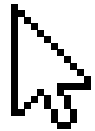
# Linking Parts: Using key phrases of descriptions to link entities (Steam Games)

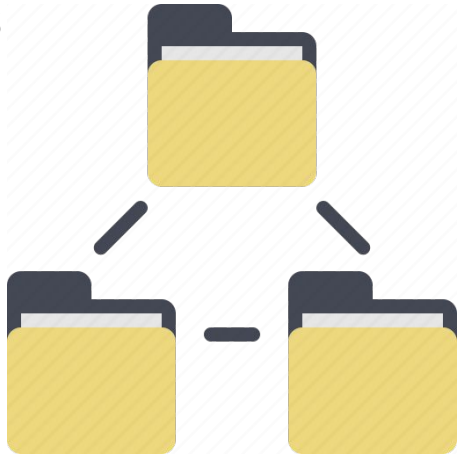
CC7220 - Linked Data



LinkingParts

# Motivation

- MultPAX shows us that we can use linked data to generate absent key phrases using present key phrases.
- We may be able to use present key phrases of entities descriptions to find interesting connections.
- Let’s create a TTL parser that does this automatically for us so we can query a dataset using SPARQL.

# Dataset

- Steam Store Dataset by Nik Davis (Kaggle)
- Steam games have short descriptions on the store so you can know a bit more about the game before buying.


# The Pipeline

Extract the present key phrases using KeyBERT

Keyphrases


It's the closest planet to the Sun

SPARQL

0

1

2

3

## Preprocessing

Get rid of empty data and problematic characters


## TTL

Parse the data, generating the ttl file

# Keyphrase and Parsing Process

| index | name | short_description | top_keyphrases |
|---|---|---|---|
| 0 | Dota 2 | Every day, millions of players worldwide enter battle as one of over a hundred Dota heroes. And no matter if it's their 10th hour of play or 1,000th, there's always something new to discover. With regular updates that ensure a constant evolution of gameplay, features, and heroes, Dota 2 has taken on a life of its own. | battle dota heroes matter, battle dota heroes, gameplay features heroes dota, dota heroes matter, dota heroes, battle dota, heroes dota, features heroes dota |
| 1 | Counter-Strike: Global Offensive | Counter-Strike: Global Offensive (CS: GO) expands upon the team-based action gameplay that it pioneered when it was launched 19 years ago. CS: GO features new maps, characters, weapons, and game modes, and delivers updated versions of the classic CS content (de_dust2, etc.). | counter strike global offensive, strike global offensive cs, counter strike global, global offensive cs, global offensive cs expands, strike global offensive, counter strike, offensive cs expands |

```
:570 rdfs:label "Dota 2";
:rating "85.20128129694771"^^xsd:float; :med_play "801"^^xsd:int; :owners "100000000"^^xsd:int;
:hasKP :battle_dota_heroes_matter ; :hasKP :battle_dota_heroes ; :hasKP :gameplay_features_heroes_dota ;
:hasKP :dota_heroes_matter ; :hasKP :dota_heroes; :dev :Valve .

:730 rdfs:label "CounterStrike Global Offensive";
:rating "86.31031244654261"^^xsd:float; :med_play "6502"^^xsd:int; :owners "50000000"^^xsd:int;
:hasKP :counter_strike_global_offensive ; :hasKP :strike_global_offensive_cs ; :hasKP :counter_strike_global ;
:hasKP :global_offensive_cs ; :hasKP :global_offensive_cs_expands; :dev :Valve ; :dev :Hidden_Path_Entertainment .
```

# Interesting queries

| | | |
|---|---|---|
| "Thief" | "Thief Deadly Shadows" | 2 |
| "Amnesia A Machine for Pigs" | "Amnesia The Dark Descent" | 2 |
| "Call of Duty 4 Modern Warfare" | "Call of Duty Black Ops" | 2 |
| "CS2D" | "Natural Selection 2" | 2 |
| "The Guild 3" | "The Guild Gold Edition" | 2 |
| "Sakura Beach" | "Sakura Fantasy" | 2 |
| "Shadow Warrior" | "Shadow Warrior 2" | 2 |
| "Call of Duty Black Ops" | "Call of Duty World at War" | 2 |
| "Call of Duty 4 Modern Warfare" | "Call of Duty World at War" | 2 |
| "FINAL FANTASY IV" | "FINAL FANTASY IV THE AFTER YEARS" | 2 |
| "NEKOPARA Vol 0" | "NEKOPARA Vol 1" | 2 |
| "Defense Grid The Awakening" | "Prime World Defenders 2" | 1 |
| "Defense Grid The Awakening" | "GUNS UP" | 1 |
| "DG2 Defense Grid 2" | "GemCraft  Chasing Shadows" | 1 |

```
SELECT ?game1Label ?game2Label
(COUNT(?keyphrase) as
?keyphraseCount)
WHERE {
   ?game1 :hasKP ?keyphrase .
   ?game2 :hasKP ?keyphrase .
   FILTER(?game1 ≠ ?game2)
   ?game1 rdfs:label ?game1Label
.
   ?game2 rdfs:label ?game2Label
.
   FILTER(?game1Label <
?game2Label)
}
GROUP BY ?game1Label ?game2Label
ORDER BY DESC (?keyphraseCount)
```

This query groups by a pair of different games and then counts their common keyphrases

# Interesting queries

| dev1 | dev2 | keyphraseCount |
|---|---|---|
| http://ex.org/Feral_Interactive_%28Mac%29 | http://ex.org/Feral_Interactive_%28Linux%29 | 106 |
| http://ex.org/Feral_Interactive_%28Linux%29 | http://ex.org/Feral_Interactive_%28Mac%29 | 106 |
| http://ex.org/Feral_Interactive_%28Mac%29 | http://ex.org/CREATIVE_ASSEMBLY | 30 |
| http://ex.org/CREATIVE_ASSEMBLY | http://ex.org/Feral_Interactive_%28Mac%29 | 30 |
| http://ex.org/Pilgrim_Adventures | http://ex.org/GrabTheGames_Studios | 25 |
| http://ex.org/Feral_Interactive_%28Linux%29 | http://ex.org/CREATIVE_ASSEMBLY | 25 |
| http://ex.org/Aspyr_%28Mac%29 | http://ex.org/Aspyr_%28Linux%29 | 25 |
| http://ex.org/Aspyr_%28Linux%29 | http://ex.org/Aspyr_%28Mac%29 | 25 |
| http://ex.org/GrabTheGames_Studios | http://ex.org/Pilgrim_Adventures | 25 |
| http://ex.org/CREATIVE_ASSEMBLY | http://ex.org/Feral_Interactive_%28Linux%29 | 25 |
| http://ex.org/Idea_Factory | http://ex.org/Compile_Heart | 20 |
| http://ex.org/Feral_Interactive_%28Mac%29 | http://ex.org/Traveller%27s_Tales | 20 |
| http://ex.org/Traveller%27s_Tales | http://ex.org/Feral_Interactive_%28Mac%29 | 20 |
| http://ex.org/Compile_Heart | http://ex.org/Idea_Factory | 20 |
| http://ex.org/Crystal_Dynamics | http://ex.org/Feral_Interactive_%28Mac%29 | 19 |
| http://ex.org/Feral_Interactive_%28Mac%29 | http://ex.org/Crystal_Dynamics | 19 |
| http://ex.org/Ubisoft_Shanghai | http://ex.org/Ubisoft_Kiev | 17 |
| http://ex.org/Back_To_Basics_Gaming | http://ex.org/SnowFlame | 17 |

```
SELECT ?dev1 ?dev2
(COUNT(?keyphrase) as
?keyphraseCount)
FROM <https://grupo1.cl>
WHERE {
   ?game1 :hasKP ?keyphrase; :dev
?dev1 .
   ?game2 :hasKP ?keyphrase; :dev
?dev2 .
   FILTER(?dev1 ≠ ?dev2)
}
GROUP BY ?dev1 ?dev2
ORDER BY DESC (?keyphraseCount)
```

This query groups by a pair of developers and counts the number of shared keyphrases between their games

# Interesting case


event[0]

We queried games with the same keyphrases from different developers and obtained information about Eidolon and Event[0], both of which share the keyphrase “narrative exploration game”

# Conclusions and further topics

1. Some queries failed due to exceeding the **time limit** in Virtuoso.
2. With **NLP** techniques, it could also be possible to classify certain key phrases and then use **OWL** to generate ontologies.
3. **SPARQL** empowers us to formulate interesting queries, as demonstrated in the case mentioned, further use of data from **remote sources** is proposed.


FIN