

Plagiat selalu menjadi sorotan, terutama di sektor akademis. Sering ditemui pekerjaan teman yang dihargai dengan nilai "0" karena diketahui telah menyalin hasil pekerjaan temannya yang lain. Praktik menyalin beberapa bagian atau keseluruhan tulisan menjadi hal biasa yang sering ditemukan pada penelitian.

Pencegahan dan pendeteksian dini merupakan cara yang dapat dilakukan untuk mengurangi plagiat. Pencegahan berarti menghalangi munculnya plagiat yang lebih ditekankan kepada moral masyarakat dan sistem pendidikan. Cara ini akan memberikan efek jangka panjang.

Duplikasi dokumen, deteksi plagiat, dan pencocokan string telah banyak dibahas pada penelitian-penelitian sebelumnya. Algoritma yang digunakan diantaranya Winnowing, Smith Waterman, Boyer Moore, dan Rabin Karp. Namun sebagian besar tanpa menggunakan preprocessing, sehingga berpengaruh pada akurasi similarity.

Pendeteksian plagiat menggunakan konsep similarity atau kemiripan dokumen merupakan salah satu cara untuk mendeteksi copy&paste plagiarism dan disguised plagiarism. Menggunakan algoritma Rabin Karp yang menerapkan metode fingerprinting yang hanya melakukan pencocokan pola string, pendeteksian plagiat ini tidak memperhatikan adanya penulisan sumber rujukan.