

**Jakub Radoszewski**

Tytuł projektu: „Nowe trendy w algorytmach tekstowych”

---

## Popularnonaukowe streszczenie projektu

W związku z rozwojem cyfryzacji i jej zastosowań w praktycznie każdej dziedzinie techniki oraz w życiu codziennym przetwarzane są obecnie ogromne ilości danych w postaci elektronicznej. Terabajty danych w postaci surowych sekwencji wartości – nazywanych dalej dla uproszczenia tekstami – każdego dnia powiększają zasoby Internetu. Aby móc dobrze wykorzystywać te zasoby, potrzebne są efektywne algorytmy (tj. metody) do przetwarzania danych w postaci tekstowej. Głównym celem projektu jest stworzenie takich właśnie algorytmów.

W projekcie będą rozpatrywane problemy wyszukiwania wzorców i różnego rodzaju regularności w tekstach. Naszym celem są zarówno algorytmy, które działają szybko w praktyce, jak i szczegółowe analizy teoretyczne ich działania, pozwalające lepiej zrozumieć te metody. Będziemy brać pod uwagę różne sytuacje wywodzące się z zastosowań praktycznych, dla których obecnie znane metody często działają w sposób niedostatecznie satysfakcjonujący. Przykładowo, rozważymy sytuację, w której musimy przetwarzać tekst w postaci strumienia danych, a nie mamy do dyspozycji dostatecznej ilości pamięci, żeby przechowywać cały ten tekst lub choćby jego znaczący fragment. Będziemy także łączyć znane metody z najnowszymi technikami uczenia maszynowego, aby otrzymać algorytmy działające możliwie dobrze na rzeczywistych danych. W szczególności, przewidywanym wynikiem pracy nad projektem będą konkretne implementacje wybranych algorytmów.