



UNIVERSITÀ
DEGLI STUDI
DI PALERMO

Progetto Open Data: SchoolRank

A.A 2018/2019

Docente: Davide Taibi

Palumbo Salvatore Danilo - 0642907

Nania Fabio - 0645452

Ragona Gabriele – 0647635

Indice

1. Descrizione
2. Accorgimenti/Premessa
3. Dataset e Licenze
 - 3.1. Dataset utilizzati
 - 3.2. Licenze
4. Analisi Codice
5. Progetto Step-by-Step
6. Query

Descrizione

“SchoolRank” nasce come progetto che usa open data e ha come scopo principale quello di definire un rank di qualità delle scuole grazie a delle caratteristiche, come il riscaldamento o la presenza di strutture a supporto dei disabili. Inoltre, consente di individuare le fermate degli autobus poste in prossimità della scuola stessa.

The project “SchoolRank” has the main purpose of using open data to define a quality rank of the school itself thanks to features, such as heating or the presence of facilities to support the disabled. Also, it allows you to identify the stops located near the school itself.



Accorgimenti/Premessa

Il progetto sfrutta l'elenco delle scuole presenti in tutta Italia, ma non essendo fornito delle coordinate è necessario estrarle utilizzando Open Street Map. OSM non è in grado di identificare molte vie (per trovare la giusta via bisogna aggiungere il CAP) e non riconosce i numeri civici (quindi la precisione è approssimativa).

OSM dopo un elevato numero di richieste consecutive blocca l'accesso.

Abbiamo deciso di utilizzare il comune di Palermo, ma è possibile utilizzare qualsiasi altro comune italiano.

I dataset presentano numerose imperfezioni e incoerenze (dovute anche ad anni di riferimento diversi dei dataset), che sono state quasi completamente corrette grazie all'utilizzo di sostituzioni ed espressioni regolari.

Queste ultime sono state rese il più generiche possibile, in modo tale da attribuire una scalabilità al progetto: basterà modificare il nome del comune di riferimento e fornire un URL o un file GTFS relativo ai trasporti del comune.

In allegato, sono disponibili anche altri comuni italiani.

Nei file delle ontologie invece abbiamo preferito agire scuola per scuola (dato che erano poche) in modo da riuscire a collegare il nome delle scuole alla persona (o luogo) a cui sono dedicate tramite interconnessione con DBpedia (attenzione: alcune interconnessioni sono errate perchè non esistono su DBpedia o conducono a riferimenti non voluti).

Dataset e Licenze

Dataset utilizzati

Abbiamo utilizzato i seguenti dataset:

Scuole statali	Informazioni anagrafiche scuole statali, anno 2018/19
Macro ambienti	Presenza di macro ambienti funzionali (palestra, mensa ecc.), anno 2017/18
Tipologia riscaldamento	Tipologia di riscaldamento, anno 2017/18
Barriere architettoniche	Accorgimenti per il superamento delle barriere architettoniche, anno 2017/18
AMAT Palermo	Amat Palermo SPA - Dati del trasporto pubblico locale, anno 2019

Il file delle scuole statali contiene tutte le informazioni relative alle scuole, come ad esempio il codice della scuola, il codice dell'istituto a cui fa riferimento, il CAP e il tipo di scuola (primaria, liceo scientifico ecc.). Gli unici campi mancanti sono latitudine e longitudine (che verranno estratti);

Il file relativo ai macro ambienti contiene il codice scuola e una serie di infrastrutture, come aula magna, mensa, spazi amministrativi ecc;

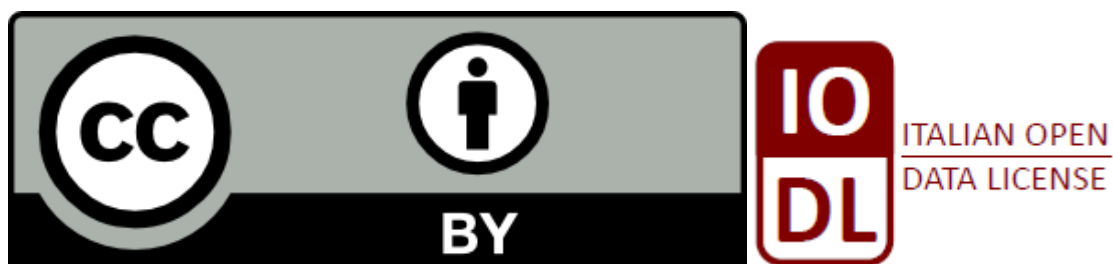
Il file relativo alla tipologia di riscaldamento contiene il codice scuola e i differenti tipi di riscaldamento, ad esempio metano, GPL, teleriscaldamento ecc;

Il file relativo alle barriere architettoniche contiene il codice scuola e vari accorgimenti per disabili, come ad esempio scale a norma, rampe per disabili ecc;

Il file relativo ai trasporti è un GTFS (General Transit Feed Specification), che definisce un formato comune, con file con estensione .txt ma formattati come CSV.

Licenze

I primi 4 dataset sono regolati dalla [Italian Open Data License v2.0](#), mentre quello dell' AMAT di Palermo dalla [CC BY 4.0](#).



Note:

Non sei tenuto a rispettare i termini della licenza per quelle componenti del materiale che siano in pubblico dominio o nei casi in cui il tuo utilizzo sia consentito da una [eccezione o limitazione](#) prevista dalla legge.

Non sono fornite garanzie. La licenza può non conferirti tutte le autorizzazioni necessarie per l'utilizzo che ti prefiggi. Ad esempio, diritti di terzi come [i diritti all'immagine, alla riservatezza e i diritti morali](#) potrebbero restringere gli usi che ti prefiggi sul materiale.

Questo è un riassunto in linguaggio accessibile a tutti (e non un sostituto) della [licenza](#). [Limitazione di responsabilità](#).

Tu sei libero di:

Condividere — riprodurre, distribuire, comunicare al pubblico, esporre in pubblico, rappresentare, eseguire e recitare questo materiale con qualsiasi mezzo e formato

Modificare — remixare, trasformare il materiale e basarti su di esso per le tue opere per qualsiasi fine, anche commerciale.

Il licenziante non può revocare questi diritti fintanto che tu rispetti i termini della licenza.



MINISTERO DELL' ISTRUZIONE, DELL'UNIVERSITÀ E DELLA RICERCA

Analisi Codice

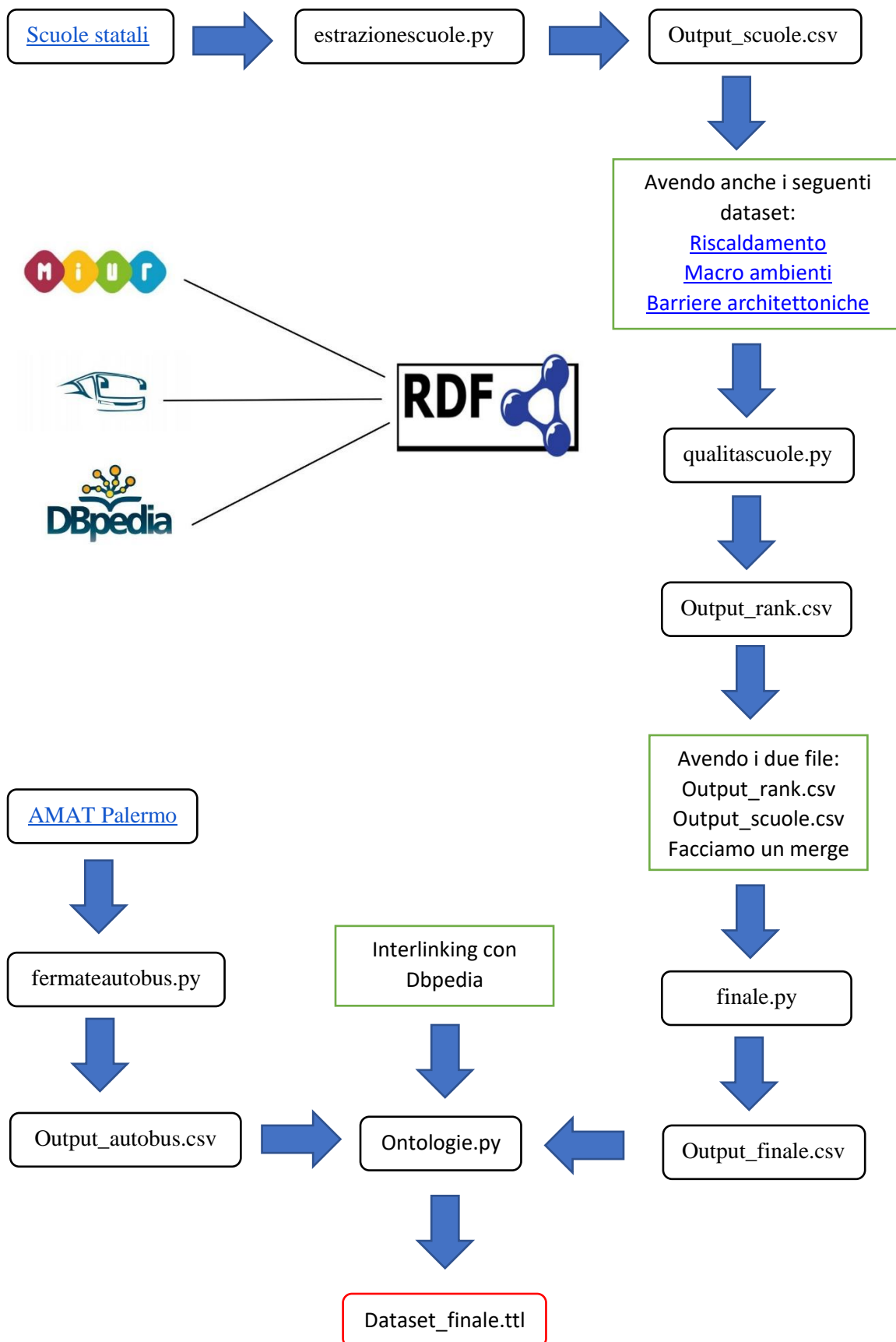
Abbiamo deciso di riportare e commentare il codice utilizzando Jupyter Notebook: riteniamo che sia uno strumento molto utile, perché consente di combinare codice vero e proprio con testo (markdown) e per evitare di rendere questa documentazione prolissa e confusionaria.

Il progetto è composto da differenti file per poter permettere il controllo degli output intermedi.

Il workflow si trova alla pagina successiva.

Progetto Step-by-Step

Il progetto inizia con la ricerca del file delle scuole e il file GTFS del comune italiano scelto. Una volta trovato l'iter che abbiamo seguito è il seguente:



Query

Ecco alcune query:

Le 10 migliori scuole

```
SELECT ?scuola ?nome_scuola ?valutazione
WHERE{ ?scuola skr:nomeScuola ?nome_scuola.
       ?scuola skr:haPlesso ?p.
       ?scuola skr:valutazione ?valutazione.
}
ORDER BY DESC (?valutazione)
LIMIT 10
```

scuola	nome_scuola	valutazione
http://www.schoolrank.org/resource/PAIC897004	I.C. GIULIANA SALADINO - PA	6.8
http://www.schoolrank.org/resource/PAIC890009	I.C. POLITEAMA -PA	6.4
http://www.schoolrank.org/resource/PAIC8A2004	I.C. SILVIO BOCCONE -PA	6.4
http://www.schoolrank.org/resource/PAIC8AA008	I.C. CRUILLAS -PA	6.4
http://www.schoolrank.org/resource/PAIC8AL00X	I.C. GIOVANNI XXIII /PIAZZI -PA	6.4
http://www.schoolrank.org/resource/PAMM03600N	SCUOLA SEC. DI 1? V.E. ORLANDO	6.4
http://www.schoolrank.org/resource/PAMM04800X	SMS DON LORENZO MILANI	6.4
http://www.schoolrank.org/resource/PARH02000A	I.P.S.S.E.O.A. "PIETRO PIAZZA"	6.4
http://www.schoolrank.org/resource/PAEE013002	D.D. A. DE GASPERI - PA	6.0
http://www.schoolrank.org/resource/PAEE02900G	D.D. MONTI IBLEI - PA	6.0

I 10 migliori plessi

```
SELECT ?plesso ?nome_plesso ?valutazione
WHERE{ ?plesso skr:nomeScuola ?nome_plesso.
       ?plesso skr:valutazione ?valutazione.
FILTER NOT EXISTS { ?plesso skr:haPlesso ?p. }
}
ORDER BY DESC (?valutazione)
LIMIT 10
```

plesso	nome_plesso	valutazione
http://www.schoolrank.org/resource/PAAA897011	I.C. GIULIANA SALADINO - PA	6.8
http://www.schoolrank.org/resource/PAAA897022	I.C. GIULIANA SALADINO - PA	6.8
http://www.schoolrank.org/resource/PAAA897033	I.C. GIULIANA SALADINO - PA	6.8
http://www.schoolrank.org/resource/PAEE897016	I.C. GIULIANA SALADINO - PA	6.8
http://www.schoolrank.org/resource/PAAA8A2011	I.C. SILVIO BOCCONE -PA	6.4
http://www.schoolrank.org/resource/PAEE8A2016	I.C. SILVIO BOCCONE -PA	6.4
http://www.schoolrank.org/resource/PAMM8A2015	I.C. SILVIO BOCCONE -PA	6.4
http://www.schoolrank.org/resource/PAAA8AA015	I.C. CRUILLAS -PA	6.4
http://www.schoolrank.org/resource/PAAA8AA026	I.C. CRUILLAS -PA	6.4
http://www.schoolrank.org/resource/PAAA8AA037	I.C. CRUILLAS -PA	6.4

Fermate vicino a scuole e plessi

```
SELECT ?scuola ?nome_scuola ?fermata
WHERE{
  ?scuola skr:nomeScuola ?nome_scuola.
  ?scuola skr:vicinoA ?fermata.
}
```

scuole_e_plessi	nome_scuola	fermata
http://www.schoolrank.org/resource/PAPC04000P	GARIBALDI	http://www.schoolrank.org/resource/point/VILLABIANCA_LAURANA
http://www.schoolrank.org/resource/PAPC04000P	GARIBALDI	http://www.schoolrank.org/resource/point/DALLA_CHIESA
http://www.schoolrank.org/resource/PAPC04000P	GARIBALDI	http://www.schoolrank.org/resource/point/DUCA DELLA VERDURA GIACHERY
http://www.schoolrank.org/resource/PAPC04000P	GARIBALDI	http://www.schoolrank.org/resource/point/DUCA DELLA VERDURA LIBERTÀ
http://www.schoolrank.org/resource/PAPC04000P	GARIBALDI	http://www.schoolrank.org/resource/point/DUCA DELLA VERDURA TOSELLI
http://www.schoolrank.org/resource/PAAA00305B	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/REGIONE_SICILIANA_ALTOFONTE
http://www.schoolrank.org/resource/PAAA00305B	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/REGIONE_SICILIANA_D_ANNA
http://www.schoolrank.org/resource/PAAA00305B	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/REGIONE_SICILIANA_PALMERINO
http://www.schoolrank.org/resource/PAAA00305B	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/REGIONE_SICILIANA_PLATANIA
http://www.schoolrank.org/resource/PAAA00305B	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/SS_MEDIATRICE_BARBARINO
http://www.schoolrank.org/resource/PAAA00305B	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/SS_MEDIATRICE_REGIONE_SICILIANA
http://www.schoolrank.org/resource/PAAA00305B	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/CASALAINA
http://www.schoolrank.org/resource/PAAA00305B	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/PALMERINO_SCONTRINO
http://www.schoolrank.org/resource/PAEE00301C	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/PIAZZA_GLE_TURBA
http://www.schoolrank.org/resource/PAEE00301C	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/SS_MEDIATRICE_BARBARINO
http://www.schoolrank.org/resource/PAEE00301C	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/CALATAFIMI_MARINUZZI
http://www.schoolrank.org/resource/PAEE00301C	D.D.E. ARCULEO - PA	http://www.schoolrank.org/resource/point/LARGO_TRAINA_CATALANO
http://www.schoolrank.org/resource/PAAA01903V	D.D.A. GABELLI - PA	http://www.schoolrank.org/resource/point/MICHELE_PIAZZA
http://www.schoolrank.org/resource/PAAA01903V	D.D.A. GABELLI - PA	http://www.schoolrank.org/resource/point/GUERRAZZI
http://www.schoolrank.org/resource/PAAA01903V	D.D.A. GABELLI - PA	http://www.schoolrank.org/resource/point/L_EMIRO_DI_SVEVIA
http://www.schoolrank.org/resource/PAAA01903V	D.D.A. GABELLI - PA	http://www.schoolrank.org/resource/point/L_EMIRO_EDRISI
http://www.schoolrank.org/resource/PAAA01903V	D.D.A. GABELLI - PA	http://www.schoolrank.org/resource/point/L_EMIRO_GUERRAZZI
http://www.schoolrank.org/resource/PAEE019001	D.D.A. GABELLI - PA	http://www.schoolrank.org/resource/point/L_EMIRO_ARISTIPPO
http://www.schoolrank.org/resource/PAEE019001	D.D.A. GABELLI - PA	http://www.schoolrank.org/resource/point/L_EMIRO_CASTELLUCCIO
http://www.schoolrank.org/resource/PAEE019001	D.D.A. GABELLI - PA	http://www.schoolrank.org/resource/point/L_EMIRO_DI_SVEVIA

Scuole e plessi dedicati a 'giuseppe'

```
SELECT ?scuola ?nome_scuola
WHERE{
  ?scuola skr:nomeScuola ?nome_scuola.
  FILTER REGEX (?nome_scuola, "giuseppe", "i")
}
```

scuole_e_plessi	nome_scuola	dedicatoa
http://www.schoolrank.org/resource/PAAA85203D	I.C. G. DI VITTORIO -PA	http://dbpedia.org/resource/Giuseppe_Di_Vittorio
http://www.schoolrank.org/resource/PAAA85204E	I.C. G. DI VITTORIO -PA	http://dbpedia.org/resource/Giuseppe_Di_Vittorio
http://www.schoolrank.org/resource/PAEE85201L	I.C. G. DI VITTORIO -PA	http://dbpedia.org/resource/Giuseppe_Di_Vittorio
http://www.schoolrank.org/resource/PAEE85202N	I.C. G. DI VITTORIO -PA	http://dbpedia.org/resource/Giuseppe_Di_Vittorio