

BIBLIOTEKA REQUESTS

CZYM JEST?

Biblioteka **Requests** zalicza się do pojęcia **web scrapingu**.

Web scraping polega na pobieraniu ze stron internetowych konkretnych interesujących nas informacji.

Requests jest powszechnie używaną biblioteką HTTP na licencji Apache2, wykorzystywaną w **Pythonie**. Służy ona (jak wskazuje nazwa) – do wykonywania requestów, czyli żądań dostępu do strony internetowej.

PRZYKŁAD

W naszym przypadku request ma zakończyć się pobraniem zawartości strony, my wysyłamy żądanie o dostęp do **kantoru-online**, po czym dostajemy odpowiedź z interesującymi nas informacjami, czyli aktualnymi kursami walut. W ten oto sposób, nasza aplikacja jest cały czas na bieżąco i jest oczywiście uzależniona od tej strony.

MIEĆ NA UWADZE

Requests pobiera cały fragment strony, który nas interesuje, czyli pobiera też wszystkie znaczniki HTML oraz zbędną treść.

Nas zazwyczaj interesuje jakiś konkretny rodzaj treści, jakaś konkretna tabela, czy opis. W tym celu dołączamy do programu bibliotekę **BS4 – BeautifulSoup 4** służącą do parsowania zawartości.

Samo pojęcie **web parsingu** odnosi się do przetwarzania obszernego tekstu na mniejsze fragmenty, łatwiejsze do czytania lub analizowania.