



A.D. 1308  
**unipg**  
DIPARTIMENTO  
DI INGEGNERIA

Tesina di progetto  
**Models and Algorithms for Data Visualization**

Corso di Laurea Magistrale in Ingegneria Informatica e Robotica

Curriculum Data Science

A.A. 2024-2025

DIPARTIMENTO DI INGEGNERIA

## MEAL GUIDE

docente

Prof. Giuseppe LIOTTA

studenti

369097	<b>Giulia</b>	<b>Giglioni</b>	giulia.giglioni@studenti.unipg.it
363551	<b>Andrea</b>	<b>Tomassoni</b>	andrea.tomassoni@studenti.unipg.it

# 0. Indice

<b>1</b>	<b>Introduzione</b>	<b>2</b>
1.1	Scopo del progetto di visualizzazione . . . . .	2
1.2	I dati da visualizzare . . . . .	2
1.3	Motivazioni alla base della visualizzazione . . . . .	2
1.4	Struttura delle sezioni successive . . . . .	3
<b>2</b>	<b>Raccolta dei Requisiti</b>	<b>4</b>
2.1	End-users . . . . .	5
2.2	Data Modeling . . . . .	6
2.2.1	Data extraction . . . . .	6
2.2.2	Data cleaning e trasformation . . . . .	6
2.2.3	Dataset types e attribute types . . . . .	7
2.3	Task Modeling . . . . .	8
<b>3</b>	<b>Design</b>	<b>10</b>
3.1	Visualization Design . . . . .	10
3.1.1	Choropleth Map e Bar Chart . . . . .	11
3.1.2	Bubble Chart . . . . .	12
3.1.3	Grafo Force-Directed . . . . .	12
3.1.4	Treemap . . . . .	13
3.2	Interaction Design . . . . .	13
3.2.1	Choropleth Map e Bar Chart . . . . .	14
3.2.2	Bubble Chart . . . . .	14
3.2.3	Grafo Force-Directed . . . . .	15
3.2.4	Treemap . . . . .	16
3.3	Architectural and Technological Choices . . . . .	16
3.4	Algorithm Engineering . . . . .	19
<b>4</b>	<b>Realization</b>	<b>22</b>

# 1. Introduzione

## 1.1 Scopo del progetto di visualizzazione

Questo progetto presenta un sistema di visualizzazione dati interattivo progettato per esplorare ed analizzare i modelli culinari globali.

Lo scopo del progetto di visualizzazione è consentire agli utenti di scoprire le relazioni tra regioni geografiche, categorie di pasti e utilizzo degli ingredienti nelle diverse cucine del mondo.

L'obiettivo principale è quello di rendere accessibili informazioni complesse attraverso rappresentazioni visuali intuitive, chiare e semplici, che permettono di esaminare connessioni tra regioni geografiche, tradizioni culinarie e utilizzo degli ingredienti nelle diverse cucine.

## 1.2 I dati da visualizzare

Le visualizzazioni si basano su un dataset completo di piatti e ingredienti provenienti da diversi paesi, che include informazioni dettagliate sulle origini dei piatti, gli ingredienti utilizzati e la loro categorizzazione.

Questi dati sono preziosi per comprendere le tradizioni culinarie, esplorare le preferenze alimentari culturali e identificare modelli di utilizzo degli ingredienti nelle diverse regioni del mondo.

Visualizzando questi dati, miriamo a scoprire tendenze nella cucina globale che potrebbero non essere evidenti attraverso la sola analisi testuale.

## 1.3 Motivazioni alla base della visualizzazione

Il sistema di visualizzazione consente di:

- comprendere la distribuzione geografica delle tradizioni culinarie
- esplorare gli ingredienti comuni in diverse cucine
- comprendere la distribuzione dei piatti per categoria
- scoprire gli ingredienti più diffusi nella cucina mondiale
- scoprire dei piatti partendo dagli ingredienti

- comprendere le categorie caratteristiche delle varie tipologie di cucina
- scoprire i piatti tipici delle specifiche tradizioni culinarie
- esplorare gli ingredienti comuni a diversi piatti

La motivazione alla base della visualizzazione di questi dati deriva dalla natura universale del cibo come elemento culturale. Il cibo non serve solo come sostentamento, ma anche come riflesso della storia culturale, delle condizioni geografiche e delle pratiche sociali. Fornendo rappresentazioni visive intuitive di questi dati, il sistema mira a rendere questi modelli culturali accessibili e coinvolgenti per gli utenti.

## 1.4 Struttura delle sezioni successive

La tesina è organizzata come segue, descrivendo nel dettaglio gli step necessari alla creazione di un sistema di information-visualization:

- **Requirements Gathering:** descrizione del dataset e delle sue caratteristiche principali
  - **End-users:** identificazione degli utenti finali e delle loro necessità
  - **Data Modeling:** processi di estrazione, pulizia e trasformazione dei dati
  - **Task Modeling:** definizione dei compiti e degli obiettivi della visualizzazione
- **Design:** dettagli sulla progettazione del sistema, incluse le scelte relative alle visualizzazioni, le interazioni, l'architettura, le tecnologie e gli algoritmi utilizzati
- **Realization:** descrizione dell'implementazione del sistema finale e dell'interfaccia utente

Nelle sezioni seguenti esploreremo il processo di raccolta dei requisiti, descriveremo in dettaglio le decisioni di progettazione che hanno dato forma al sistema di visualizzazione e mostreremo l'implementazione finale delle nostre visualizzazioni interattive.

## 2. Raccolta dei Requisiti

Il dataset utilizzato in questo progetto è stato fornito dal Graph Drawing Contest 2019, una competizione annuale che si svolge nell'ambito della conferenza internazionale Graph Drawing & Network Visualization (GD).

Nell'edizione del 2019, il contest ha proposto come creative topics un dataset ispirato agli ingredienti dei piatti. Il dataset modella le occorrenze degli ingredienti nei piatti popolari di diversi paesi, fornendo così un'opportunità per analizzare le connessioni tra gli ingredienti utilizzati nei vari piatti dei diversi paesi.

### Descrizione del dataset

Il dataset fornito contiene tre file JSON.

Il file "ingredients" elenca 297 ingredienti, ad esempio "Beef", "Flour", "Red Wine", "Onion", "Carrot", "Thyme", "Mustard".

Il file "areas" contiene 11 aree (paesi) che sono popolari per i loro piatti in tutto il mondo:

- American
- British
- Chinese
- French
- Greek
- Indian
- Italian
- Japanese
- Mexican
- Spanish
- Thai

Infine, il file "meals" contiene 151 ricette. Ogni record contiene le seguenti proprietà:

- **category:** (opzionale) la categoria a cui appartiene il piatto
- **ingredients:** la lista degli ingredienti del piatto

- tutti gli ingredienti presenti in questo campo si trovano nel file "ingredients"
- **name:** il nome del piatto
- **area:** la nazione di provenienza del piatto
  - tutte le nazioni sono elencate nel file "areas"

Questo è un esempio di record contenuto nel file "meals":

```
Record file "meals"

"Chocolate Gateau": {
  "category": "Desert",
  "ingredients": [
    "Chocolate",
    "Butter",
    "Milk",
    "Egg",
    "Sugar",
    "Flour"
  ],
  "name": "Chocolate Gateau",
  "area": "French"
}
```

Il dataset è scaricabile dal sito del contest:

<https://mozart.diei.unipg.it/gdcontest/contest2019/data/meal-db.zip>.

## 2.1 End-users

Il sistema di visualizzazione è progettato per diversi gruppi di utenti:

- **chef e appassionati di cucina** che desiderano esplorare modelli di cucina globali e scoprire nuovi ingredienti da cucine diverse
- **ricercatori alimentari** che studiano le tendenze culinarie interculturali e l'utilizzo degli ingredienti
- **nutrizionisti e dietisti** che analizzano i modelli alimentari in diverse regioni
- **professori e studenti** che possono utilizzare il sistema come strumento didattico per approfondire concetti legati ai dati e alla cucina

Gli utenti hanno diverse necessità legate all'analisi e alla comprensione di questi dati:

- confronti tra paesi per analizzare la distribuzione delle ricette e degli ingredienti tra le varie nazioni
- esplorazione tendenze culturali per identificare gli ingredienti più utilizzati nei vari paesi

- approfondimento di cucina per comprendere l'utilizzo degli ingredienti nelle diverse cucine
- accesso intuitivo a dati complessi per semplificare la comprensione di dataset articolati attraverso rappresentazioni visive chiare ed interattive
- personalizzazione dell'analisi per selezionare ed ottenere informazioni specifiche in base ai propri interessi

## 2.2 Data Modeling

Vediamo le operazioni effettuate durante le fasi del processo di modellazione dei dati.

### 2.2.1 Data extraction

In questa fase convertiamo i dati disponibili in dati strutturati.

Il dataset scaricato è in formato JSON, un formato semi-strutturato e leggibile che consente di rappresentare dati complessi. Il dataset viene caricato ed analizzato per produrre dati strutturati.

Questa fase viene effettuata attraverso uno script Python `dataCleaning.py` attraverso il quale leggiamo e cariciamo il dataset originale utilizzando la libreria Python `json`, che consente di convertire il file in una struttura dati gestibile (dizionari e liste) per l'elaborazione successiva.

Abbiamo lavorato principalmente con due file:

- `meals.json`: contiene informazioni sui pasti, inclusi nome, categoria e ingredienti
- `ingredients.json`: contiene una lista di ingredienti disponibili

### 2.2.2 Data cleaning e trasformazione

In questa fase cerchiamo di rendere adatto il dataset all'analisi e alla visualizzazione.

Per garantire la qualità e l'omogeneità dei dati, è stato implementato un processo di pulizia e trasformazione attraverso lo script `dataCleaning.py`. Le principali operazioni eseguite sono state:

#### 1. Correzione di errori ortografici e tipografici:

- sostituzione di "Challots" con "Shallots"
- sostituzione di "Cacao" con "Cocoa"
- sostituzione di "Zucchini" con "Courgettes"

#### 2. Standardizzazione delle categorie:

- correzione della categoria "Desert" in "Dessert"

#### 3. Rimozione di elementi duplicati o ridondanti:

- eliminazione di voci duplicate dall'elenco degli ingredienti
- pulizia di stringhe con spazi superflui

#### 4. Gestione dei valori nulli:

- sono stati cercati eventuali valori nulli nei dati perché se presenti, devono essere gestiti consapevolmente, ma il dataset non presenta valori mancanti

Il codice è strutturato per caricare i file JSON originali, apportare le correzioni necessarie e salvare i risultati in nuovi file (`meals_corrected.json` e `ingredients_corrected.json`), preservando così i dati originali.

Dopo la pulizia dei dati, è stata implementata una categorizzazione degli ingredienti tramite lo script `ingredientsCategorization.py`. Lo scopo era organizzare gli ingredienti in gruppi semanticamente coerenti per facilitare l'analisi e la visualizzazione.

Le categorie definite sono:

- Meat (Carne)
- Fish and Seafood (Pesce e Frutti di mare)
- Dairy (Latticini)
- Vegetables (Verdure)
- Fruits (Frutta)
- Spices and Herbs (Spezie ed Erbe)
- Grains and Pasta (Cereali e Pasta)
- Sauces and Condiments (Salse e Condimenti)
- Nuts and Seeds (Noci e Semi)
- Others (Altri)

Ogni ingrediente è stato assegnato a una categoria appropriata, creando un nuovo file JSON strutturato (`categorized_ingredients.json`) che associa ciascun ingrediente alla sua categoria.

### 2.2.3 Dataset types e attribute types

I dataset risultanti dopo le operazioni di pulizia e trasformazione presentano le seguenti strutture:

#### 1. Dataset dei piatti (`meals_corrected.json`):

- **Tipo:** tabella
- **Items:** piatti culinari
- **Attributi:**
  - `category` (categorico): categoria del piatto (es. "Chicken", "Dessert")
  - `ingredients` (categorico): lista di ingredienti necessari per il piatto
  - `name` (categorico): nome del piatto
  - `area` (categorico): nazione di cui fa parte il piatto



2. **Dataset degli ingredienti (ingredients\_corrected.json):**
  - **Tipo:** insieme
  - **Items:** ingredienti
3. **Dataset dei paesi (areas.json):**
  - **Tipo:** lista
  - **Items:** paesi
4. **Dataset degli ingredienti categorizzati (categorized\_ingredients.json):**
  - **Tipo:** tabella
  - **Items:** ingredienti
  - **Attributi:**
    - ingredient (categorico): nome dell'ingrediente
    - category (categorico): categoria a cui appartiene l'ingrediente
5. **Dataset geospaziale (world.geojson):**
  - **Tipo:** geometria
  - **Items:** paesi
  - **Posizioni:** coordinate geografiche dei paesi

## 2.3 Task Modeling

La fase di Task Modeling definisce formalmente i task che un utente può svolgere attraverso le visualizzazioni sviluppate.

Questi task sono stati identificati e formulati come coppie {action, target}.

Il sistema di visualizzazione supporta i seguenti task primari:

- {Explore, Distribuzione geografica dei piatti} - L'utente esplora la mappa interattiva per scoprire la distribuzione dei piatti nei diversi paesi
- {Lookup, Ingredienti di un paese} - L'utente clicca su un paese per vedere gli ingredienti più utilizzati nei suoi piatti
- {Compare, Numero dei piatti nei diversi paesi} - L'utente confronta la prevalenza dei piatti documentati nei diversi continenti
- {Discover, Distribuzione dei piatti} - L'utente osserva la mappa per inferire quali paesi abbiano più piatti tipici (colore più intenso) e quali meno
- {Locate, Piatti di una cucina nota} - L'utente cerca nella mappa un paese noto per la sua cucina per analizzarne la quantità dei piatti e i principali ingredienti
- {Explore, Ingredienti e categorie} - L'utente esplora il treemap per scoprire quali ingredienti sono più usati e come sono distribuiti nelle varie categorie

- {Discover, Distribuzione categorie} - L'utente osserva la dimensione e il colore dei rettangoli per inferire quali categorie hanno più ingredienti rappresentati
- {Browse, Categorie da legenda} - L'utente interagisce con la legenda per evidenziare una specifica categoria e analizzarne visivamente la distribuzione e presenza nel treemap
- {Identify, Occorrenze di un ingrediente} - L'utente passa il mouse su un ingrediente per leggere nel tooltip il numero di volte in cui è presente nei piatti
- {Compare, Numero di piatti delle diverse categoria} - L'utente confronta il numero di piatti documentati per categoria
- {Summarize/Discover , Distribuzione categorie e Piatti} - L'utente osserva la dimensione dei cerchi per scoprire quali categorie hanno più piatti documentati
- {Browse, Categorie piatti e Aree} - L'utente può filtrare per area le categorie e i piatti
- {Browse, Categorie piatti e Piatti} - L'utente può filtrare per categoria i piatti
- {Discover, Ingrediente e Categorie Ingredienti} - L'utente può esplorare gli ingredienti divisi per categoria e scoprire quali i piatti utilizzano gli ingredienti selezionati o parte di essi
- {Explore, Piatti e Ingredienti} - L'utente può esplorare i piatti visibili e visualizzare quali ingredienti sono stati selezionati e quali no
- {Summarize, Piatti e Ingredienti} - L'utente può identificare l'ingrediente più utilizzato nei piatti visibili in base alla dimensione del nodo ingrediente

## 3. Design

Il processo di progettazione rappresenta una fase cruciale nella creazione del sistema di visualizzazione, in cui si definiscono le modalità con cui i dati saranno rappresentati, le interazioni offerte agli utenti, le scelte tecnologiche e architettoniche, e gli algoritmi utilizzati per implementare le funzionalità previste.

Questa sezione descrive le scelte progettuali adottate, focalizzandosi su quattro aree principali:

1. **Visualization Design:** definisce come i dati saranno visualizzati, specificando marks e channels, basandosi sui risultati della fase di Data Modeling. La semantica associata ad ogni channel viene definita per garantire coerenza e significato. Successivamente si descrivono gli idiomi visuali scelti e le motivazioni delle scelte effettuate.
2. **Interaction Design:** descrive le modalità di interazione offerte agli utenti che consentono di esplorare e analizzare i dati in modo dinamico e personalizzato.
3. **Architectural and Technological Choices:** illustra l'architettura del sistema, le tecnologie adottate e i criteri di progettazione scelti per garantire scalabilità, efficienza e facilità d'uso.
4. **Algorithm Engineering:** analizza gli algoritmi implementati per elaborare, trasformare e rappresentare i dati in modo ottimale, supportando sia le funzionalità visive che quelle interattive.

Nel nostro progetto, sono stati adottati approcci diversi e ibridi per supportare più tipologie di visualizzazioni e rispondere ai task definiti, offrendo soluzioni specifiche per le diverse esigenze analitiche degli utenti.

### 3.1 Visualization Design

Questa sezione definisce come i dati saranno rappresentati visivamente, includendo la scelta dei marks e dei channels, con una semantica associata per garantire coerenza e significato. Successivamente vengono descritti gli idiomi visuali selezionati e le motivazioni delle scelte.

### 3.1.1 Choropleth Map e Bar Chart

#### Choropleth Map

**Mark:** interlocking marks. Ogni paese è rappresentato da poligono che rappresenta i confini geografici di una nazione definiti nel dataset in formato GeoJSON.

**Channels:**

- Posizione: definisce la collocazione geografica dei paesi sul piano della mappa
  - Semantica: rappresenta la corrispondenza tra la posizione geografica reale e la posizione visiva sulla mappa
- Forma: segue i confini dei paesi, riflettendo la loro forma geografica reale
  - Semantica: aiuta ad identificare le regioni tramite la loro forma unica
- Colore: utilizza una colormap sequenziale per rappresentare il valore dell'attributo quantitativo (il numero di piatti associati a ciascun paese)
  - Semantica: differenze di tonalità indicano variazioni nel valore quantitativo, facilitando il confronto tra paesi

La Choropleth Map è stata selezionata per fornire una visione globale della distribuzione geografica dei piatti di ciascun paese. Grazie all'utilizzo del colore tramite una colormap sequenziale, è possibile rappresentare visivamente l'intensità dell'attributo quantitativo, consentendo agli utenti di identificare rapidamente le nazioni con numerosi piatti.

Inoltre, la mappa permette di effettuare inferenze basate sulla vicinanza geografica, come il confronto tra continenti o regioni, facilitando l'esplorazione di pattern e relazioni spaziali. Questa rappresentazione si rivela particolarmente utile per il confronto immediato dei valori tra diversi paesi.

#### Bar Chart

**Mark:** barre (rettangoli).

**Channels:**

- Posizione verticale (asse y): definisce l'altezza della barra in funzione della frequenza dell'ingrediente
  - Semantica: rappresenta l'intensità dell'attributo quantitativo, ovvero il numero di volte in cui l'ingrediente appare nei piatti
- Posizione orizzontale (asse x): organizza le barre lungo l'asse orizzontale in base all'identità nominale degli ingredienti
  - Semantica: permette di distinguere visivamente ciascun ingrediente tramite la sua posizione

Il Bar Chart è stato selezionato per fornire una rappresentazione chiara e comparativa dei 5 ingredienti più utilizzati all'interno di un singolo paese. L'uso dell'altezza come canale

visivo principale facilita la comparazione immediata tra gli ingredienti, permettendo agli utenti di identificare rapidamente quelli predominanti.

L'interazione con la mappa consente di aggiornare dinamicamente il bar chart, rendendolo uno strumento efficace per esplorare le differenze culturali nelle abitudini culinarie dei diversi paesi. L'aggiunta di tooltip migliora ulteriormente la leggibilità dei dati, offrendo informazioni precise durante il confronto visivo.

### 3.1.2 Bubble Chart

**Mark:** Area (cerchi). Ogni cerchio rappresenta una categoria di piatti.

**Channels:**

- Dimensione dell'area: l'area di ciascun cerchio è proporzionale (scala di Flannery) al numero di piatti di ogni categoria
  - Semantica: la dimensione del cerchio comunica il valore quantitativo dell'attributo (frequenza), facilitando l'identificazione delle categorie con più piatti

Il Bubble Chart è stato scelto per mostrare la distribuzione e la quantità di elementi (in questo caso, piatti) raggruppati per categoria. Le dimensioni delle bolle rappresentano il numero di piatti in ogni categoria, rendendo immediata la comprensione delle proporzioni. L'aggiunta di filtri per area permette di andare a filtrare dinamicamente i piatti e le categorie in modo da poter osservare i dati relativi alle sole tipologie culinarie di interesse. Un tooltip permette anche di visualizzare il numero preciso di piatti per categoria.

### 3.1.3 Grafo Force-Directed

**Mark:** Punti (Nodi) e Linee (Archi). I nodi sono utilizzati per rappresentare piatti e ingredienti, mentre gli archi rappresentano l'utilizzo di un ingrediente in una ricetta.

**Channels (Punti):**

- Colore: i nodi ingrediente sono di colore verde mentre i nodi piatto sono rossi oppure grigi, in base al numero di nodi ingrediente a cui sono collegati: se è presente un solo nodo ingrediente, tutti i nodi piatto sono rossi; se sono presenti due o più nodi ingrediente, i nodi piatto sono rossi se collegati ad almeno due dei nodi ingrediente, grigi altrimenti
  - Semantica: il colore comunica l'attributo categorico del nodo, nel caso dei nodi piatto comunica anche un valore quantitativo legato al numero di ingredienti presenti
- Forma: i nodi ingrediente sono dei quadrati mentre i nodi piatto sono dei cerchi
  - Semantica: la forma del nodo rafforza la comunicazione dell'attributo categorico data dal colore
- Dimensione: la dimensione dei nodi piatto è proporzionale (scala di Flannery) al numero di nodi ingrediente collegati

- Semantica: la dimensione del nodo comunica il valore quantitativo dell'attributo (frequenza), facilitando l'identificazione dei piatti con più ingredienti tra quelli selezionati

In questa visualizzazione è stato utilizzato un grafo perchè permette di rappresentare al meglio le relazioni tra oggetti, in questo caso tra piatti e ingredienti. Il grafo viene costruito incrementalmente in base agli ingredienti selezionabili dal menù. Inoltre è presente una scheda in cui sono mostrati i piatti che comprendono tutti gli ingredienti selezionati o almeno due di essi (nel caso ci sia più di un ingrediente selezionato). È presente anche un tooltip che mostra l'etichetta del nodo nel grafo e l'elenco degli ingredienti se posizionato su un elemento della colonna dei piatti.

### 3.1.4 Treemap

**Mark:** Area (rettangoli). Ogni rettangolo rappresenta un ingrediente.

**Channels:**

- Dimensione dell'area: l'area di ciascun rettangolo è proporzionale al numero di volte in cui l'ingrediente appare nei piatti
  - Semantica: la dimensione del rettangolo comunica il valore quantitativo dell'attributo (frequenza), facilitando l'identificazione degli ingredienti più comuni
- Colore: ogni categoria ha un colore assegnato (blu, rosso, verde per le tre più frequenti)
  - Semantica: il colore identifica la categoria dell'ingrediente, rendendo immediatamente distinguibili i gruppi principali

Il treemap è stato scelto per rappresentare in modo compatto e gerarchico la distribuzione degli ingredienti, categorizzati secondo la loro natura (ad esempio, Dairy, Vegetables, Spices...). L'uso dell'area come encoding principale consente un confronto immediato tra la frequenza di utilizzo degli ingredienti, mentre la suddivisione per categorie agevola l'analisi semantica.

L'aggiunta dello slider permette di filtrare gli ingredienti meno frequenti, migliorando la leggibilità e mettendo in risalto quelli più significativi. La legenda interattiva consente inoltre un'esplorazione focalizzata sulle categorie, offrendo all'utente strumenti per un'analisi approfondita sia quantitativa che strutturale del dataset.

## 3.2 Interaction Design

La progettazione dell'interazione per il progetto è stata guidata dal *task modeling*, che ha evidenziato l'esigenza di operazioni centrate sull'esplorazione, il confronto e l'accesso ai dettagli. Sono stati implementati diversi meccanismi di interazione che supportano in modo diretto i task principali dell'utente. Di seguito si riportano la descrizione delle operazioni di interazione progettate e il paradigma di interazione adottato per ciascuna visualizzazione.

### 3.2.1 Choropleth Map e Bar Chart

#### Operazioni di interazione

- **Selezione:** l'utente può cliccare su un paese nella mappa per attivare la visualizzazione del bar chart associato. Questa operazione di *point-and-click* consente di selezionare un'istanza specifica (il paese) per accedere ad un sottoinsieme di dati correlati (ingredienti).
- **Esplorazione:** l'utente può navigare la mappa visualizzando diversi paesi e confrontando la quantità di piatti tramite la colorazione. È possibile eseguire *panning* e *zoom* attraverso pulsanti predefiniti (America, Europa, Asia).
- **Abstract zooming (Details-on-demand):** al passaggio del cursore sulle barre del bar chart, un tooltip mostra i dettagli del dato (nome dell'ingrediente e occorrenze), senza modificare la visualizzazione.
- **Zooming:** tramite i controlli specifici è possibile focalizzare l'attenzione su aree geografiche specifiche della mappa.

#### Paradigma di interazione

È stato adottato il paradigma **full view + details-on-demand**, in linea con il mantra di Ben Shneiderman: "*overview first, zoom and filter, then details-on-demand*".

La mappa fornisce una visione complessiva del dataset, mentre il bar chart si attiva in risposta all'interazione, permettendo un *drill-down* efficace nei dati.

L'approccio preserva il contesto e guida l'utente in modo fluido ed efficiente.

### 3.2.2 Bubble Chart

#### Operazioni di interazione

- **Filtraggio:** tramite un menù con *checkbox* si può filtrare per tipologia di cucina. Il bubble chart si aggiorna dinamicamente tramite un'operazione di *dynamic filtering*.
- **Selezione:** l'utente può cliccare su una categoria del bubble chart per aprire un menù che mostra tutti i piatti della categoria selezionata.
- **Abstract zooming (Details-on-demand):** al passaggio del cursore sui cerchi del bubble chart è possibile visualizzare il numero esatto di occorrenze per ogni categoria.
- **Esplorazione:** nel menù laterale, attivabile tramite la selezione di una categoria, è possibile eseguire un'operazione di esplorazione mediante scrolling per visualizzare i piatti della categoria selezionata, se questi sono più di quanti entrano nel display.

#### Paradigma di interazione

In questa visualizzazione è stato adottato il paradigma **full-view + vista coordinata**, combinando:

- *Dynamic Filtering* (filtraggio dinamico tramite checkbox)

- *Details-on-demand* (tooltip al passaggio del mouse)
- *Riconfigurazione* (aggiornamento automatico del layout)
- *Selezione + Details-on-demand* (apertura menu laterale con piatti della categoria selezionata)

### 3.2.3 Grafo Force-Directed

#### Operazioni di interazione

- **Selezione:** tramite un menù laterale è possibile selezionare gli ingredienti di interesse con cui si andrà a costruire il grafo. Inoltre, diventerà visibile un elenco di piatti che utilizzano gli ingredienti selezionati.
- **Esplorazione** tramite Panning: l'utente può eseguire un'operazione di esplorazione tramite panning per muovere la visualizzazione, concentrandosi su una specifica sezione del grafo.
- **Esplorazione** tramite Scrolling: nel caso in cui gli elementi del menù laterale, che viene utilizzato per mostrare i piatti relativi agli ingredienti selezionati, non possano entrare tutti nel display, sarà possibile esplorare tutto il menù tramite scrolling.
- **Zooming:** ogni volta che l'utente aggiunge dei nuovi nodi al grafo tramite l'operazione di selezione, il grafo si ridimensiona automaticamente in modo da poter essere tutto visibile all'interno della visualizzazione. Se lo desidera, l'utente può eseguire un'operazione di zooming controllata dal mouse per ingrandire una determinata porzione del grafo.
- **Abstract zooming (Details-on-demand):** quando il cursore si posiziona sopra un nodo, un tooltip mostra il tipo di nodo e il nome del piatto o ingrediente a cui fa riferimento. Se invece il cursore si posiziona sopra uno dei piatti del menù laterale, vengono mostrati gli ingredienti, mettendo in evidenza quelli presenti nel grafo.
- **Connessione** (Brushing e Linking): quando l'utente posiziona il cursore sopra un nodo di tipo piatto, vengono messi in evidenza tutti i nodi ingrediente che fanno parte di quel piatto; inoltre viene messo in evidenza anche il piatto presente nel menù laterale.
- **Rinconfigurazione:** il layout del grafo viene aggiornato automaticamente quando si va ad aggiungere/rimuovere un nodo oppure lo si sposta tramite drag and drop.

#### Paradigma di interazione

In questa visualizzazione è stato utilizzato il paradigma di **interazione incrementale**, in quanto si parte da una piccola porzione del grafo, in questo caso da un singolo nodo e tutti i suoi vicini, e si espande il grafo on-demand aggiungendo nodi e link, collegandoli a quelli già presenti. Viene inoltre utilizzato il paradigma di **viste coordinate** in quanto vengono combinate le seguenti operazioni:

- *Brushing e Linking* (interazione con il grafo per evidenziare i vicini di un nodo piatto e l'etichetta del piatto nel menù laterale)



- *Details-on-demand* (tooltip al passaggio del mouse che mostra informazioni sul nodo)
- *Riconfigurazione* (aggiornamento automatico del layout all'aumentare del numero dei nodi)

### 3.2.4 Treemap

#### Operazioni di interazione

- **Filtraggio:** tramite uno *slider*, l'utente imposta una soglia minima di occorrenze. Il treemap si aggiorna dinamicamente mostrando solo gli ingredienti più significativi. Questa è un'operazione di *dynamic filtering*, con risposta immediata alla variazione dei parametri.
- **Abstract zooming (Details-on-demand):** al passaggio del cursore su un rettangolo, viene mostrato un tooltip contenente il nome dell'ingrediente, la categoria e il numero di occorrenze. L'utente può così ottenere informazioni puntuali senza modificare il layout.
- **Connessione (Brushing e Linking):** interagendo con la legenda, l'utente può evidenziare una singola categoria di ingredienti. Tutti gli elementi appartenenti a quella categoria sono messi in evidenza nella visualizzazione principale, mentre gli altri vengono attenuati.
- **Riconfigurazione:** al variare del filtro o della categoria attiva, la disposizione spaziale dei rettangoli viene aggiornata automaticamente per riflettere la nuova struttura informativa.

#### Paradigma di interazione

È stato adottato il paradigma di **vista coordinata**, che combina:

- *Dynamic Filtering* (filtraggio dinamico tramite slider)
- *Brushing e Linking* (interazione con la legenda per evidenziare una categoria)
- *Details-on-demand* (tooltip al passaggio del mouse)
- *Riconfigurazione* (aggiornamento automatico del layout)

Questo paradigma supporta un'interazione esplorativa avanzata, in cui più viste e controlli interagiscono in modo coordinato per facilitare l'analisi dei dati sia a livello globale che locale.

## 3.3 Architectural and Technological Choices

Questa sezione descrive le fondamentali tecniche del sistema di visualizzazione, illustrando l'architettura progettuale, le tecnologie adottate e le soluzioni implementate per supportare le interazioni utente. L'obiettivo principale è garantire efficienza, scalabilità e usabilità, offrendo strumenti intuitivi per l'analisi dei dati. Questa parte esplora le scelte tecnologiche, i framework utilizzati e le funzionalità progettate per ottimizzare l'esperienza dell'utente e soddisfare i requisiti dei task definiti.

## Tipo di sistema progettato

Il sistema sviluppato è una **applicazione web interattiva**, progettata per consentire agli utenti di esplorare informazioni culinarie provenienti da diversi paesi.

Il sistema fornisce quattro visualizzazioni interattive:

- una **Choropleth Map** collegata ad un **Bar Chart**, per visualizzare la distribuzione geografica dei piatti e degli ingredienti principali per paese
- un **Bubble Chart**, per conoscere le distribuzioni delle categorie di piatti potendo filtrare per paese
- uno **Grafo Force-Directed**, che permette di selezionare gli ingredienti e conoscere i piatti che li utilizzano
- un **Treemap interattivo**, che mostra la distribuzione e la categorizzazione degli ingredienti in base alla frequenza di utilizzo

L'obiettivo è supportare l'analisi visiva e comparativa tra paesi, ingredienti e categorie, attraverso strumenti intuitivi e reattivi.

## Tecnologie utilizzate

Il sistema è stato sviluppato interamente per il web, utilizzando le seguenti tecnologie:

- **HTML, CSS, JavaScript**: per la struttura, lo stile e la logica di base delle visualizzazioni e dell'interfaccia
- **D3.js**: libreria JavaScript di basso livello, scelta per la sua elevata *espressività* nella creazione di visualizzazioni personalizzate. È stata utilizzata per tutte le visualizzazioni, gestendo l'intero flusso dei dati (caricamento, trasformazione e rendering)
- **Python**: viene utilizzato per elaborare il dataset originale, rimuovendo dati superflui e garantendo un formato coerente
- **GeoJSON**: formato per la definizione dei confini geografici dei paesi, utilizzato nella choropleth map
- **JSON**: utilizzato per strutturare i dati relativi a piatti e ingredienti, facilitandone l'integrazione e la manipolazione lato client
- **Node.js**: per simulare un server

## Tecniche e strumenti di interazione forniti all'utente

Le interazioni sono state progettate per supportare i principali task identificati nella fase di task modeling. Le tecniche per interagire con il sistema sono le seguenti:

### Per la Choropleth Map + Bar Chart:

- **Point-and-click selection:** selezione di un paese per attivare dinamicamente la visualizzazione degli ingredienti in un bar chart
- **Geographical zoom buttons:** pulsanti per eseguire *zoom* in regioni predefinite (es. Europa, Asia, America)
- **Tooltip informativi:** *details-on-demand* sia sulla mappa che sul bar chart

### Per il Bubble Chart:

- **Point-and-click Selection:** selezione di una categoria per attivare dinamicamente l'elenco dei piatti di quella categoria
- **Dynamic filtering:** filtraggio tramite *checkbox* per selezionare il paese di cui conoscere le distribuzioni delle varie categorie e di cui mostrare i piatti della categoria selezionata

### Per il Grafo Force-Directed:

- **Esplorazione tramite Panning:** l'utente può cliccare e trascinare tramite mouse muovendo la telecamera su una differente porzione di grafo
- **Selezione tramite Browsing:** il grafo viene costruito incrementalmente selezionando gli ingredienti suddivisi per categoria
- **Brushing + linking:** quando il cursore passa sopra un nodo piatto viene messo in evidenza insieme ai nodi ingrediente collegati e all'etichetta piatto presente sul menù laterale
- **Tooltip informativi:** *details-on-demand* sia sul grafo che sul menù laterale dei piatti
- **Mouse Zoom:** operazione di *zoom* eseguita tramite rotella del mouse
- **Riconfigurazione:** aggiornamento del layout del grafo nel caso di spostamento di un nodo tramite drag and drop

### Per il Treemap:

- **Dynamic filtering:** slider interattivo per selezionare la soglia minima di occorrenze degli ingredienti
- **Brushing + linking:** interazione con la legenda per evidenziare le categorie e mettere in risalto i corrispondenti rettangoli nel treemap
- **Tooltip al passaggio del mouse:** forniscono informazioni dettagliate su ciascun ingrediente (nome, categoria, frequenza)
- **Riconfigurazione automatica:** aggiornamento del layout in risposta a modifiche nei filtri o nella categoria selezionata

Gli strumenti forniti all'utente includono il mouse e la tastiera.

## 3.4 Algorithm Engineering

Per rappresentare i dati in modo chiaro e intuitivo, sono state sviluppate quattro visualizzazioni principali, ognuna delle quali utilizza algoritmi specifici per elaborare e presentare le informazioni in modo efficace. Di seguito, una panoramica di ogni visualizzazione e delle sue funzionalità principali.

### Choropleth Map e Bar Chart

#### Pre-elaborazione dati

I dati sui piatti sono stati caricati dal file `meals_corrected.json`.

Per ogni paese è stato calcolato il numero totale di piatti presenti nel dataset. Questo valore è stato utilizzato per assegnare l'intensità cromatica nella mappa.

È stato inoltre costruito un dizionario `ingredientsByCountry`, che per ciascun paese mappa memorizza la lista degli ingredienti usati nei piatti di quel paese e quante volte ciascun ingrediente compare nei piatti associati a quel paese.

#### Rendering della mappa

Il rendering della mappa è stato fatto usando Leaflet.js. Viene caricato il file `world.geojson` tramite `fetch` e, una volta ottenuti i dati, viene creato un layer GeoJSON.

I paesi sono colorati in base al numero di piatti associati, usando la funzione `getColor`.

#### Interazione e aggiornamento del Bar Chart

Al click su un paese, viene eseguita una funzione di `updateBarChart()`, che aggiorna dinamicamente il contenuto del bar chart.

Gli ingredienti vengono ordinati per frequenza decrescente e rappresentati con rettangoli SVG. L'altezza della barra è calcolata linearmente rispetto al massimo locale.

Viene inoltre gestita l'animazione di ingresso e il tooltip informativo al passaggio del mouse.

### Bubble Chart

#### Pre-elaborazione dati

I dati sono caricati dal file `meals_corrected.json` che contiene un record per ogni piatto contenente le informazioni sul nome del piatto, la categoria, gli ingredienti e il paese. Successivamente si calcola il numero di pasti per categoria, viene creata una lista di tutti i paesi presenti e creato un mappaggio dei pasti suddivisi per categoria.

#### Rendering

Vengono creati i filtri basandosi sulla lista dei paesi. Si crea una gerarchia in base alle occorrenze di una categoria e tramite la funzione `d3.pack()` si va a creare il bubble chart in cui il raggio delle bolle è proporzionale alle occorrenze della categoria che rappresenta (viene usata la *scala di Flannery*). Successivamente vengono definite delle forze per tenere le bolle vicine e non sovrapposte.

## Interazione

Cambiando i filtri, il chart si aggiorna dinamicamente. Cliccando su una delle bolle si visualizza un menù contenente i piatti della categoria selezionata, filtrati per i paesi selezionati nel menù di filtraggio. Cliccando fuori dalle bolle si nasconde la lista dei piatti. Passando con il cursore sopra una bolla è possibile vedere il numero esatto di occorrenze.

## Grafo Force-Directed

### Pre-elaborazione dati

I dati vengono caricati dai seguenti file:

- `categorized_ingredients.json`: contiene ingredienti e categoria
- `ingredients.json`: contiene i dettagli degli ingredienti
- `meals_corrected`: contiene i piatti e gli ingredienti associati

Per quanto riguarda il processamento dei dati, inizialmente si estraggono le categorie e si crea una lista di ingredienti categorizzati. Successivamente si costruisce una struttura `mealsData` che mappa ogni piatto ai suoi ingredienti e infine si inizializza la struttura del grafo nel modo seguente:

```
let graphData = {nodes: [], links: []}
```

### Rendering

Nella fase di rendering vengono creati tre componenti principali:

- Sidebar: contiene gli ingredienti suddivisi per categoria
- Ranking: contiene i piatti che fanno uso degli ingredienti selezionati
- SVG: in questa sezione viene disegnato e aggiornato dinamicamente il grafo

## Interazione

Selezionando un ingrediente dal menù dalla sidebar, viene aggiunto/rimosso dal grafo. Posizionando il cursore su un nodo o su una voce della colonna ranking vengono evidenziati i collegamenti e mostrati i tooltip informativi. Tramite *Drag and Drop* è possibile spostare la visualizzazione oppure un singolo nodo. È possibile zoommare tramite il mouse e la visualizzazione fa *auto-zoom* quando si aggiunge un nuovo nodo al grafo e questo esce dalla visualizzazione.

## Treemap

### Aggregazione e categorizzazione

All'avvio, viene caricato il file `categorized_ingredients.json`, che associa ogni ingrediente a una categoria, e il file `meals_corrected.json` che contiene l'elenco di tutti i piatti.

Costruiamo una mappa dove per ogni ingrediente viene calcolato il numero totale di occorrenze.

Una volta che ogni ingrediente ha la sua categoria e il suo numero di occorrenze, raggruppiamo gli ingredienti per categoria e contiamo quanti ingredienti appartengono a ciascuna categoria.

### Layout del Treemap

È stato utilizzato l'algoritmo **Squarified Treemap**, implementato tramite la funzione `d3.treemap()` della libreria D3.js. Questo algoritmo suddivide ricorsivamente lo spazio in rettangoli proporzionali al valore dei nodi, cercando di mantenere un aspect ratio vicino al quadrato, per favorire il confronto visivo.

Il layout viene ricalcolato ogni volta che cambia la soglia minima di occorrenze, impostata dallo slider.

### Gestione dell'interazione

Il filtro tramite slider attiva un ricalcolo del dataset filtrato e un aggiornamento completo del layout.

L'interazione con la legenda è realizzata tramite `mouseover` e `mouseout` events. Viene gestita la variazione dell'opacità dei rettangoli per evidenziare solo gli elementi appartenenti alla categoria selezionata.

I tooltip sono generati dinamicamente per mostrare nome, categoria e occorrenze di ciascun ingrediente al passaggio del mouse.

## 4. Realization

La fase di realizzazione ha portato allo sviluppo di un'applicazione web interattiva che consente di esplorare diversi aspetti dei dati culinari attraverso quattro visualizzazioni interattive. Il sistema è accessibile tramite una homepage principale da cui l'utente può scegliere quale visualizzazione esplorare. Tutte le componenti sono sviluppate in HTML, CSS e JavaScript, con il supporto della libreria `D3.js` per la gestione delle visualizzazioni grafiche.

### Struttura generale del sistema

All'avvio, l'utente accede a una pagina iniziale dal titolo *Meal Guide*, che funge da dashboard di navigazione per le quattro visualizzazioni offerte:

- **Visualization 1 – Global comparison and single nation trend**
- **Visualization 2 – Meals categorization**
- **Visualization 3 – Ingredient-Meal network explorer**
- **Visualization 4 – Ingredients usage Treemap**

Ogni sezione è rappresentata da una card grafica con immagine, titolo e pulsante “View Visualization” che reindirizza alla relativa pagina HTML.

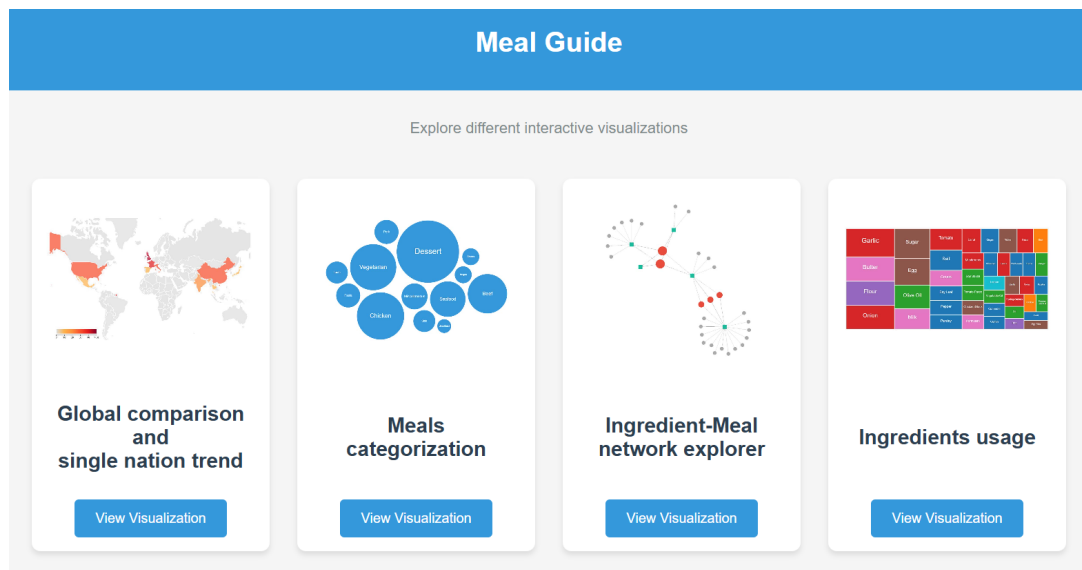


Figura 4.1: Homepage dell'applicazione con selezione delle quattro visualizzazioni.

## Visualizzazione 1 – Global comparison and single nation trend

Questa visualizzazione combina una **choropleth map** con un **bar chart**. L'utente può esplorare la distribuzione geografica dei piatti tipici nel mondo: ogni paese è colorato in funzione del numero di piatti presenti nel dataset, consentendo confronti visivi immediati. Cliccando su un paese, il sistema mostra i cinque ingredienti più utilizzati nella cucina di quella nazione tramite un bar chart interattivo. L'utente può così osservare, confrontare e analizzare le differenze tra le cucine dei vari paesi, supportato anche da controlli che permettono lo zoom diretto su specifiche aree geografiche (Europa, Asia, America). I tooltip attivati al passaggio del mouse forniscono dettagli puntuali sia sui paesi che sugli ingredienti.

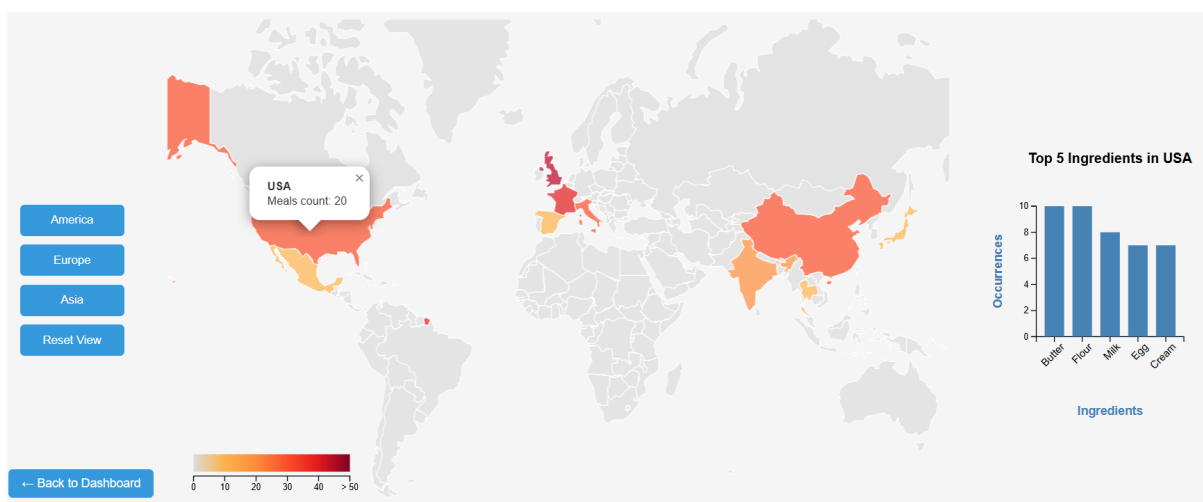


Figura 4.2: Choropleth map e bar chart degli ingredienti per nazione.



## Visualizzazione 2 – Meals categorization

Questa visualizzazione utilizza un **bubble chart** per esplorare la distribuzione delle categorie di piatti tipici. Ogni bolla rappresenta una categoria culinaria e la sua dimensione è proporzionale al numero di piatti presenti nel dataset per quella categoria, permettendo confronti visivi immediati tra le diverse tipologie di cucina. L'utente può filtrare le categorie visualizzate in base alla tipologia di cucina tramite un pannello di controllo laterale. Cliccando su una bolla, viene mostrata una lista dei piatti appartenenti a quella categoria in una colonna laterale, facilitando l'esplorazione dettagliata delle ricette. Tooltip interattivi forniscono informazioni aggiuntive sulle categorie al passaggio del mouse, mentre la visualizzazione si adatta dinamicamente ai filtri selezionati, supportando un'analisi approfondita delle tendenze culinarie globali.

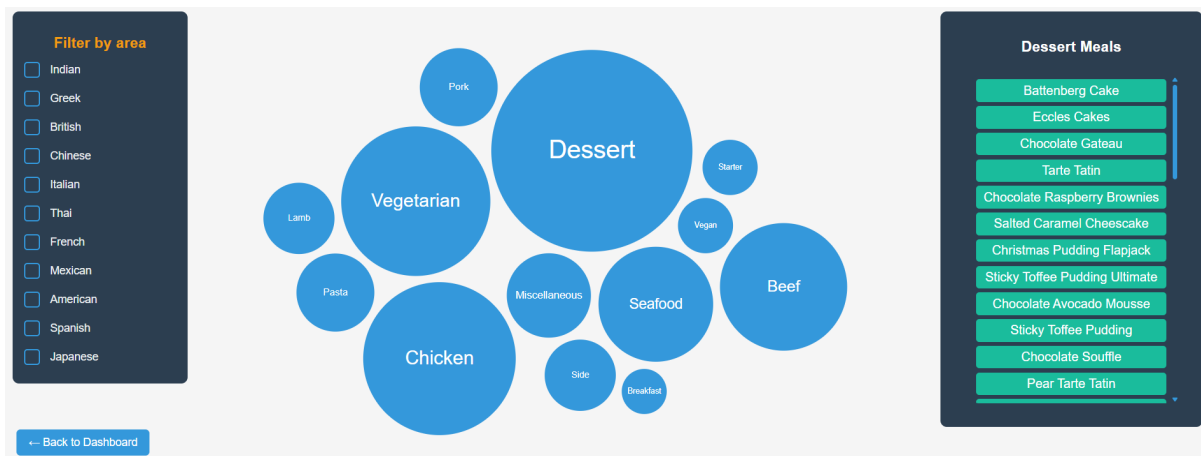


Figura 4.3: Bubble Chart per la visualizzazione delle categorie di piatti in base alla tipologia di cucina.

## Visualizzazione 3 – Ingredient-Meal network explorer

Questa visualizzazione utilizza un **grafo force-directed** che collega ingredienti e piatti tipici. Gli ingredienti sono rappresentati come rettangoli verdi, mentre i piatti sono cerchi rossi; un collegamento tra i due indica la presenza dell'ingrediente nella ricetta. L'utente può selezionare ingredienti dalla sidebar per aggiungerli alla rete e vedere immediatamente i piatti che li contengono, con la possibilità di esplorare le combinazioni e le sovrapposizioni tra ingredienti. Una colonna laterale mostra i piatti che contengono tutti o almeno due degli ingredienti selezionati, ordinati per rilevanza. Tooltip e highlight dinamici facilitano l'esplorazione delle connessioni, mentre controlli di zoom e drag permettono di navigare agevolmente nella rete.

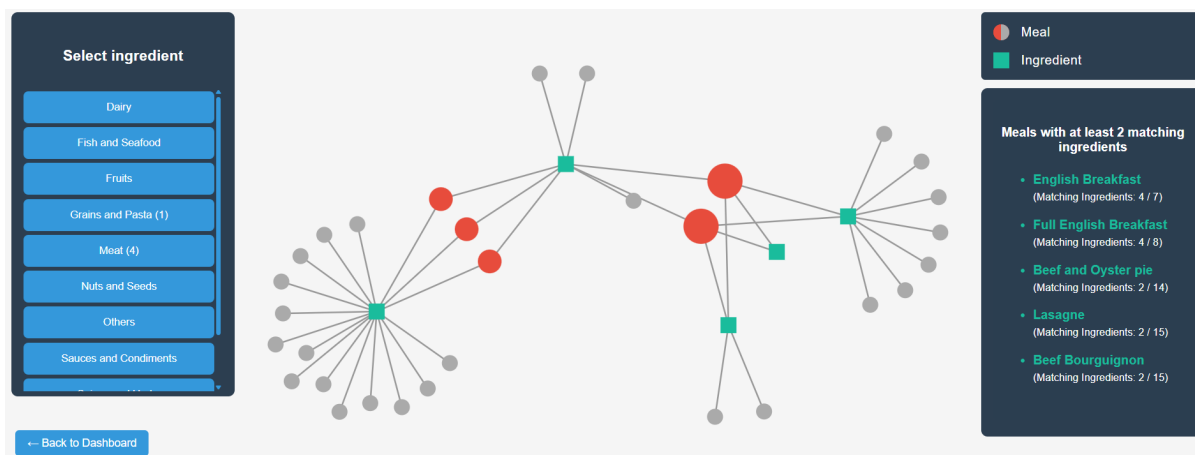


Figura 4.4: Grafo Force-Directed per la visualizzazione delle relazioni tra piatti e ingredienti.

## Visualizzazione 4 – Ingredients usage Treemap

Questa visualizzazione è costituita da un **treemap gerarchico**, che rappresenta gli ingredienti più utilizzati nel dataset, organizzati per categoria semantica (come ad esempio *Vegetables*, *Dairy*, *Spices*). La dimensione di ciascun rettangolo è proporzionale alla frequenza di utilizzo dell'ingrediente, mentre il colore ne identifica la categoria. L'interfaccia offre all'utente uno slider per filtrare dinamicamente gli ingredienti meno frequenti, concentrando l'attenzione solo su quelli rilevanti. Inoltre, grazie alla legenda interattiva, è possibile mettere in risalto una singola categoria, facilitando l'analisi semantica del contenuto. Infine, il passaggio del mouse su ciascun ingrediente attiva un tooltip informativo con nome, categoria e occorrenze.

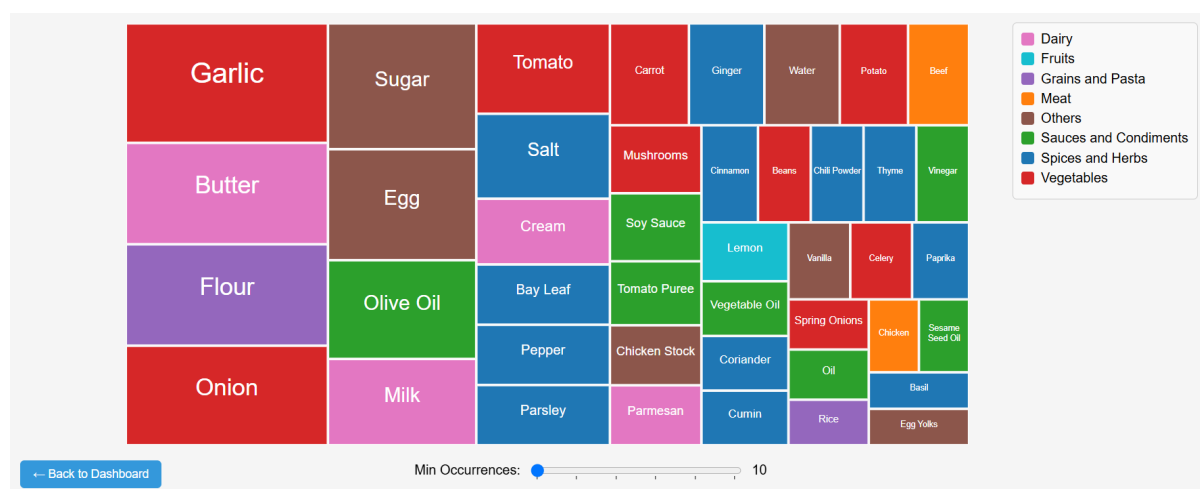


Figura 4.5: Treemap degli ingredienti con filtraggio e highlighting per categoria.