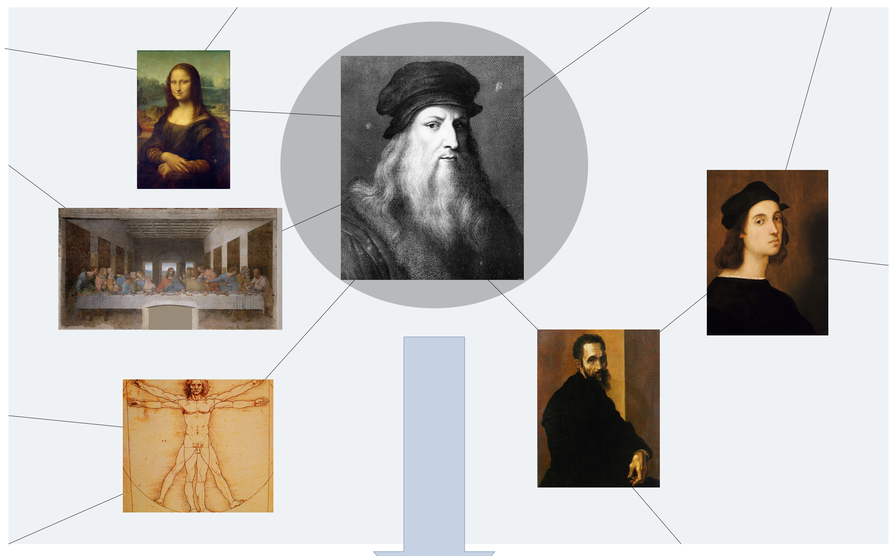
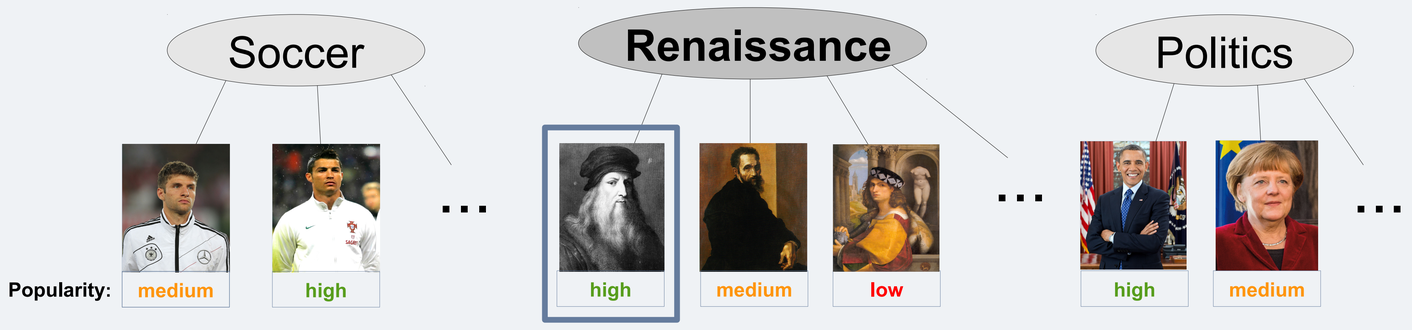
Generating Quiz Questions from Knowledge Graphs

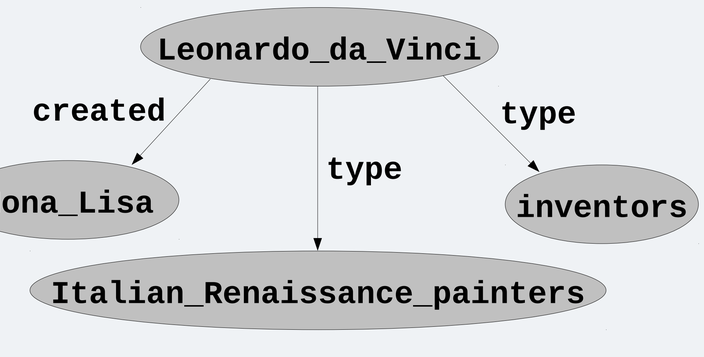
Dominic Seyler, Mohamed Yahya, Klaus Berberich
{dseyler, myahya, kberberi}@mpi-inf.mpg.de



Answer Selection

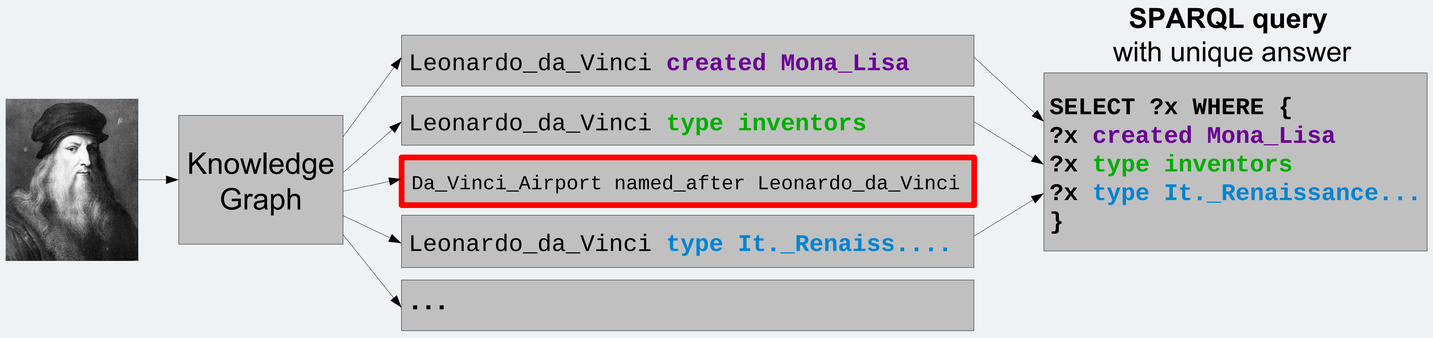
*Select named entity **e** as answer to question for topic: **Renaissance***

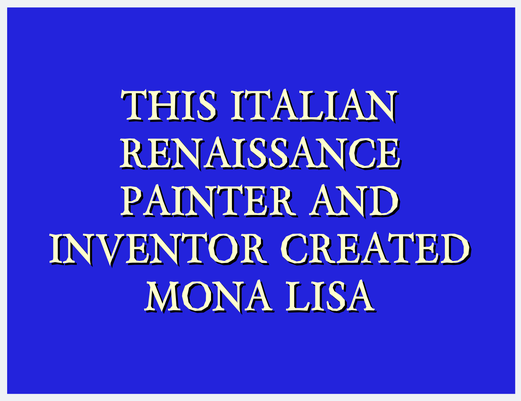




Query Generation

Generate SPARQL query for a specific difficulty





Question Verbalization

Verbalize SPARQL query yielding a natural language question

-Turn **type** to singular : `inventors` → `inventor`
-Construct dictionary to verbalize predicates **p** : `created` → `has creator`
-Use canonical surface form for objects **o** : `Mona_Lisa` → `Mona Lisa`
-Verbalize using pattern:

$$\textit{This}\ \texttt{type}_1, \ldots,\ \textit{and}\ \texttt{type}_m\ \texttt{p}_1\ \texttt{o}_1, \ldots,\ \textit{and}\ \texttt{p}_n\ \texttt{o}_n\ .$$

Examples

easy

```
Leonardo_da_Vinci type painter .
Leonardo_da_Vinci created Mona_Lisa .
Leonardo_da_Vinci created Vitruvian Man .
Leonardo_da_Vinci created The_Last_Supper
```

This painter created Mona Lisa, Vitruvian Man, and The Last Supper.

hard

```
Leonardo_da_Vinci type scientist .
Leonardo_da_Vinci type engineer .
Leonardo_da_Vinci influences Victor_Bregeda .
Leonardo_da_Vinci created Portrait_of_a_Musician
```

This scientist and engineer influences Victor Bregeda and created Portrait of a Musician.

Question Difficulty

Popularity: fraction of links in Wikipedia which point to the target entity's article.

$$\textit{Difficulty} = p(\texttt{e}) + \frac{1}{n}\sum_{i=1}^{n} s(\texttt{s}_\texttt{i}\ \texttt{p}_\texttt{i}\ \texttt{o}_\texttt{i}) + \frac{1}{n}\sum_{i=1}^{n} c(\texttt{s}_\texttt{i}\ \texttt{p}_\texttt{i}\ \texttt{o}_\texttt{i})\ .$$

Selectivity: reciprocal number of answer triples in the knowledge graph

Coherence: Jaccard coefficient of the sets of Wikipedia articles pointing to **s** and **o**

https://gate.d5.mpi-inf.mpg.de/q2g/

MAX-PLANCK-GESELLSCHAFT

max planck institut informatik