

## Literature Review on Text Comparison Techniques: Analysis of Text Extraction, Main Comparison and Visual Representation Tools

**Authors :** Andriana Mkrtchyan, Vahe Khlgatyan

**Abstract :** The choice of a profession is one of the most important decisions people make throughout their life. With the development of modern science, technologies, and all the spheres existing in the modern world, more and more professions are being arisen that complicate even more the process of choosing. Hence, there is a need for a guiding platform to help people to choose a profession and the right career path based on their interests, skills, and personality. This review aims at analyzing existing methods of comparing PDF format documents and suggests that a 3-stage approach is implemented for the comparison, that is - 1. text extraction from PDF format documents, 2. comparison of the extracted text via NLP algorithms, 3. comparison representation using special shape and color psychology methodology.

**Keywords :** color psychology, data acquisition/extraction, data augmentation, disambiguation, natural language processing, outlier detection, semantic similarity, text-mining, user evaluation, visual search

**Conference Title :** ICNLPPTA 2024 : International Conference on Natural Language Processing and Text Processing Aids

**Conference Location :** Amsterdam, Netherlands

**Conference Dates :** January 15-16, 2024