

POPULARNONAUKOWE STRESZCZENIE PROJEKTU (W JĘZYKU POLSKIM)

Tematem projektu jest opracowanie metod umożliwiających efektywną organizację i przetwarzanie zbiorów danych pochodzących z doświadczeń naukowych zawierających dane o charakterze ilościowym, tzn. obserwacje cech charakteryzujących badane obiekty. Podstawowym celem pracy jest zdefiniowanie semantycznej miary podobieństwa zbiorów danych naukowych, która stanowić będzie podstawę ich klasyfikacji i wyszukiwania. Miara ta będzie uwzględniać dane nie tylko pod kątem ich wartości numerycznych, ale również znaczenia obserwacji, zarówno dla rozważanego obszaru wiedzy, jak i znaczenia statystycznego. Innymi słowy, zastosowane zostanie podejście wykorzystujące tzw. warstwę semantyczną danych. Opracowane rozwiązania zostaną wykorzystane do przygotowania prototypu systemu semantycznej organizacji i przetwarzania zbiorów danych naukowych o charakterze ilościowym.

Zagadnienie zautomatyzowanego przetwarzania zbiorów danych naukowych jest istotne i aktualne o tyle, że ilość publicznie dostępnych zasobów szybko się zwiększa. Publikowanie danych eksperymentalnych staje się powszechnym elementem prowadzenia badań w związku z wymaganiami instytucji finansujących naukę oraz czasopism naukowych, jak również z powodu korzyści związanych z cytowaniem zbiorów danych. Warunkiem użyteczności publicznych zasobów danych jest możliwość ich sprawnego wyszukiwania, poprawnej interpretacji i integracji celem prowadzenia dalszych analiz. Semantyczna reprezentacja wiedzy umożliwia formalizację i unifikację opisu zasobów, tym samym zapewniając ich zrozumiałość dla automatycznych metod przetwarzania.

Zastosowanie opracowanych rozwiązań, umożliwiających wyszukiwanie i porównywanie zbiorów danych ilościowych, zostanie zademonstrowane na przykładzie zbiorów danych naukowych reprezentujących wyniki eksperymentów z dziedziny biologii roślin polegających na fenotypowaniu, czyli obserwacji cech ilościowych ważnych dla oceny przystosowania roślin do środowiska i ich wartości ekonomicznej.