

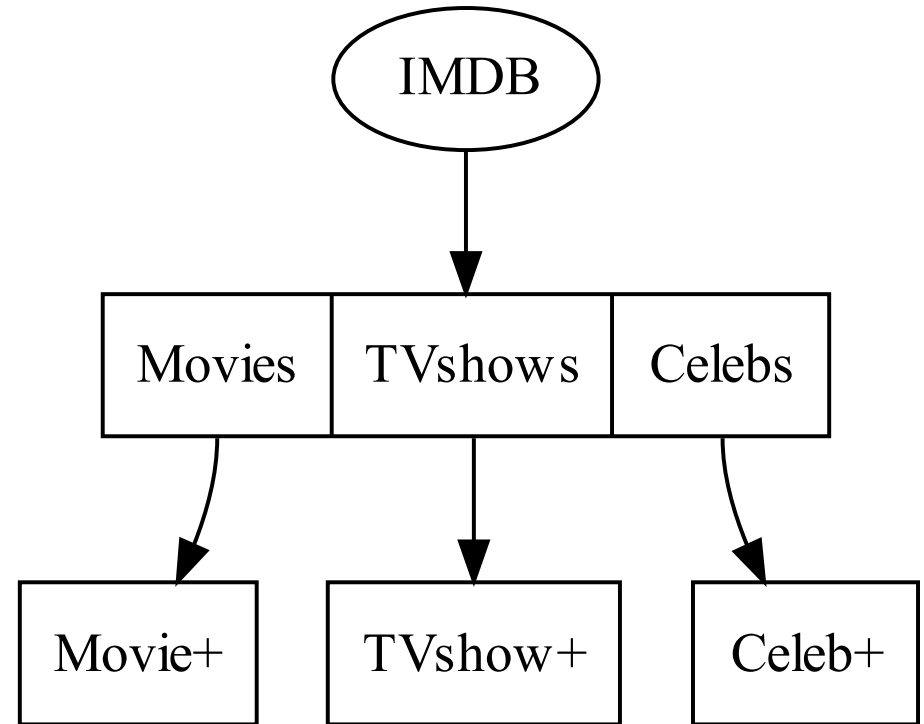
The IMDb logo is centered on the left side of the slide. It consists of the letters "IMDb" in a bold, black, sans-serif font. The letters are set against a bright yellow rectangular background with rounded corners. The entire logo is positioned on a light blue background with a fine, pixelated texture.

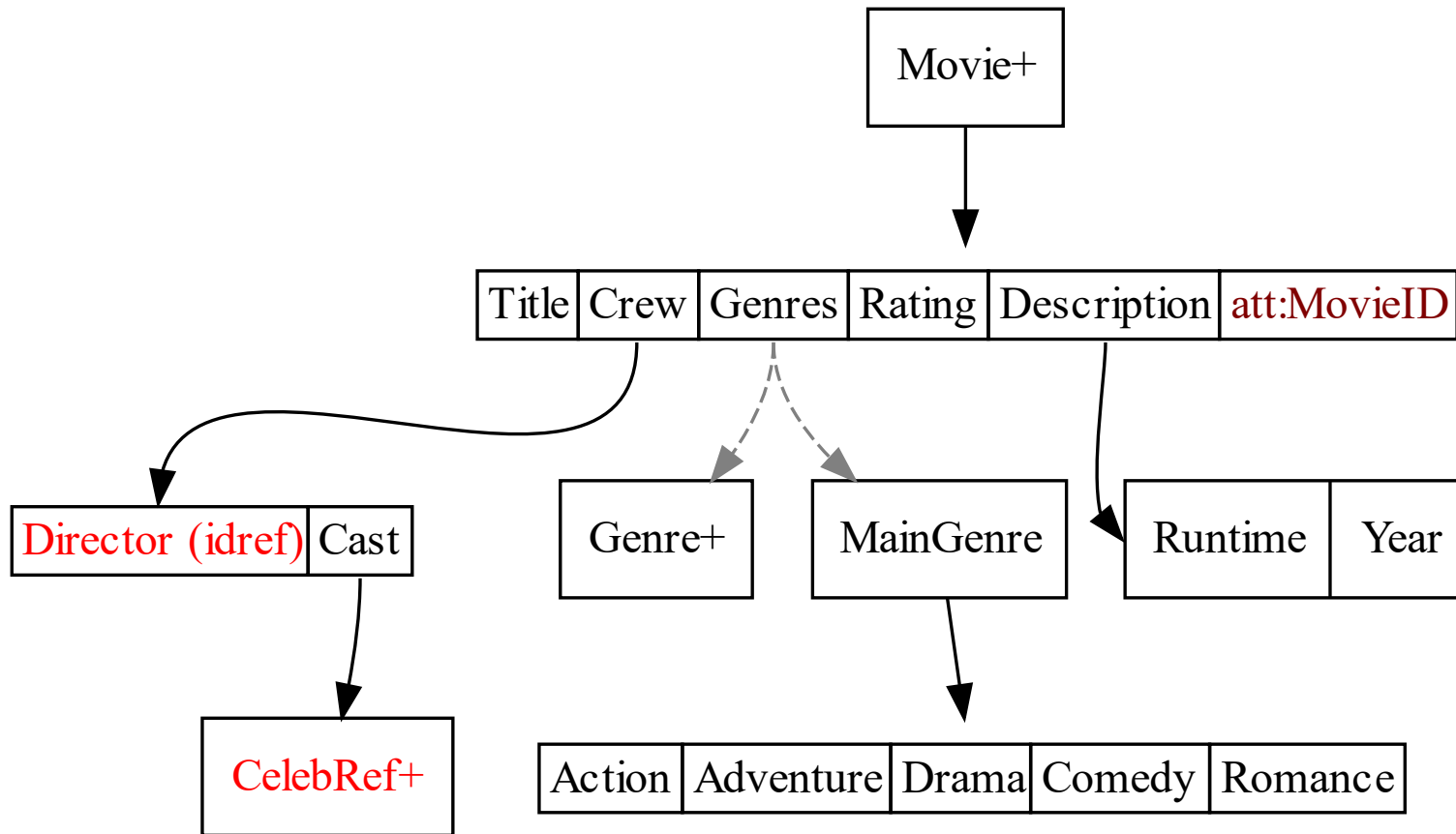
# IMDb

IMDB domain model in XML.  
Lokesh, Nikhil, Akhil, Avinash,  
Anirudh, Dinesh, Shyam

# IMDB

- IMDB contains Movies, Tvshows, Celebs.



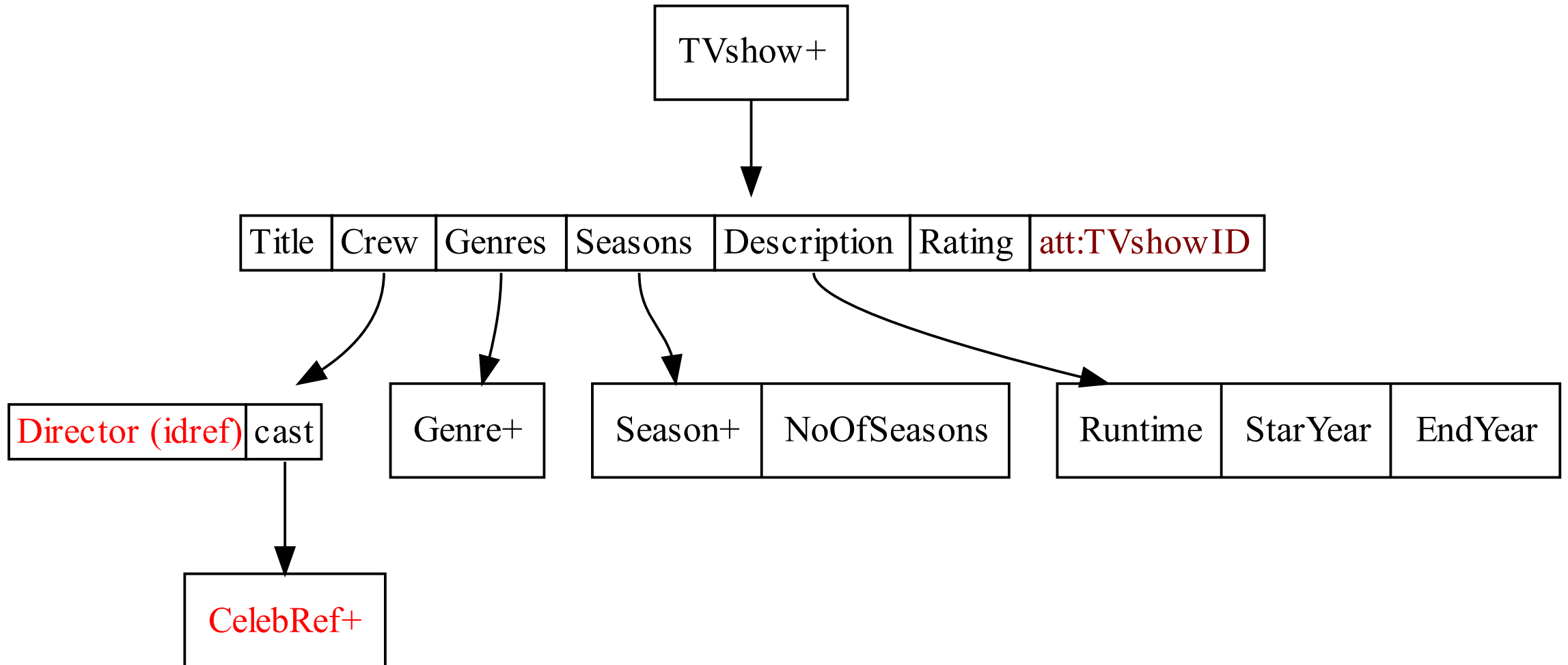


- Movie contains, Title, Crew, Genres, Description, Rating, ID.

Movie

# Tvshow

- Tvshow contains, Title, Crew, Genres, Seasons, Description, Rating, ID.



Continue  
Tvshow

Season+

NoOfSeasons

Episodes

NoOfEpisodes

Rating

Year

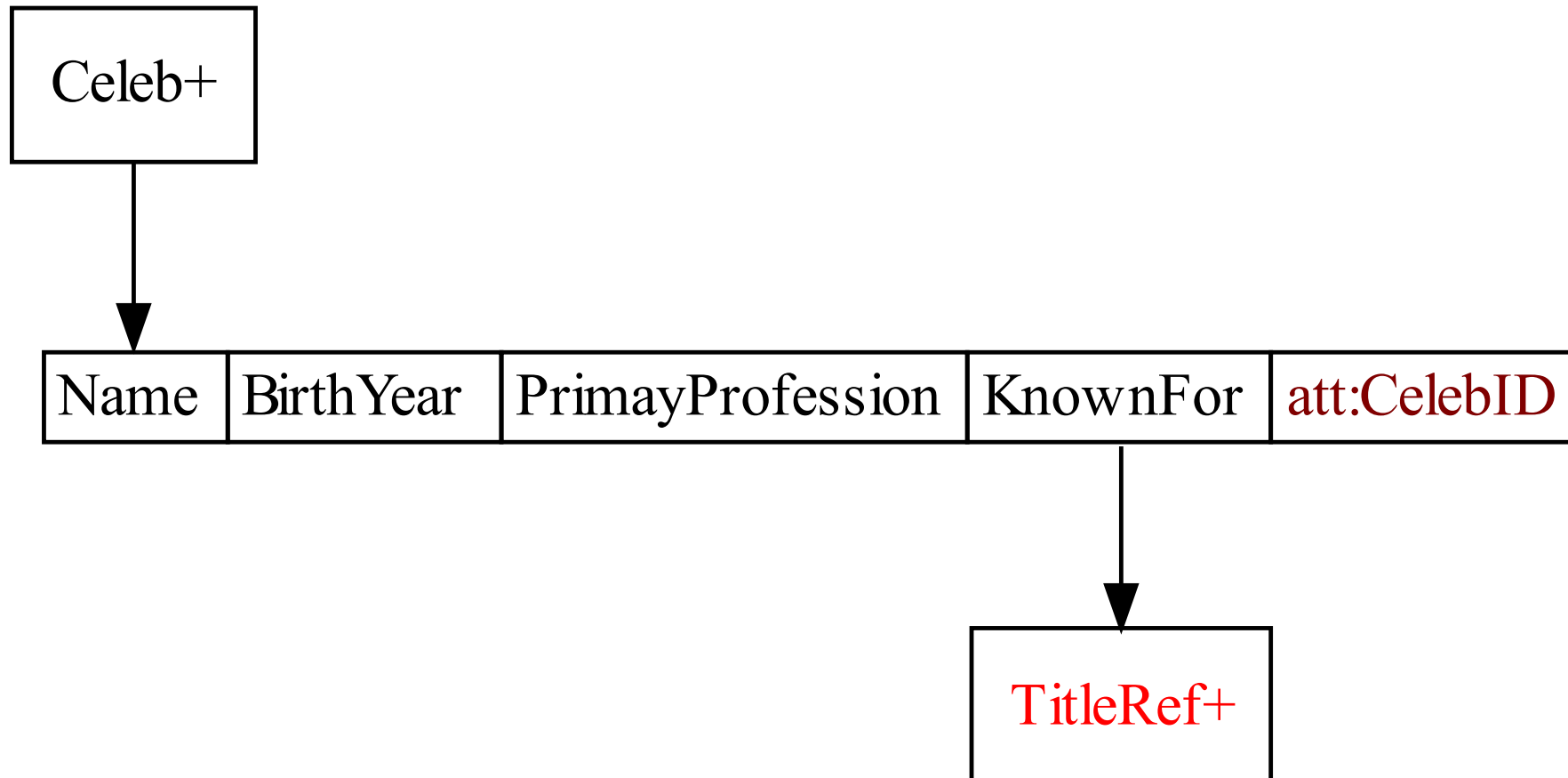
Episode+

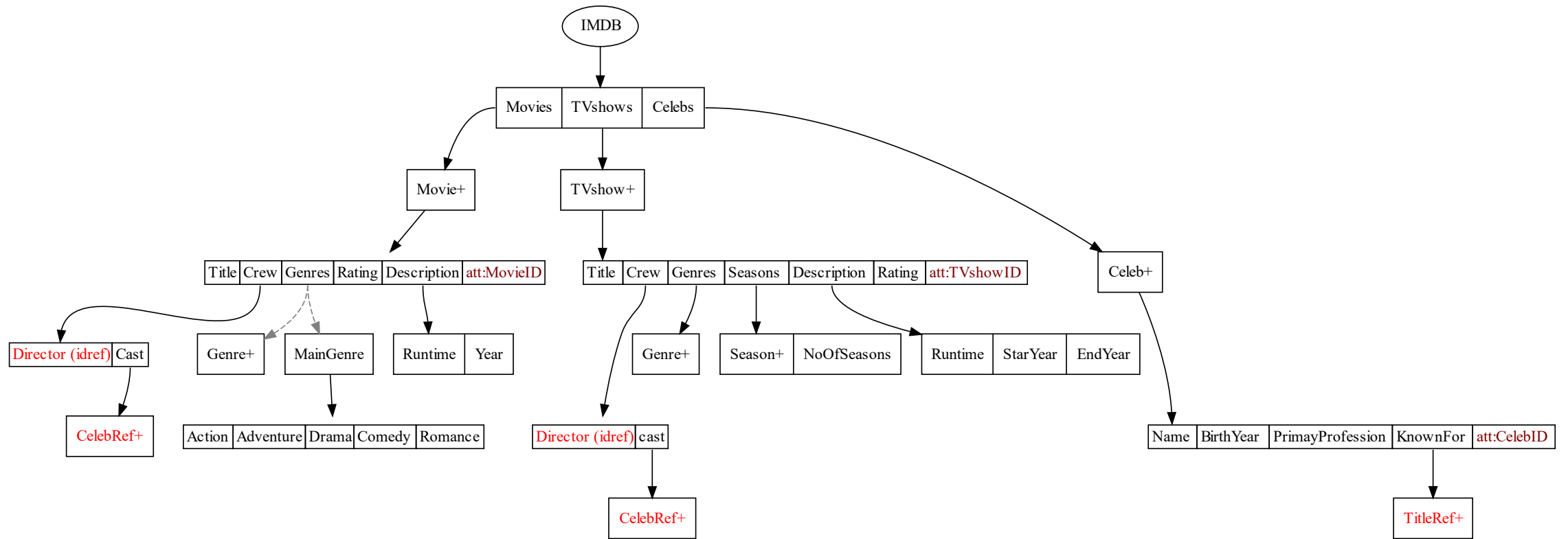
Title

Rating

# Celeb

- Celeb contains, Name, BirthYear, PrimaryProfession, KnownFor, ID.





# SSD properties

- Heterogeneity
  - Genres can have multiple Genre tags or a single MainGenre tag in which there is a tag for each type of genre.
- Cross-Referencing
  - Tvshows, Movies has reference to Celebs.
  - Celebs has reference to Tvshows, Movies.



# Heterogeneity

- Genres may contain many Genre tags or a single MainGenre.

```
<Genres>
  <MainGenre>
    <Action>1</Action>
    <Adventure>1</Adventure>
    <Drama>0</Drama>
    <Comedy>1</Comedy>
    <Romance>0</Romance>
  </MainGenre>
</Genres>
```

```
<Genres>
  <Genre>Action</Genre>
  <Genre>Adventure</Genre>
  <Genre>Comedy</Genre>
</Genres>
```

# Generating xml data

We got data from [datasets.imdbws.com](https://datasets.imdbws.com).

We used pandas and `xml.etree.ElementTree` for generating xml data.

## **Data Validation:**

We used xmllint, visual studio, Liquid Studio 2017 for data validation.

Thank You