



# SpaghettiCode

spaghetti.code.g6@gmail.com

VERBALE ESTERNO  
2020-12-17

<b>Versione</b>	v1.0.0
<b>Approvazione</b>	Giorgia Paparazzo
<b>Redazione</b>	Masevski Martin
<b>Verifica</b>	Cognome Nome
<b>Uso</b>	Esterno
<b>Destinato a</b>	prof. Vardanega Tullio prof. Cardin Riccardo SpaghettiCode Zucchetti S.p.A.

## Descrizione

Riassunto dell'incontro realizzato dal gruppo SpaghettiCode tenutosi il 17 Dicembre 2020 in forma di meeting online con Piccoli Gregorio.



## Registro delle modifiche

Versione	Nominativo	Ruolo	Data	Descrizione
v0.0.1	Masevski Martin	Amministratore	2020-12-17	Creazione del documento e prima stesura



## Indice

<b>1</b>	<b>Informazioni generali</b>	<b>1</b>
1.1	Informazioni incontro . . . . .	1
1.2	Riferimenti . . . . .	1
<b>2</b>	<b>Domande poste</b>	<b>1</b>
<b>3</b>	<b>Resoconto</b>	<b>2</b>
3.1	Gli utenti si devono poter registrare ed autenticare? . . . . .	2
3.2	Oltre ai file .csv dobbiamo prevedere altri formati obbligatori? . . . . .	2
3.3	Dobbiamo supportare determinate versioni di browser? . . . . .	2
3.4	Dobbiamo permettere la visualizzazione di più grafici contemporaneamente? . . . . .	2
3.5	Dobbiamo aiutare l'utente a capire quello che sta vedendo? . . . . .	2
3.6	Deve essere disponibile per l'utente una breve guida introduttiva? . . . . .	2
3.7	Il progetto avrà un uso interno a scopo didattico o esterno? . . . . .	2
3.8	È desiderabile salvare il lavoro in corso oppure si comincia una nuova sessione ogni volta? . . . . .	2
3.9	È desiderabile prevedere un sistema di esportazione? . . . . .	2
3.10	I dati forniti saranno da filtrare prima di visualizzare i grafici oppure sono già pronti all'uso? . . . . .	2
3.11	È importante l'aspetto estetico del software? . . . . .	3
3.12	Avremo file o database da cui prendere i dati? . . . . .	3
3.13	Potremmo avere un esempio tipico di utilizzo del prodotto finale? . . . . .	3
3.14	Limite delle dimensioni: sono massimo 15 o almeno 15? . . . . .	3
3.15	Verrà fornita qualche tipo di formazione o è parte del nostro auto apprendimento? . . . . .	3
3.16	È possibile usare TypeScript o è un requisito JavaScript? . . . . .	3
3.17	In merito a ml.js di D3.js consigliate altre librerie? . . . . .	3
3.18	Ci sono requisiti particolari per il Front-End? . . . . .	3
3.19	Dobbiamo usare un grafico a seconda del tipo di dato? . . . . .	3
<b>4</b>	<b>Conclusione dell'incontro</b>	<b>4</b>



# 1 Informazioni generali

## 1.1 Informazioni incontro

- **Luogo:** Applicazione desktop SkypeG;
- **Data:** 2020-12-17;
- **Ora:** 10:15-11:30
- **Partecipanti:**
  - SpaghettiCode
  - Gruppo 5
  - Gruppo 10
  - Gruppo 15
  - Piccoli Gregorio

## 1.2 Riferimenti

- **link1:** <https://observablehq.com/@mbostock/the-wealth-health-of-nations>;
- **link1:** <https://github.com/mljs/distance>;
- **link1:** <https://cs.stanford.edu/people/karpathy/convnetjs>;
- **link1:** <https://github.com/karpathy/tsnejs>;
- **link1:** <https://github.com/PAIR-code/umap-js>.

# 2 Domande poste

Vengono riportate le domande che sono state poste nel corso del meeting:

- Gli utenti si devono poter registrare ed autenticare?;
- Oltre ai file .csv dobbiamo prevedere altri formati obbligatori?;
- Dobbiamo supportare determinate versioni di browser?;
- Dobbiamo permettere la visualizzazione di più grafici contemporaneamente?;
- Dobbiamo aiutare l'utente a capire quello che sta vedendo?;
- Deve essere disponibile per l'utente una breve guida introduttiva?;
- Il progetto avrà un uso interno a scopo didattico o esterno?;
- È desiderabile salvare il lavoro in corso oppure si comincia una nuova sessione ogni volta?;
- È desiderabile prevedere un sistema di esportazione?;
- I dati forniti saranno da filtrare prima di visualizzare i grafici oppure sono già pronti all'uso?;
- È importante l'aspetto estetico del software?;
- Avremo file o database da cui prendere i dati?;
- Potremmo avere un esempio tipico di utilizzo del prodotto finale?.



### **3 Resoconto**

#### **3.1 Gli utenti si devono poter registrare ed autenticare?**

No, siamo liberi di fare quello che riusciamo, non dobbiamo gestire la sicurezza. Le visualizzazioni dei grafici hanno priorità assoluta. Se vogliamo aggiungere l'autenticazione è una scelta nostra.

#### **3.2 Oltre ai file .csv dobbiamo prevedere altri formati obbligatori?**

In qualche modo il dato deve essere portato nell'ambiente client, così come richiesto nei requisiti. Il modo più naturale è avere un server, tipicamente i dati possono provenire da una query su database oppure da un file. Il formato .csv è quello più comune e più facile da ottenere. Si può anche fare l'upload diretto del file senza passare per il server. I dati possono essere estrapolati da un database SQL o NoSQL. I dati possono provenire da più fonti.

#### **3.3 Dobbiamo supportare determinate versioni di browser?**

Non è un requisito necessario. I browser che possiamo prendere in considerazione sono Firefox e Google Chrome. Sarebbe desiderabile poter usare il software anche da iPad e simili quindi si potrebbe ampliare lì, ma è un requisito non obbligatorio.

#### **3.4 Dobbiamo permettere la visualizzazione di più grafici contemporaneamente?**

La documentazione richiede almeno un grafico per volta. La possibilità di mettere insieme più grafici è cosa gradita. Poter mettere due grafici affiancati sarebbe carino.

#### **3.5 Dobbiamo aiutare l'utente a capire quello che sta vedendo?**

È un'idea interessante, se si riesce a far dire al sistema cosa vede.

#### **3.6 Deve essere disponibile per l'utente una breve guida introduttiva?**

Sì non fa mai male. I meno esperti potrebbero essere aiutati con delle interfacce introduttive o dei percorsi di presentazione.

#### **3.7 Il progetto avrà un uso interno a scopo didattico o esterno?**

Si vogliono raccogliere idee, se la cosa deve uscire dallo scopo didattico il progetto verrà rimaneggiato.

#### **3.8 È desiderabile salvare il lavoro in corso oppure si comincia una nuova sessione ogni volta?**

Dato che dobbiamo fare test automatici, preparare l'ambiente e poi testarlo diventa difficile, quindi meglio di no.

#### **3.9 È desiderabile prevedere un sistema di esportazione?**

No non è previsto, se si vuole condividere un grafico si fa uno screenshot.

#### **3.10 I dati forniti saranno da filtrare prima di visualizzare i grafici oppure sono già pronti all'uso?**

Ci verranno forniti dati anonimi a molte dimensioni. Possiamo partire con set di dati che troviamo liberamente su Internet, ad esempio quello dell'iris. Se ci saranno da fare scremature ci verrà detto più avanti, tendenzialmente comunque ci danno dati da visualizzare così come sono. L'analisi che faranno non è un'analisi sul contenuto dei dati, bensì sulla visualizzazione. Quindi anche una possibile filtratura sarà sempre finalizzata alla visualizzazione grafica.



### 3.11 È importante l'aspetto estetico del software?

D3 fa già bei grafici, quindi è meglio fare qualcosa di visivamente gradevole anche per il resto del software. Che il grafico sia informativo è più importante, poi se è carino ancora meglio.

### 3.12 Avremo file o database da cui prendere i dati?

Non dovete collegarvi a database esterni. Vi saranno forniti file .csv, se poi volete caricarli su un database ok.

### 3.13 Potremmo avere un esempio tipico di utilizzo del prodotto finale?

Si parte da una query o da un file .csv, lo si importa nel sistema, il quale mi mostra i grafici che producono le proiezioni, quindi a questo punto ho un grafico in 2D o 3D prodotto a partire da N dimensioni iniziali. Se sono contento del grafico bene, altrimenti proseguo con altre modifiche sulla proiezione dei dati per ottenere una riduzione che sia utile.

### 3.14 Limite delle dimensioni: sono massimo 15 o almeno 15?

Se si fa un'analisi di riduzione dimensionale è facile escludere delle dimensioni trovando delle scorciatoie. Il numero 15 è per dire che effettivamente c'è stata una operazione di riduzione dimensionale. Anche se ci fermiamo a 9 dimensioni va bene lo stesso. Ad esempio Gli scatter plot sono grafici animati che riescono a includere più dimensioni come: continente, paese, popolazione, tempo, vita media, PIL. Ci sono 6 dimensioni messe in uno scatter plot normale, per le altre 9 dimensioni rimanenti si useranno altri grafici.

### 3.15 Verrà fornita qualche tipo di formazione o è parte del nostro auto apprendimento?

Provate, se avete difficoltà ci si sente. È già capitato in passato di fare un pochino di formazione, ci si mette d'accordo.

### 3.16 È possibile usare TypeScript o è un requisito JavaScript?

Meglio lasciare stare TypeScript. TypeScript permette di fare controlli sul tipo rispetto a JavaScript. Purtroppo ad oggi TypeScript si sta scontrando con i limiti dell'approccio.

### 3.17 In merito a ml.js di D3.js consigliate altre librerie?

È comodo ml.js perché c'è una grande raccolta di distanze. D3.js è l'asse portante perché permette le visualizzazioni. Sicuramente ci sono librerie per leggere file .csv, se vogliamo possiamo usare altre librerie di grafici. D3.js è un consiglio non un obbligo. Se volete avventurarvi con il machine learning ml.js è una raccolta di algoritmi interessante. Se seguite carpati c'è convnet.js per le reti neurali, che può essere un mezzo di riduzione dimensionale. Con auto-encoder si costruiscono reti neurali. L'esempio di convnet.js è un auto-encoder che fa riduzione dimensionale. Mentre karpathy-tsne è uno di quelli che mantiene vicine le distanze piccole ed evidenzia le distanze grandi. Invece umap.js è un metodo simile a tsne con basi più teoriche.

### 3.18 Ci sono requisiti particolari per il Front-End?

No, nessuno.

### 3.19 Dobbiamo usare un grafico a seconda del tipo di dato?

Se ho un tipo di dati vorrei vederli con vari grafici. Li devo provare perché questa attività è molto esplorativa.



## 4 Conclusione dell'incontro

Zucchetti cercherà di anonimizzeranno i dati, però dovremo comunque poter capire di cosa si sta parlando. Se queste dimensioni sono troppo indicative le dovranno scartare o rinominare per non farci capire a cosa sono riferite. La scelta del posizionamento dei dati sui vari assi è una scelta che spetta all'utente. Potrebbe anche essere la macchina che sottolinea delle somiglianze tra i dati. È probabile che chi ha fatto il grafico dell'aspettativa di vita, preso d'esempio durante le domande, abbia fatto delle prove per evidenziare certe proprietà. Ad esempio ci siamo resi conto che la popolazione non influenza molto il PIL.

Zucchetti chiude per Natale, quindi non li troveremo fino al loro rientro.