



**Program szkolenia stacjonarnego  
„Korpusy językowe: metodologia badań korpusowych i narzędzia  
do analizy danych. Wprowadzenie do platformy SketchEngine”  
realizowanego w ramach programu Inicjatywa Doskonałości  
– Uczelnia Badawcza (IDUB)**

W szkoleniu mogą uczestniczyć nauczyciele akademicki oraz pracownicy badawczo-techniczni zajmujący się badaniami korpusowymi.

Prowadząca: **dr hab. Agnieszka Leńko-Szymańska, Wydział Lingwistyki Stosowanej UW**

Szkolenie obejmujące 15 godz. dydaktycznych będzie realizowane **w sali komputerowej**.

**Celem szkolenia jest** wprowadzenie dotyczące zagadnienia korpusów i językoznawstwa korpusowego. W szczególności podczas szkolenia uczestnicy zapoznają się z:

- najważniejszymi pojęciami związanymi z korpusami i ich analizą;
- korpusowymi metodami badawczymi;
- zasobami korpusowymi dostępnymi w sieci;
- zasadami tworzenia własnych korpusów;
- najpopularniejszymi narzędziami do analizy danych korpusowych;
- podstawowymi opisami korpusowymi;
- testami statystycznymi wykorzystywanymi w badaniach;
- możliwościami i zasadami korzystania z platformy Sketch Engine.

**Po ukończeniu szkolenia uczestnik:**

a) będzie znał:

- najważniejsze pojęcia dotyczące korpusów i ich analizy,
- metody badawcze z zakresu językoznawstwa korpusowego,
- podstawowe opisy i testy statystyczne używane do analizy danych korpusowych,
- najważniejsze zasoby korpusowe dostępne w sieci,
- najpopularniejsze narzędzie do analizy danych korpusowych;

b) będzie potrafił:

- zrozumieć szczegóły analizy i wyniki w opublikowanych badaniach korpusowych,
- tworzyć i analizować własne korpusy,
- korzystać z platformy Sketch Engine w podstawowym zakresie,
- kształcić się samodzielnie w zakresie korzystania z narzędzie korpusowych;

c) będzie doceniał:

- miejsce i rolę korpusów w badaniach językoznawczych,
- zalety i wady danych korpusowych.

**Ramowy program:**

**Zajęcia I**

1. miejsce korpusów w badaniach językoznawczych
2. rodzaje korpusów językowych
3. podstawowe pojęcia związane a analiza korpusową
4. najważniejsze projekty korpusowe języka polskiego i angielskiego

## Zajęcia II

1. wprowadzenie do platformy SketchEngine
2. zasoby korpusowe dostępne na platformie SketchEngine
3. generowanie listy słów oraz wiązek leksykalnych
4. analiza częstotliwości, normalizacja częstotliwości, porównawcze testy statystyczne

## Zajęcia III

1. generowanie konkordancji i ich analiza
2. budowanie zapytań
3. łączliwość wyrazów i struktur, kolokacje i profile wyrazów (wordsketches)
4. statystyczne miary łączliwości wyrazów

## Zajęcia IV

1. generowanie listy słów i wyrażenia kluczowych – ekstrakcja terminologii
2. przegląd zaawansowanych funkcji do analizy danych dostępnych na platformie *Sketch Engine*
3. tworzenie własnych korpusów

## Zajęcia V.

1. tworzenie i analiza korpusów wielojęzycznych
2. inne narzędzia do analizy korpusów dostępne w sieci (analizatory morfosyntaktyczne, lematyzatory, parsery i inne).
3. podsumowanie, zakończenie, przekierowanie uczestników do wypełnienia post-testu na platformę [www.szkolonia-rozwoj.uw.edu.pl](http://www.szkolonia-rozwoj.uw.edu.pl) (10 min.)