

Baza danych jest narzędziem służącym do zbierania i organizowania informacji. Kiedy tymi informacjami są dane biologiczne, mówimy o bioinformatycznych bazach danych. Ich tworzenie jest jednym z najważniejszych zastosowań bioinformatyki.

Wyniki eksperymentów biologicznych oraz rezultaty przetwarzania tych wyników gromadzone są od wielu lat w różnych repozytoriach danych. Z czasem ilość posiadanych informacji jest coraz większa, wtedy pojawia się potrzeba stworzenia efektywnego systemu, który pozwoli na ich szybkie wyszukiwanie i analizę. Bioinformatyczne bazy danych pozwalają na przeszukiwanie dostępnych w repozytoriach informacji, przetwarzanie ich, a często również i przesyłanie nowych danych.

Każdego roku liczba baz danych wzrasta, a ich większość jest dostępna darmowo i bez ograniczeń. Korzystają z nich również osoby, które nie posiadają wykształcenia informatycznego, przez co charakteryzują się prostym i intuicyjnym interfejsem, a także często wyposażone są w oprogramowanie, które umożliwia wstępną analizę danych.

Bioinformatyka znajduje zastosowanie w niezliczonej ilości zadań, jednym z nich jest identyfikacja, klasyfikacja oraz analiza funkcjonalna białek. Stąd też białkowe bazy danych pełnią bardzo istotną rolę jako repozytoria, w których możliwe jest deponowanie odkrytych białek, ich struktury, umieszczanie informacji na temat ich funkcji oraz udostępnianie zgromadzonej wiedzy szerokiemu środowisku naukowemu. Przykładem takiej bazy danych jest PDB (Protein Data Bank). Klasyfikuje ona struktury białek w oparciu o ich ewolucyjne pokrewieństwo i reguły, które determinują ich trójwymiarową strukturę.