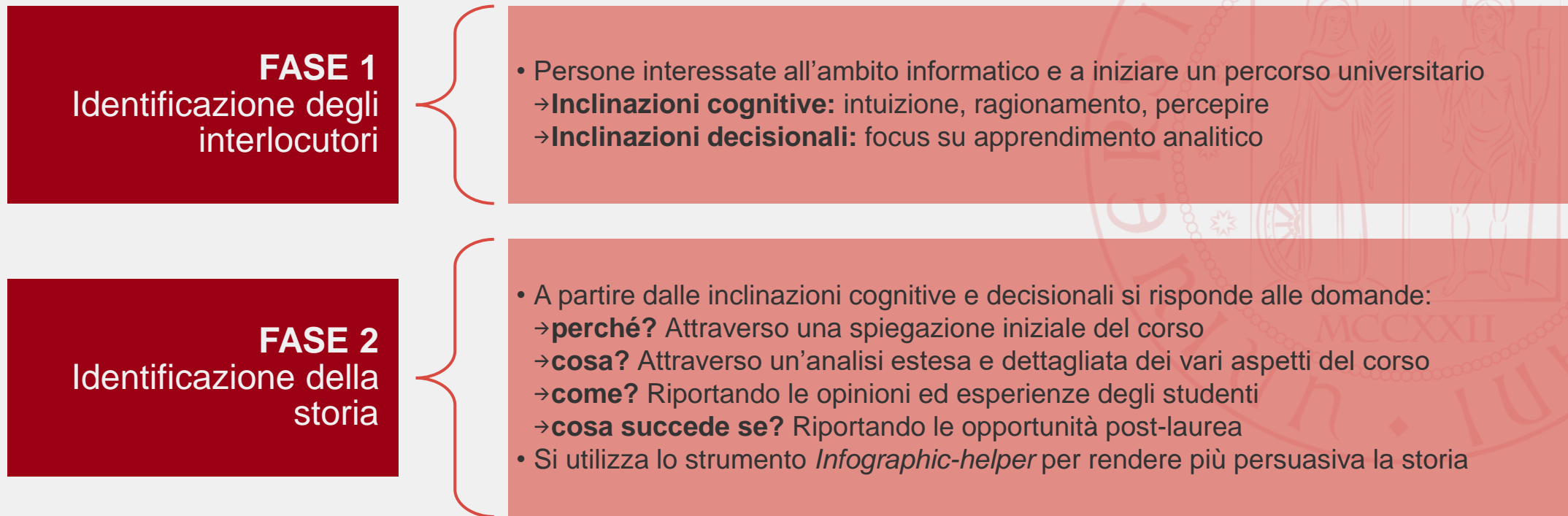
- » **Strumento** prototipale che **assiste l'utente nella creazione della storia** dell'infografica **e nella definizione della** sua futura **struttura grafica**
- » L'**utente** fornisce una **possibile storia** la quale viene **elaborata da un LLM** (Large Language Model), restituendo:
  - **Gli elementi che possono rendere la narrativa persuasiva** (i.e. elementi di *ethos*, *pathos* e *logos*)
  - **Gli elementi 'estranei'**, i.e. scollegati dal tema principale
  - **Le diverse parti** in cui è possibile dividere la storia
  - **L'obiettivo di visualizzazione** dell'infografica



Data visualization e infografiche:  
studio e applicazione attraverso D3.js

# ESEMPIO PRATICO: PROGETTAZIONE FASI 1-2

- » Infografica sul **Corso di Laurea di Informatica dell'Università degli Studi di Padova**
- » Per la progettazione si segue l'**Infomodel**:



# ESEMPIO PRATICO: PROGETTAZIONE FASI 3-5

## FASE 3 Identificazione del dataset

- Dati forniti dai report dell'**Università** e dalle statistiche di **Almalaurea**
- Dati ricavati manualmente dalla **pagina ufficiale del CdL**

## FASE 4 Analisi dei dati

- Si effettua il parsing dei dati
- Si analizzano i dati con *Chart-chooser* per la scelta del grafico

## FASE 5 Rappresentazione della storia

- Dall'obiettivo generale dell'infografica si intuisce il layout a **spiral**, che allinea simmetricamente elementi a destra e sinistra
- Si ricava la palette di **colori dal logo dell'Università**
- Si utilizzano caratteri **sans-serif**, più leggibili
- Si dà maggiore importanza agli **elementi di risposta alla domanda 'cosa?'**, realizzandoli **di una grandezza maggiore**
- Si inseriscono elementi grafici quali **icone**, piuttosto che immagini realistiche
- Si inseriscono **blocchi di testo coincisi** divisi tra le varie sezioni
- Si inseriscono elementi di interattività (**pan, zoom, filtro, ricerca, piccole animazioni**)

Data visualization e infografiche:  
studio e applicazione attraverso D3.js

# ESEMPIO PRATICO: LE TECNOLOGIE

## Premessa:

*Esistono due versioni dell'esempio che si distinguono in base all'implementazione del chatbot:*

- » **Versione base:** l'algoritmo del chatbot è implementato interamente in JavaScript
- » **Versione integrata:** integra del codice prodotto dal collega Fabio Meneghini in Python e SQL per realizzare una ricerca ibrida

### Struttura, stile e interattività

HTML

CSS

JavaScript

### Data visualization

D3.js (con d3-sankey e d3-legend)

Venn.js

Leaflet (con  
OverlappingMarkerSpiderfier)

### Chatbot e altro

Per *versione base*:

→ **Node.js** (con http-server)

Per *versione integrata*:

→ **Python** (con psycpg2, txtai,  
flask, flask-CORS)

→ **PostgreSQL** (con pgvector)

→ **pgAdmin**

Data visualization e infografiche:  
studio e applicazione attraverso D3.js

# ESEMPIO PRATICO: IMPLEMENTAZIONE

## Struttura generale

- » `<head>` contenente i metadati del documento
- » `<body>` racchiude il contenuto visivo dell'infografica diviso in `<div>` per ciascuna sezione
  - Si utilizza un'apposita classe per le sezioni interne per strutturare il layout a 'spiral'

## Data visualization

- » Si utilizzano **funzioni di D3.js** (`d3.json()`, `d3.dsv()` e `d3.text()`) **per leggere i dati**
- » Si utilizzano dei **contenitori `<div>` con delle classi apposite per creare contemporaneamente gli SVG** con diverse dimensioni in base all'importanza della visualizzazione

## Chatbot

- » Implementato come `<form>` che all'invio richiama funzioni diverse in base alla versione:
  - **VERSIONE BASE:** funzione JS che effettua una ricerca *term-based*
  - **VERSIONE INTEGRATA:** funzione Python che realizza una ricerca ibrida, combinando un modello di *sentence similarity* con il metodo Ranking BM25



# RENDICONTO FINALE DEI RISULTATI

## Risultato ottenuto

[https://drive.google.com/file/d/1iI7fJgplSvF9b5KNPc-F9b7tX1OI0Fz5/view?usp=drive\\_link](https://drive.google.com/file/d/1iI7fJgplSvF9b5KNPc-F9b7tX1OI0Fz5/view?usp=drive_link)

## Raggiungimento degli obiettivi

- » Tutti gli obiettivi **obbligatori** sono stati soddisfatti
- » Per quanto riguarda gli obiettivi **desiderabili**, questi sono stati **completati parzialmente**:
  - Non è stato soddisfatto l'obiettivo di creazione di un'infografica alternativa
  - L'obiettivo di creare un chatbot con LLM è stato realizzato parzialmente, in quanto il chatbot è stato creato però non utilizza LLM

