



UNIwersYTET
IM. ADAMA MICKIEWICZA
W POZNANIU

Komputerowe wspomaganie tłumaczenia Sylabus zajęć

Informacje podstawowe

Kierunek studiów Informatyka	Cykl dydaktyczny 2023/24
Specjalność -	Kod zajęć 06INFS.41S.01006.23
Jednostka organizacyjna Wydział Matematyki i Informatyki	Języki wykładowe polski
Poziom studiów studia drugiego stopnia poinżynierskie	Obligatoryjność Fakultatywny
Forma studiów studia stacjonarne	Blok zajęciowy Przedmioty specjalnościowe
Profil studiów profil ogólnoakademicki	
Koordinator zajęć	Rafał Jaworski
Prowadzący zajęcia	Rafał Jaworski
Okres Semestr 1	Forma zajęć / liczba godzin / forma zaliczenia • Laboratorium: 30, Zaliczenie z oceną
	Liczba punktów ECTS 3

Cele kształcenia dla zajęć

Kod	Cel
C1	Poznanie idei komputerowego wspomaganie tłumaczenia w odniesieniu do alternatywnych technik (m.in. tłumaczenia automatycznego).
C2	Uzyskanie wiedzy o najczęściej stosowanych w praktyce technikach wspomaganie tłumaczenia.
C3	Poznanie różnych systemów wspomaganie tłumaczenia (Trados, memoq).
C4	Poznanie i wykorzystanie narzędzi do diagnozowania produktywności tłumaczenia (translog).
C5	Poznanie rozwiązań wspomaganie tłumaczenia opartych na uczeniu maszynowym.
C6	Nabywanie umiejętności projektowania i ewaluowania własnych narzędzi wspomaganie tłumaczenia.

Wymagania wstępne

Umiejętność programowania na poziomie inżyniera informatyki. Znajomość podstaw inżynierii programowania. Znajomość podstaw systemu kontroli wersji Git. Umiejętność korzystania z systemu typu Linux z wiersza poleceń. Podstawowa znajomość języków programowania Python, Bash oraz Java.

Efekty uczenia się dla zajęć

Kod	Efekty uczenia się dla zajęć w zakresie	Efekty uczenia się dla kierunku	Metody weryfikacji osiągnięcia efektów uczenia się dla zajęć
Wiedzy - Student/ka:			
W1	Zna podstawowe techniki komputerowego wspomaganie tłumaczenia przy użyciu pamięci tłumaczeń.	INF_K4_W03	Raport
W2	Zna zaawansowane techniki użycia pamięci tłumaczeń oraz kierunki badań w tej dziedzinie.	INF_K4_W02	Projekt
Umiejętności - Student/ka:			
U1	Potrafi zaimplementować działający ekstraktor terminologii używając różnych technik rozwiązania problemu.	INF_K4_U03	Raport
U2	Potrafi zbudować słownik dziedzinowy używając samodzielnie zaprojektowanych algorytmów.	INF_K4_U06	Projekt
U3	Potrafi używać wyrażeń regularnych do zadań związanych ze wspomaganie pracy tłumacza.	INF_K4_U02	Projekt
U4	Zna kryteria użyteczności systemów tłumaczenia automatycznego oraz potrafi dokonać ewaluacji tych systemów.	INF_K4_U07	Raport
U5	Potrafi używać technik web scrapingu do pozyskiwania zasobów lingwistycznych.	INF_K4_U02	Projekt
U6	Posiada wiedzę i umiejętności w dziedzinie automatycznego uśredniania tekstów.	INF_K4_U07	Raport
U7	Potrafi wykonać ewaluację przy użyciu oprogramowania typu key logger.	INF_K4_U03	Raport
U8	Zna techniki korekty pisowni, potrafi zaimplementować własny korektor pisowni.	INF_K4_U02	Projekt

Kod	Efekty uczenia się dla zajęć w zakresie	Efekty uczenia się dla kierunku	Metody weryfikacji osiągnięcia efektów uczenia się dla zajęć
U9	Zna algorytmy korekty gramatycznej tekstu, potrafi zaimplementować własny korektor gramatyczny na podstawie znanych rozwiązań.	INF_K4_U02	Raport
U10	Potrafi skonstruować własny system komputerowego wspomaganie tłumaczenia.	INF_K4_U02, INF_K4_U03, INF_K4_U04, INF_K4_U05	Projekt

Treści programowe dla zajęć

Lp.	Treści programowe dla zajęć	Efekty uczenia się dla zajęć	Formy zajęć
1.	Zapoznanie z podstawowymi pojęciami wspomaganie tłumaczenia. Testowanie działania pamięci tłumaczeń w popularnych programach do wspomaganie tłumaczenia	W1	Laboratorium
2.	Zaawansowane użycie pamięci tłumaczeń - ICE matching, fuzzy matching. Implementacja modułu pamięci tłumaczeń w oparciu o bibliotekę Lucene.	W2	Laboratorium
3.	Techniki automatycznego zarządzania terminologią, w tym ekstrakcja terminologii. Ewaluacja narzędzi do ekstrakcji fraz i terminologii.	U1	Laboratorium
4.	Klasyfikacja dziedzinowa terminologii - techniki automatyczne i półautomatyczne. Tworzenie słowników dziedzinowych.	U2	Laboratorium
5.	Preprocessing i postprocessing tłumaczonych tekstów - automatyczne wstawianie elementów formatujących, konwersje dat i liczb. Zastosowanie wyrażeń regularnych do postprocessingu.	U3	Laboratorium
6.	Tłumaczenie automatyczne jako technika wspomaganie tłumaczenia. Ewaluacja jakości oraz możliwości wykorzystania tłumaczenia automatycznego do wspomaganie tłumaczenia.	U4	Laboratorium
7.	Web scraping - pozyskiwanie danych na potrzeby wspomaganie tłumaczenia. Praktyczne ćwiczenia z web scrapingu.	U5	Laboratorium
8.	Urównoleglanie jako technika tworzenia pamięci tłumaczeń. Eksperymenty z wykorzystaniem dostępnych narzędzi do urównoleglania.	U6	Laboratorium
9.	Techniki badania wydajności procesu tłumaczenia - key logging, eye tracking. Uruchomienie keyloggera (Translog) i analiza wyników.	U7	Laboratorium
10.	Algorytmy automatycznej korekty pisowni. Ewaluacja/implementacja korektora pisowni.	U8	Laboratorium
11.	Automatyczna korekta gramatyczna tekstu. Ewaluacja wybranych narzędzi do korekty gramatycznej (Grammarly, MS Word).	U9	Laboratorium
12.	Projekt własnego mechanizmu wspomaganie tłumaczenia. Implementacja i ewaluacja wybranej techniki wspomaganie tłumaczenia.	U10	Laboratorium

Informacje dodatkowe

Forma zajęć	Metody i formy prowadzenia zajęć
Laboratorium	Metoda analizy przypadków, Uczenie problemowe (Problem-based learning), Metoda laboratoryjna, Metoda warsztatowa, Praca w grupach

Forma zajęć	Warunki zaliczenia zajęć
Laboratorium	Końcowa ocena składa się z następujących elementów: 1. projekt - 20%, 2. raport - 80%. Skala ocen: 1. bardzo dobry (bdb; 5,0) - od 90% punktów, 2. dobry plus (db plus; 4,5) - od 80% punktów, 3. dobry (db; 4,0) - od 70% punktów, 4. dostateczny plus (dst plus; 3,5) - od 60% punktów, 5. dostateczny (dst; 3,0) - od 50% punktów, 6. niedostateczny (ndst; 2,0) - poniżej 50% punktów.

Literatura

Obowiązkowa

1. Lynne Bowker: Computer-aided Translation Technology: A Practical Introduction
2. Dokumentacja systemu Trados <https://docs.sdl.com/>
3. Dokumentacja systemu memoq <https://docs.memoq.com/>
4. Opis systemu Translog: <http://www.translog.dk/>
5. Dokumentacja frameworków NLP: <https://stanfordnlp.github.io/CoreNLP/> oraz <https://spacy.io/>
6. Marcello Federico: Measuring User Productivity in Machine Translation Enhanced Computer Assisted Translation

Nakład pracy studenta i punkty ECTS

Rodzaje zajęć studenta	Średnia liczba godzin* przeznaczonych na zrealizowane rodzaje zajęć
Laboratorium	30
Przygotowanie do zajęć	5
Czytanie wskazanej literatury	10
Przygotowanie projektu	10
Inne	20
Łączny nakład pracy studenta	Liczba godzin 75
Liczba punktów ECTS	ECTS 3

* godzina (lekcyjna) oznacza 45 minut

Efekty uczenia się dla kierunku

Kod	Treść
INF_K4_U02	Absolwent/ka potrafi adaptować istniejące oraz tworzyć nowe metody informatyczne do rozwiązywania nieszablonowych problemów praktycznych i teoretycznych
INF_K4_U03	Absolwent/ka potrafi stosować zaawansowane metody budowy oprogramowania, rozstrzyga o ich przydatności, w tym podejmuje decyzje dotyczące wyboru technik prowadzących do otrzymania oprogramowania wysokiej jakości
INF_K4_U04	Absolwent/ka potrafi projektować i implementować systemy informatyczne o różnej złożoności i różnych architekturach
INF_K4_U05	Absolwent/ka potrafi formułować i testować nowe algorytmy i metody rozwiązywania problemów w wybranych obszarach informatyki na potrzeby prowadzenia prac badawczo-rozwojowych z uwzględnieniem aktualnego stanu wiedzy
INF_K4_U06	Absolwent/ka potrafi rozwiązywać złożone problemy z wybranych obszarów informatyki oraz proponować nowe algorytmy, narzędzia i metody wykorzystując odpowiednio dobrane źródła, które poddaje krytycznej analizie, syntezie i twórczej interpretacji
INF_K4_U07	Absolwent/ka potrafi wyrażać krytyczne opinie na temat architektury oraz użyteczności wykorzystywanych systemów informatycznych
INF_K4_W02	Absolwent/ka zna i rozumie współczesny stan badań i tendencje rozwojowe w wiodących obszarach informatyki
INF_K4_W03	Absolwent/ka zna i rozumie w pogłębionym stopniu współczesne metody, narzędzia i technologie informatyczne właściwe dla wybranych obszarów zastosowań niezbędne przy budowie złożonych systemów informatycznych oraz przy prowadzeniu prac badawczo-rozwojowych