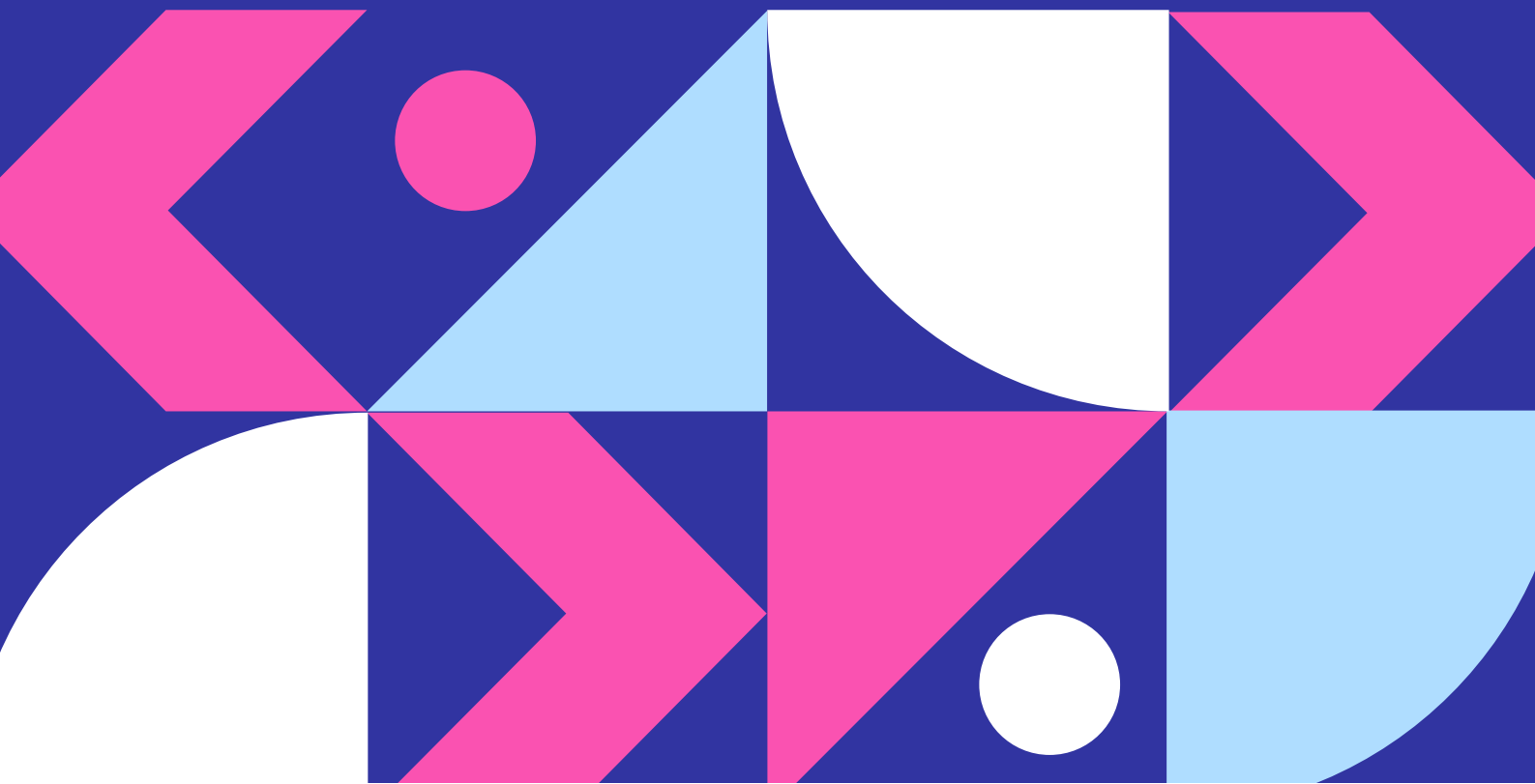
SoCK Use Cases

Explore the use of completeness patterns



#1
G20 Nations

Meet Lisa, a web developer who is working on a web development project for the G20 Presidency event in Indonesia. Lisa is responsible for creating a feature that displays all the countries which are members of the G20. Gradually, Lisa decides to use the DBpedia knowledge graph. However, Lisa is doubtful whether all the data she needs is available on DBpedia. Therefore, Lisa decides to do a validation check first to make sure all information related to the countries that are members of the G20 is in the knowledge graph.

Figure 1 Form to create completeness pattern instance

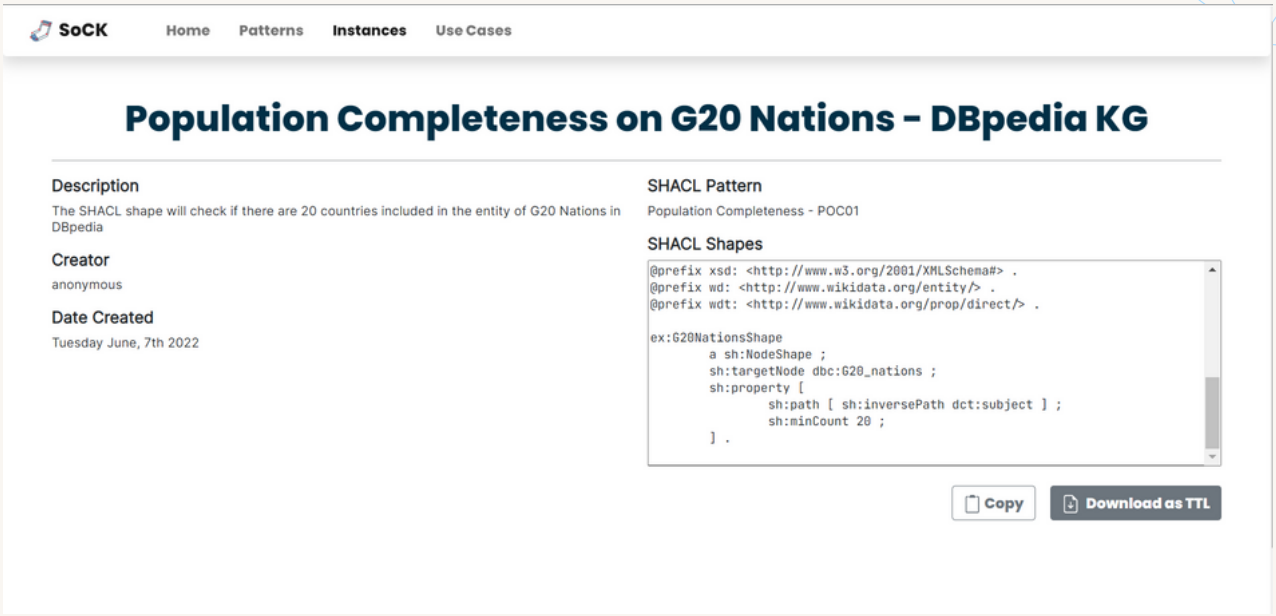


Figure 2 Instance pattern for G20 nations

Through the SoCK website, Lisa looks for a pattern of completeness that matches the case she is facing. Finally, Lisa finds that the completeness pattern which fits her case is the **population completeness** pattern. Through this pattern, it creates an instance that will check whether the dbc:G20_nations entity has 20 countries belonging to the G20 entity. The process of creating an instance of the pattern can be seen in Figure 1. After completing the instance, Lisa can check the result by searching for the name of the instance that was just created. Then, Lisa can view the instance in Figure 2 and download the SHACL shape to use as validation against the dbc:G20_nations entity on DBpedia.

#2
American Films

Meet Max, a freelancer who recently landed a film website development project. In particular, all films shown on the website are from the United States of America. At first, Max was confused about where he got the film data. Eventually, Max discovers that Wikidata has the movie data he needs. He has the idea to display the title of the film along with a brief description or synopsis of the film. Max needs to check if each entity from a movie on Wikidata contains a label and a description.

Using the SoCK website, Max is able to find a pattern of completeness that fits his case. Max chooses the **label and description completeness**. Through this pattern, Max creates an instance that checks whether the entities from "American Film" class have label and description properties in English according to their country of origin. In the process of creating the instance, Max customizes the target section so that it fits into the structure of the Wikidata. The result of creating an instance by Max can be seen in Figure 3.

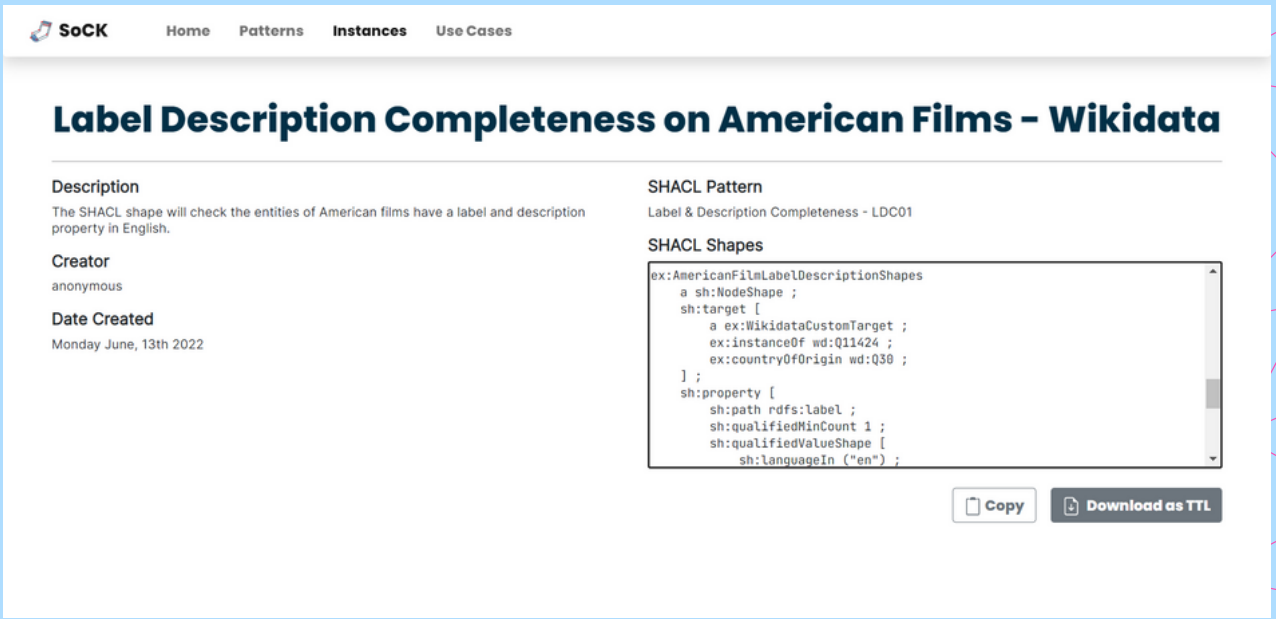


Figure 3 Instance pattern for American films label and description

#3
Number of TV Series Seasons

Meet Mino, an entry level data scientist. Mino has just been given an assignment by his boss to analyze data related to a TV series. At the initial stage, he needed to do data crawling on the TV series data. One of the data needed was the number of seasons of the series. Mino found that the Wikidata knowledge graph stores information related to TV series, one of them was the number of seasons along with information from that season. However, Mino was worried that the information contained in the Wikidata knowledge graph was incomplete. Therefore, he decided to validate the completeness of the data first.

Mino used the SoCK website to find out which pattern of completeness fits the problem. Mino finally got a completeness pattern that matched was **property completeness** as a check for the completeness of a specific property (in this case the number of seasons) on a particular entity (TV series). Helped by the provided pattern, Mino successfully created an instance of property completeness for checking number of seasons from his favorite TV series, Attack on Titan. The instance could check whether its entity has complete information for each season according to the real situation.

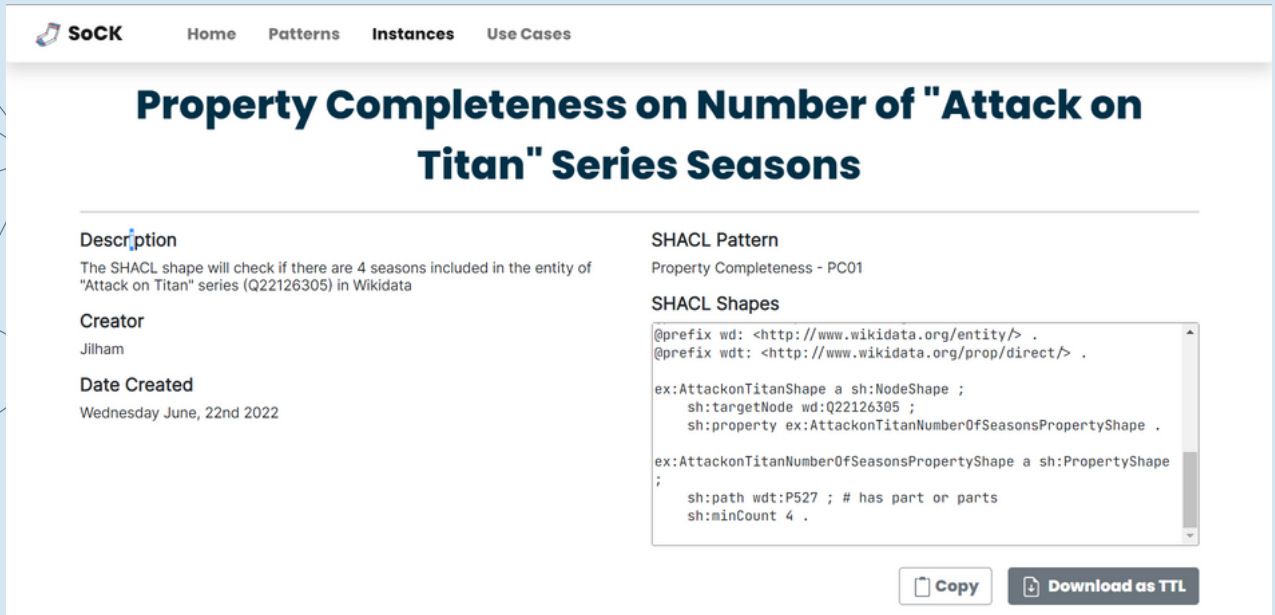


Figure 4 Instance pattern for number of TV series seasons