



Smart Data

Saqib Razzaq

Datasets

CSV file Containing Video games Dataset

Columns: 17

Dataset: 19714

Dataset:

Id, name title, IMDB URL, Year of release, certificate ratings , average rating, number of votes, summary plot, Action is game of action genre?, Adventure is game of adventure genre?,Comedy, Crime, Family, Fantasy, Mystery, Sci-Fi, Thriller.

Mapping

- Prompt Engineering :
- Using ChatGPT Version GPT-4

Convert CVS file dataset into **RDF.**

- RML Generation
- One-Shot
- Few-Shot

Convert CVS file dataset into **RDF.**

- RML Generation
- Zero-Shot
- One-Shot
- Few-Shot

RDF Generation.

- One-shot

SR

You

Convert the following data into RDF.

id, name, url, year, certificate, rating, votes, plot, Action, Adventure, Comedy, Crime, Family, Fantasy, Mystery, Sci-Fi, Thriller

"1, Red Dead Redemption II, https://www.imdb.com/title/tt6161168/?ref=adv_li_tt, 2018, M, 9.7, 35,703, Amidst the decline of the Wild West at the turn of the 20th century, TRUE, TRUE, FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, FALSE"

@prefix ex: <http://example.org/game/> .

@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

ex:1 a ex:VideoGame ;

ex:id "0" ;

ex:name "Spider-Man" ;

ex:url <https://www.imdb.com/title/tt5807780/?ref=adv_li_tt> ;

ex:year "2018" ;

ex:certificate "M" ;

ex:rating "9.2" ;

ex:votes "20,759" ;

ex:plot "When a new villain threatens New York City" ;

ex:genre ex:Action, ex:Adventure, ex:Crime ;

ex:hasFamilyGenre "FALSE" ;

ex:hasFantasyGenre "TRUE" ;

ex:hasMysteryGenre "FALSE" ;

ex:hasSciFiGenre "FALSE" ;

ex:hasThrillerGenre "FALSE" .

Given CSV data:

less

Copy code

"1, Red Dead Redemption II, https://www.imdb.com/title/tt6161168/?ref=

RDF representation:

turtle

Copy code

@prefix ex: <http://example.org/game/> .

@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

ex:1 a ex:VideoGame ;

ex:id "1" ;

ex:name "Red Dead Redemption II" ;

ex:url <https://www.imdb.com/title/tt6161168/?ref=adv_li_tt> ;

ex:year "2018" ;

ex:certificate "M" ;

ex:rating "9.7" ;

ex:votes "35,703" ;

ex:plot "Amidst the decline of the Wild West at the turn of the 20t

ex:genre ex:Action, ex:Adventure, ex:Crime ;

ex:hasFamilyGenre "FALSE" ;

ex:hasFantasyGenre "FALSE" ;

ex:hasMysteryGenre "FALSE" ;

ex:hasSciFiGenre "FALSE" ;

ex:hasThrillerGenre "FALSE" .

RDF Generation.

- Few-shots

- Give Multiple Input to get the for Completing The request

SR You

convert the provided CSV data for provided RDF structure. Combine all of the three dataset of id '0' id '1' and id '2'

id, name, url, year, certificate, rating, votes, plot, Action, Adventure, Comedy, Crime, Family, Fantasy, Mystery, Sci-Fi, Thriller

"1, Red Dead Redemption II, https://www.imdb.com/title/tt6161168/?ref_=adv_li_tt, 2018, M, 9.7, 35,703, Amidst the decline of the Wild West at the turn of the 20th century, TRUE, TRUE, FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, FALSE"

@prefix ex: <http://example.org/game/> .

@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

```
ex:1 a ex:VideoGame ;
  ex:id "0" ;
  ex:name "Spider-Man" ;
  ex:url <https://www.imdb.com/title/tt5807780/?ref_=adv_li_tt> ;
  ex:year "2018" ;
  ex:certificate "M" ;
  ex:rating "9.2" ;
  ex:votes "20,759" ;
  ex:plot "When a new villain threatens New York City" ;
  ex:genre ex:Action, ex:Adventure, ex:Crime ;
  ex:hasFamilyGenre "FALSE" ;
  ex:hasFantasyGenre "TRUE" ;
  ex:hasMysteryGenre "FALSE" ;
  ex:hasSciFiGenre "FALSE" ;
  ex:hasThrillerGenre "FALSE" .
```

###

id, name, url, year, certificate, rating, votes, plot, Action, Adventure, Comedy, Crime, Family, Fantasy, Mystery, Sci-Fi, Thriller

"0, Spider-Man, https://www.imdb.com/title/tt5807780/?ref_=adv_li_tt, 2018, T, 9.2, 20,759, When a new villain threatens New York City, TRUE, TRUE, FALSE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE"

id '2' is: "2, Grand Theft Auto V, https://www.imdb.com/title/tt2103188/?ref_=adv_li_tt, 2013, M, 9.5, 59,986, Three very different criminals team up for a series of heists and walk into some of the most thrilling experiences in the corrupt city of Los Santos., TRUE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, FALSE"

turtle

Copy code

@prefix ex: <http://example.org/game/> .

@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

```
ex:0 a ex:VideoGame ;
  ex:id "0" ;
  ex:name "Spider-Man" ;
  ex:url <https://www.imdb.com/title/tt5807780/?ref_=adv_li_tt> ;
  ex:year "2018" ;
  ex:certificate "T" ;
  ex:rating "9.2" ;
  ex:votes "20,759" ;
  ex:plot "When a new villain threatens New York City" ;
  ex:genre ex:Action, ex:Adventure, ex:Fantasy ;
  ex:hasFamilyGenre "FALSE" ;
  ex:hasMysteryGenre "FALSE" ;
  ex:hasSciFiGenre "FALSE" ;
  ex:hasThrillerGenre "FALSE" .
```

```
ex:1 a ex:VideoGame ;
  ex:id "1" ;
  ex:name "Red Dead Redemption II" ;
  ex:url <https://www.imdb.com/title/tt6161168/?ref_=adv_li_tt> ;
  ex:year "2018" ;
  ex:certificate "M" ;
  ex:rating "9.7" ;
  ex:votes "35,703" ;
  ex:plot "Amidst the decline of the Wild West at the turn of the 20th century" ;
  ex:genre ex:Action, ex:Adventure, ex:Crime ;
  ex:hasFamilyGenre "FALSE" ;
  ex:hasFantasyGenre "FALSE" ;
  ex:hasMysteryGenre "FALSE" ;
  ex:hasSciFiGenre "FALSE" ;
  ex:hasThrillerGenre "FALSE" .
```

```
ex:2 a ex:VideoGame ;
  ex:id "2" ;
  ex:name "Grand Theft Auto V" ;
  ex:url <https://www.imdb.com/title/tt2103188/?ref_=adv_li_tt> ;
  ex:year "2013" ;
  ex:certificate "M" ;
  ex:rating "9.5" ;
```

Mapping

- Prompt Engineering :
- Using ChatGPT Version GPT-4

Convert CVS file dataset into **RML**.

- RML Generation
- Zero-Shot
- One-Shot
- Few-Shot

RML Generation.

- Zero-shot
 - No previous information and data is provided for the completing request.

SR

You



imdb-videogames.csv
Spreadsheet

Generate RML file to convert it to RDF.

```
turtle
Copy code

@prefix rr: <http://www.w3.org/ns/r2rml#>.
@prefix rml: <http://semweb.mmlab.be/ns/rml#>.
@prefix ql: <http://semweb.mmlab.be/ns/ql#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix ex: <http://example.com/ns#>.

<#VideoGameMapping>
  rml:logicalSource [
    rml:source "/path/to/imdb-videogames.csv";
    rml:referenceFormulation ql:CSV
  ];
  rr:subjectMap [
    rr:template "http://example.com/videogame/{id}";
    rr:class ex:VideoGame
  ];
  rr:predicateObjectMap [
    rr:predicate ex:name;
    rr:objectMap [ rr:column "name" ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:url;
    rr:objectMap [ rr:column "url" ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:year;
    rr:objectMap [ rr:column "year"; rr:datatype xsd:integer ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:certificate;
    rr:objectMap [ rr:column "certificate" ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:rating;
    rr:objectMap [ rr:column "rating"; rr:datatype xsd:decimal ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:votes;
    rr:objectMap [ rr:column "votes" ]
  ];
  rr:predicateObjectMap [
```

```
rr:predicateObjectMap [
  rr:predicate ex:plot;
  rr:objectMap [ rr:column "plot" ]
];
# Mapping for genres (e.g., Action, Adventure, etc.)
rr:predicateObjectMap [
  rr:predicate ex:genre;
  rr:objectMap [
    rr:column "Action";
    rr:datatype xsd:boolean;
    rr:termType rr:Literal
  ]
];
```

RML Generation.

- Zero-shot
- Applying Scheme.org Ontologies for rml



You

As you mentioned the column of CSV file are as follows.

id: A unique identifier for each video game.

name: The name of the video game.

url: A URL to the video game's page on IMDb.

year: The release year of the video game.

certificate: The game's rating certificate.

rating: The IMDb rating of the video game.

votes: The number of votes the game received on IMDb.

plot: A brief plot description of the video game.

Action, Adventure, Comedy, Crime, Family, Fantasy, Mystery, Sci-Fi, Thriller: Boolean

columns indicating the genres of the video game.

Apply sheme.org ontologies to them.

Now, let's adjust the RML mapping to use these Schema.org properties. I'll modify the previous RML template to reflect these changes.

```
turtle
Copy code

@prefix rr: <http://www.w3.org/ns/r2rml#>.
@prefix rml: <http://semweb.mmlab.be/ns/rml#>.
@prefix ql: <http://semweb.mmlab.be/ns/ql#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix schema: <http://schema.org/>.

<#VideoGameMapping>
  rml:logicalSource [
    rml:source "/path/to/imdb-videogames.csv";
    rml:referenceFormulation ql:CSV
  ];
  rr:subjectMap [
    rr:template "http://example.com/videogame/{id}";
    rr:class schema:VideoGame
  ];
  rr:predicateObjectMap [
    rr:predicate schema:name;
    rr:objectMap [ rr:column "name" ]
  ];
  rr:predicateObjectMap [
    rr:predicate schema:url;
    rr:objectMap [ rr:column "url" ]
  ];
  rr:predicateObjectMap [
    rr:predicate schema:datePublished;
    rr:objectMap [ rr:column "year"; rr:datatype xsd:gYear ]
  ];
  rr:predicateObjectMap [
    rr:predicate schema:contentRating;
    rr:objectMap [ rr:column "certificate" ]
  ];
  rr:predicateObjectMap [
    rr:predicate schema:aggregateRating;
    rr:objectMap [
      rr:parentTriplesMap <#RatingMapping>
    ]
  ];
  rr:predicateObjectMap [
```

```
rr:objectMap [ rr:column "rating"; rr:datatype xsd:decimal ]
];
rr:predicateObjectMap [
  rr:predicate schema:reviewCount;
  rr:objectMap [ rr:column "votes" ]
];
```

n this revised RML mapping:

RML Generation.

- One-shot

- One Specific Input is provided and given out in scheme.org ontologies.



You
convert this Data into RML format.

"1, Red Dead Redemption II, https://www.imdb.com/title/tt6161168/?ref_=adv_li_tt,
2018, M, 9.7, 35703, Amidst the decline of the Wild West at the turn of the 20th
century, TRUE, TRUE, FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, FALSE"

turtle

Copy code

```
<http://example.com/videogame/1> a schema:VideoGame;  
    schema:name "Red Dead Redemption II";  
    schema:url <https://www.imdb.com/title/tt6161168/?ref_=adv_li_tt>;  
    schema:datePublished "2018"^^xsd:gYear;  
    schema:contentRating "M";  
    schema:description "Amidst the decline of the Wild West at the turn  
    schema:aggregateRating [  
        schema:ratingValue "9.7"^^xsd:decimal;  
        schema:reviewCount "35703"^^xsd:integer;  
    ];  
    schema:genre "Action", "Adventure", "Crime".
```

RML Generation.

- Few-shots
- Multiple pieces of data or guidelines given before completing request
- Generate Mapping for first two Unique ID: id '0' and id '1'



You

Generate RML mapping for these two id '0' and id '0'

"0, Spider-Man, https://www.imdb.com/title/tt5807780/?ref_=adv_li_tt, 2018, T, 9.2, 20,759, When a new villain threatens New York City, TRUE, TRUE, FALSE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE, FALSE"

id '2' is: "2, Grand Theft Auto V, https://www.imdb.com/title/tt2103188/?ref_=adv_li_tt, 2013, M, 9.5, 59,986, Three very different criminals team up for a series of heists and walk into some of the most thrilling experiences in the corrupt city of Los Santos., TRUE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, FALSE"

RML Generation.

- Few-shots

```
turtle Copy code

@prefix rr: <http://www.w3.org/ns/r2rml#>.
@prefix rml: <http://semweb.mmlab.be/ns/rml#>.
@prefix ql: <http://semweb.mmlab.be/ns/ql#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix ex: <http://example.org/game/>.

<#VideoGameMapping>
  rml:logicalSource [
    rml:source "/path/to/your/csvfile.csv";
    rml:referenceFormulation ql:CSV;
    rml:iterator "row"
  ];
  rr:subjectMap [
    rr:template "http://example.org/game/{id}";
    rr:class ex:VideoGame
  ];
  rr:predicateObjectMap [
    rr:predicate ex:id;
    rr:objectMap [ rml:reference "id" ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:name;
    rr:objectMap [ rml:reference "name" ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:url;
    rr:objectMap [ rml:reference "url" ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:year;
    rr:objectMap [ rml:reference "year"; rr:datatype xsd:gYear ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:certificate;
    rr:objectMap [ rml:reference "certificate" ]
  ];
  rr:predicateObjectMap [
    rr:predicate ex:rating;
    rr:objectMap [ rml:reference "rating"; rr:datatype xsd:decimal
  ];
  rr:predicateObjectMap [
```

```
rr:predicateObjectMap [
  rr:predicate ex:plot;
  rr:objectMap [ rml:reference "plot" ]
];
# Genre mappings for Action, Adventure, etc.
rr:predicateObjectMap [
  rr:predicate ex:genre;
  rr:objectMap [
    rml:reference "Action";
    rr:datatype xsd:boolean
  ]
];
```



Thank You