

DIT045 Exemplar 3: Originality Checker

Fall 2017

Summary

Coming up with new and unique ideas is difficult. Given the large amount of creative content, including books, movies, blogs, and video games – most now easily accessible via the Internet, it's particularly difficult to ensure your new idea is really new. How do we know that our new idea, story, or plot is novel enough? Some overlap with existing works is probably unavoidable, but how much overlap is OK? How do we find similarities between our creative work and existing work of others? This is important not only for the quality of our work, but to avoid copyright infringements, problems with plagiarism rules, and lawsuits for improper use of existing work.

More Detail

Existing tools like [quertext](#) and [urkund](#) help to check for plagiarism by looking for copied or similar text. But we are interested in something deeper, not just to know if the text itself is copied, but if the ideas, concepts, and style are sufficiently unique. We need a tool that breaks text down to its essence, focusing on concepts, actions, and relationships, to find other material which bears similarity. We are interested in similarity in a variety of other artistic works, such as books, scripts, movie synopses, or other ways to capture stories and plots. What inputs would an application like this need? Where do these inputs come from? Who are the different types users, what are they trying to do with such a tool? How can we ensure that we don't break any copyright rules while trying to detect potential copyright violations? How can we protect the privacy and intellectual property of the work which is used as an input to our product? How can we balance between the speed of our responses and the thoroughness of our results?

Novelty

It's fine to take inspiration from existing software and tools, but the requirements for your products should finally describe something which is sufficiently unique and could be a potential competitor to existing applications. In this domain, the idea is to focus beyond plagiarism and simple copied text to look for similarity in ideas.

Further Information

This description is intended to give a good start on developing requirements for the potential system. It's possible that further information or clarifications may be needed in order to model and specify the requirements. You can ask the various stakeholders involved in the domain questions using the discussion forum on GUL. Find the discussion forum for this exemplar. Ask a clarifying question, making sure that you indicate which stakeholder you are asking. The "stakeholder" will provide you an answer. Answers and responses are shared between all teams working with this exemplar.