



UNIVERSITY
OF WARSAW

Faculty of Modern Languages
School for International Science

R M

Research Methodology

**Korpusy językowe:
metodologia badań korpusowych
i narzędzia do analizy danych.
Wprowadzenie do platformy
Sketch Engine**

Wydział Neofilologii UW zaprasza pracowników naszego Uniwersytetu na kurs prowadzenia badań korpusowych – **Korpusy językowe: metodologia badań korpusowych i narzędzia do analizy danych. Wprowadzenie do platformy Sketch Engine**. Korpusy językowe są coraz częściej wykorzystywane w badaniach z zakresu językoznawstwa, ale także innych nauk społecznych oraz literatury. Dają one dostęp do ogromnych zasobów autentycznych i naturalnych wypowiedzi pisemnych i ustnych, dzięki czemu umożliwiają bardziej trafną i rzetelną analizę języka na jego wielu poziomach: fonetycznym i fonologicznym, morfologicznym, syntaktycznym, leksykalnym, frazeologicznym, semantycznym, pragmatycznym oraz na poziomie dyskursu. Dzięki korpusom powstały zupełnie nowe metody analizy danych językowych uwypuklające pojęcie częstotliwości i zjawisko powtarzalności schematów. Językoznawstwo korpusowe proponuje także nowe podejście do opisu języka, oparte na prawdopodobieństwie raczej niż regułach.

W Internecie dostępnych jest wiele narzędzi służących do różnorodnej analizy danych korpusowych. Umożliwiają one samodzielną pracę badawczą naukowcom, którzy nie posiadają zaawansowanych umiejętności informatycznych, takich jak programowanie. Jednak korzystanie z tych aplikacji może być początkowo trudne i zniechęcić potencjalnych użytkowników. Zasady funkcjonowania takich programów oraz możliwości, jakie oferują, nie są intuicyjne i wymagają wprowadzenia oraz omówienia.

Przykładem narzędzi tego typu jest platforma Sketch Engine. Nie jest ona ogólnie dostępna, jednak pracownicy i studenci UW mogą korzystać z niej nieodpłatnie i bez ograniczeń w ramach europejskiego programu ELEXIS. Jest to obecnie jedna z najlepszych aplikacji do analizy danych korpusowych. Dodatkowo oferuje ona dostęp do ponad 500 korpusów w prawie 100 różnych językach.

Kurs Korpusy językowe: metodologia badań korpusowych i narzędzia do analizy danych. Wprowadzenie do platformy Sketch Engine adresowany jest do osób, które:

- chcą prowadzić badania korpusowe,
- chcą efektywnie wykorzystać platformę Sketch Engine,
- chcą opanować podstawowe testy statystyczne używane w badaniach korpusowych.

Kurs Korpusy językowe: metodologia badań korpusowych i narzędzia do analizy danych. Wprowadzenie do platformy Sketch Engine obejmuje 10 godzin zegarowych (5 x 2 godziny). Zajęcia prowadzone będą w języku polskim, zaś dane omawiane w przykładach będą pochodzić z różnych języków, głównie z języka angielskiego. Kurs będzie zawierał prezentacje omawianych zagadnień uzupełnione zadaniami warsztatowymi. Główny nacisk zostanie położony raczej na przedstawienie i przegląd dziedziny aniżeli na techniczne szczegóły korzystania z aplikacji. Jednak w przypadku zapotrzebowania, istnieje możliwość organizacji dalszych kursów ukierunkowanych na naukę obsługi konkretnego programu do analizy danych korpusowych.

CELE KURSU:

Kurs Korpusy językowe: metodologia badań korpusowych i narzędzia do analizy danych. Wprowadzenie do platformy Sketch Engine jest wprowadzeniem w meandry korpusów i językoznawstwa korpusowego. Jego zakres obejmuje prezentację najważniejszych pojęć z tej dziedziny, omówienie korpusowych metod badawczych stosowanych w językoznawstwie oraz przegląd najpopularniejszych narzędzi do analizy danych korpusowych. Przedstawione zostaną także podstawowe opisy i testy statystyczne wykorzystywane w badaniach tego typu. Uczestnicy będą mieli także okazję poznać możliwości i zasady korzystania z platformy Sketch Engine.

Uczestnicy kursu poznają:

- najważniejsze pojęcia związane z korpusami i ich analizą,
- korpusowe metody badawcze,

- zasoby korpusowe dostępne w sieci,
- zasady tworzenia własnych korpusów,
- najpopularniejsze narzędzia do analizy danych korpusowych,
- podstawowe opisy i testy statystyczne wykorzystywane w badaniach korpusowych,
- możliwości i zasady korzystania z platformy Sketch Engine.

EFEKTY KSZTAŁCENIA:

Po ukończeniu kursu *Korpusy językowe: metodologia badań korpusowych i narzędzia do analizy danych*. Wprowadzenie do platformy Sketch Engine **uczestnik 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.

Uczestnik 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.

Uczestnik będzie doceniał:

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

SYLABUS (10 h)

MODUŁ I:

2
godziny

TREŚĆ:

- miejsce korpusów w badaniach językoznawczych
- rodzaje korpusów językowych
- podstawowe pojęcia związane z analizą korpusową
- najważniejsze projekty korpusowe języka polskiego i angielskiego

MODUŁ II:

2
godziny

TREŚĆ:

- wprowadzenie do platformy Sketch Engine
- zasoby korpusowe dostępne na platformie Sketch Engine
- generowanie listy słów oraz wiązek leksykalnych
- analiza częstotliwości, normalizacja częstotliwości, porównawcze testy statystyczne

MODUŁ III:

2
godziny

TREŚĆ:

- generowanie konkordancji i ich analiza
- budowanie zapytań
- łączliwość wyrazów i struktur, kolokacje i profile wyrazów (wordsketches)
- statystyczne miary łączliwości wyrazów

MODUŁ IV:

2
godziny

TREŚĆ:

- generowanie listy słów i wyrażeń kluczowych – ekstrakcja terminologii
- przegląd zaawansowanych funkcji do analizy danych dostępnych na platformie Sketch Engine
- tworzenie własnych korpusów

MODUŁ V:

2
godziny

TREŚĆ:

- tworzenie i analiza korpusów wielojęzycznych
- inne narzędzia do analizy korpusów dostępne w sieci (analizatory morfosyntaktyczne, lematyzatory, parsery i inne)

KONTAKT

School for International Science
Wydział Neofilologii | Uniwersytet Warszawski
www.sis.uw.edu.pl | mail: sis@uw.edu.pl