

## PROGRAM DOKTORANCKIEJ SZKOŁY LETNIEJ

1. NAZWA:

*Cyfrowe metody analizy tekstu w badaniach humanistycznych i społecznych*

2. NAZWA W J. ANG.:

*Digital methods of text analysis in humanities and social science*

3. JĘZYK WYKŁADOWY: język polski

4. JEDNOSTKA PROWADZĄCA SZKOŁĘ LETNIĄ: Instytut Socjologii, Wydział Filozofii i Socjologii Uniwersytetu Warszawskiego (jednostka wiodąca), Instytut Języka Polskiego, Wydział Polonistyki Uniwersytetu Warszawskiego oraz konsorcjum CLARIN-PL

5. CELE KSZTAŁCENIA:

Celem szkoły jest wyposażenie doktorantów i doktorantek w specjalistyczne instrumentarium badawcze do analizy różnego typu danych tekstowych (artykułów prasowych, wywiadów, tekstów literackich, dokumentów urzędowych i archiwalnych, zapisów języka mówionego, zapisów komunikacji cyfrowej). W trakcie zajęć, prowadzonych przez doświadczonych badaczy, reprezentujących różne dyscypliny i patrzących na tekst z różnych perspektyw – językoznawców, socjologów, literaturoznawców, informatyków i filologów – uczestnicy i uczestniczki poznają wiodące w danej dyscyplinie metodologie analizy tekstu i uzupełnią arsenal metod i narzędzi dobrze ugruntowanych w naukach humanistycznych i społecznych (m.in. analiza semantyczno-pragmatyczna czy analiza treści) o metody i narzędzia cyfrowe (m.in. analiza korpusowa i text mining). Zostaną również wprowadzeni/wprowadzone w podstawy programowania w języku Python. Nacisk położony zostanie z jednej strony na rozwój kompetencji analitycznych, z drugiej zaś – na zrozumienie specyfiki danych tekstowych, w tym zwłaszcza tzw. danych nieustrukturyzowanych. Poza zaawansowaną wiedzą teoretyczną i metodologiczną uczestnicy/uczestniczki uzyskają umiejętność łączenia różnych metodologii, dostosowania metody do problemu badawczego (*problem-driven research*) oraz – co niezwykle istotne z perspektywy publikacji naukowych – prowadzenia badań w sposób uporządkowany i systematyczny.

6. POZIOM KSZTAŁCENIA ZGODNY Z PRK: 8 PRK
7. NAKŁAD PRACY UCZESTNIKA SZKOŁY LETNIEJ WYRAŻONY W PKT ECTS:  
6 ECTS
8. EFEKTY KSZTAŁCENIA DLA SZKOŁY LETNIEJ:

Efekty kształcenia	Odniesienie do charakterystyk PRK
<b>WIEDZA</b>	
Doktorant/Doktorantka po ukończeniu szkoły:	
zna i rozumie założenia metodologiczne stojące za wybranymi metodami analizy danych tekstowych;	P8S_WG, P8S_WK
posiada zaawansowaną wiedzę na temat wybranych narzędzi służących do pracy z danymi tekstowymi;	P8S_WG, P8S_WK
zna i rozumie mechanizmy działania poszczególnych narzędzi, ich możliwości, ograniczenia i obszary zastosowania;	P8S_WG, P8S_WK
posiada pogłębioną wiedzę z zakresu metod statystycznych wykorzystywanych w lingwistyce korpusowej;	P7S_WG
zna podstawy programowania w języku Python oraz sposoby jego wykorzystania w przygotowaniu, przetwarzaniu i przeprowadzeniu analiz danych tekstowych.	P7S_WG
<b>UMIEJETNOŚCI</b>	
Doktorant/Doktorantka po ukończeniu szkoły:	
potrafi wykorzystywać wiedzę teoretyczną i metodologiczną z zakresu analizy tekstu do identyfikowania, formułowania i rozwiązywania problemów badawczych;	P8S_UW
dysponuje umiejętnością krytycznego doboru metody analizy oraz selekcji danych o charakterze tekstowym umożliwiających opracowanie wybranego problemu badawczego;	P8S_UW, P8S_UO
doskonali umiejętność poruszania się pomiędzy odmiennymi metodologiami analizy danych tekstowych;	P8S_UW, P8S_UU
potrafi wykorzystywać dostępne zasoby (m.in. korpusy i wyszukiwarki korpusowe, skrypty, aplikacje) związane z analizą tekstu i przetwarzaniem języka naturalnego;	P8S_UW, P8S_UO
potrafi posługiwać się wybranymi narzędziami cyfrowymi do przetwarzania języka, korzystając z ich zaawansowanych funkcji;	P8S_UW
posiada umiejętność pozyskania danych tekstowych oraz ich przetworzenia na potrzeby analizy ilościowej i jakościowej z wykorzystaniem specjalistycznych narzędzi;	P8S_UW
potrafi wykorzystać język Python jako środowisko programistyczne do analizy danych tekstowych.	P8S_UW

<b>KOMPETENCJE SPOŁECZNE:</b>	
Doktorant/Doktorantka po ukończeniu szkoły:	
jest gotów/gotowa do pracy zespołowej i współpracy, również międzydziedzinowej, w ramach prowadzonych badań;	P8S_KK
jest gotów/gotowa do krytycznej oceny własnego projektu badawczego w świetle wybranych metodologii analizy tekstu wypracowanych na gruncie humanistyki cyfrowej i cyfrowych nauk społecznych;	P8S_KK
jest gotów/gotowa do krytycznej oceny podejść badawczych wypracowanych w ramach reprezentowanej przez siebie dyscypliny naukowej.	P8S_KK

## 9. UZASADNIENIE URUCHOMIENIA SZKOŁY LETNIEJ W ODNIESIENIU DO:

### a) GRUPY DOCELOWEJ:

Szkoła letnia stanowi odpowiedź na wyzwania związane z postępującym procesem cyfryzacji, gwałtowny przyrost danych o charakterze tekstowym, trudności związane z analizą tzw. danych nieustrukturyzowanych oraz rozwój infrastruktury naukowej, otwierającej nowe możliwości pogłębionej analizy tekstu. Doktoranci/doktorantki w swojej pracy badawczej coraz częściej wykorzystują duże zbiory tekstów (artykułów prasowych, zapisów debat politycznych czy dyskusji w mediach społecznościowych). Cyfrowe metody analizy tekstu mogą być stosowane w badaniach realizowanych w ramach niemal wszystkich dyscyplin humanistycznych i społecznych, stanowiąc doskonałe uzupełnienie tradycyjnego warsztatu badawczego. Jest to pole nowe, rozwijające się niezależnie w różnych dziedzinach nauki, zgodnie z właściwymi im podejściami. Aby w pełni wykorzystać ich potencjał badawczy, niezbędne jest połączenie kompetencji z wielu dyscyplin i rozwinięcie umiejętności sprawnego poruszania się pomiędzy odmiennymi metodologiami.

Szkoła pozwoli uczestnikom/uczestniczkom na uzyskanie zaawansowanej wiedzy na temat potencjału różnego rodzaju danych tekstowych i umiejętności systematycznej analizy dużych zbiorów tekstów przy użyciu specjalistycznych narzędzi. Doktoranci/doktorantki nauczą się, jak w sposób uporządkowany i przemyślany pozyskiwać, a następnie wykorzystywać dane tekstowe. Poznają założenia teoretyczne poszczególnych metodologii analizy tekstu, dowiedzą się, jakie pytania badawcze się z nimi łączą, nauczą się również, jak dobierać dane, co w danym przypadku stanowi jednostkę analizy i jak powinien przebiegać proces interpretacji wyników. Jako platforma umożliwiająca wymianę doświadczeń między badaczami z różnych obszarów nauki szkoła może stać się podstawą do nawiązania współpracy interdyscyplinarnej, niezbędnej do prowadzenia pogłębionych analiz tekstu.

## b) OFERTY DYDAKTYCZNEJ UW:

Obecna oferta UW w zakresie zajęć dotyczących analizy tekstu jest stosunkowo wąska i nie odpowiada rosnącemu zapotrzebowaniu na zajęcia wysokospecjalistyczne, kształcące badaczy wyposażonych w wiązkę kompetencji z różnych dziedzin, niezbędnych do projektowania i realizacji badań z pogranicza nauk społecznych, humanistycznych i świata nowych technologii.

Po pierwsze, brakuje zajęć dotyczących narzędzi, które pozwalają w sposób uporządkowany analizować jakościowo teksty, np. Atlas.ti czy MAXQDA. Doktoranci nie uczą się metodologii komputerowo wspomaganey analizy danych jakościowych (ang. Computer-Assisted Qualitative Data Analysis). Brakuje zwłaszcza wiedzy na temat wskaźników zgodności kodowania, których zastosowanie jest wymagane do publikacji artykułów naukowych z wykorzystaniem tego rodzaju analiz.

Po drugie, nie oferuje się zaawansowanych zajęć z metodologii analizy jakościowej i ilościowej dużych zbiorów danych tekstowych. Problem ten dotyczy zarówno nauk humanistycznych, jak i społecznych.

Po trzecie, brakuje zajęć na temat badań korpusowych, skierowanych do doktorantów. Kursów dla studentów studiów licencjackich i magisterskich jest bardzo niewiele i są prowadzone z reguły na kierunkach lingwistycznych – przez i dla językoznawców. Nie ma również kursów na temat text mining, a więc grupy metod służących do eksploracji danych tekstowych, które są intensywnie rozwijane w zachodnich ośrodkach badawczych.

Po czwarte, dostępne zajęcia dotyczące metod analizy tekstu ograniczone są zwykle do prezentacji wybranych narzędzi (ich repertuar jest z reguły bardzo ograniczony) i nie oferują całościowej metodologii badań zorientowanych na analizę tekstów oraz namysłu teoretycznego nad specyfiką poszczególnych podejść badawczych. Sprzyja to prowadzeniu badań w paradygmacie *method-driven research* i utwierdza ich dominację nad badaniami nastawionymi na rozwiązywanie problemów badawczych (*problem-driven research*).

Szkoła letnia pozwoli wypełnić wskazane luki. Kurs rozpocznie się od prezentacji perspektywy językoznawczej, sięgającej w głąb tekstu i skupionej analizie jego elementów leksykalnych. Następnie uczestnicy/uczestniczki dowiedzą się, jak spojrzeć na słowa (i język) w szerszym kontekście i poznają metody służące do jakościowej analizy większych zbiorów danych tekstowych wykorzystywane w naukach społecznych. Kolejnym krokiem będzie zaprezentowanie metodologii lingwistyki korpusowej i komputerowej oraz text mining, nauka

podstaw programowania oraz posługiwania się narzędziami cyfrowymi (webowymi i desktopowymi), które pozwalają na ilościową analizę wielu tekstów jednocześnie. Każdy blok zajęciowy będzie obejmował prezentację danej metody wraz z omówieniem jej teoretycznych założeń, zakresu możliwych zastosowań i dostępnych narzędzi do analizy danych. W celu pokazania uczestnikom/uczestniczkom, jak w praktyce wygląda proces badawczy *problem-driven research* poszczególne metody zostaną przedstawione na konkretnym materiale i w odniesieniu do konkretnego pytania badawczego.

## 10. WYKAZ PRZEDMIOTÓW PRZOWADZONYCH W RAMACH SZKOŁY LETNIEJ:

Nazwa przedmiotu	Forma zajęć	Efekty kształcenia	Metody dydaktyczne wykorzystywane podczas zajęć
<b>Analiza leksykalno-semantyczna</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna podstawy semantyki (teoria znaczenia, struktura znaczenia wraz z elementami ocennymi i ekspresywnymi);</li> <li>– posiada umiejętność analizy leksykalno-semantycznej (z elementami analizy gramatycznej) tekstów;</li> <li>– identyfikuje nośniki sensu na różnych poziomach systemu (od morfologicznego, przez leksykalny, do syntaktycznego);</li> <li>– potrafi wykonać analizę struktury znaczeniowej jednostek leksykalnych.</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach</p> <p>Udział dwóch prowadzących wynika z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów doktorantów na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego i koordynacji pracy w grupach.</p>
<b>Analiza semantyczno-pragmatyczna</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna najważniejsze teorie pragmatyczne (teorie dotyczące funkcji języka, teorię aktów mowy, teorię implikatur konwersacyjnych Grice’a i in.);</li> <li>– dostrzega wpływ kontekstu (rozumianego szeroko – jako kontekst leksykalny, kontekst genologiczny, kontekst intertekstualny, kontekst dyskursowy,</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach</p> <p>Udział dwóch prowadzących wynika z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów doktorantów na bieżąco, udzielania przez prowadzących</p>

		<p>kontekst pozajęzykowy itd.) na sens i wartość illokucyjną jednostki w tekście;</p> <ul style="list-style-type: none"> <li>– potrafi przeprowadzić analizę semantyczno-pragmatyczną, zmierzającą do rozpoznania niektórych typów specyficznych działań społeczno-językowych (agresji językowej, manipulacji, mowy nienawiści itd.).</li> </ul>	indywidualnego wsparcia merytorycznego i koordynacji pracy w grupach.
<b>CAQDAS: analiza treści z wykorzystaniem Atlas.ti</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna i rozumie założenia wspomaganą komputerowo analizy treści;</li> <li>– potrafi wykorzystać program Atlas.ti do analizy treści;</li> <li>– potrafi opracować klucz kodowy oraz zakodować materiał tekstowy;</li> <li>– zna i rozumie znaczenie wskaźników zgodności kodowania;</li> <li>– zna zaawansowane funkcje programu Atlas.ti służące do wizualizacji, rekonstrukcji pól semantycznych oraz ram tematów i przygotowania biblioteki projektu do pracy z danymi;</li> <li>– potrafi połączyć ilościową i jakościową analizę tekstu.</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z wykorzystaniem programu Atlas.ti</p> <p>Udział dwojga prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektyw socjologicznej i językoznawczej oraz z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.</p>
<b>Wprowadzenie do lingwistyki korpusowej</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna podstawy teoretyczne lingwistyki korpusowej oraz najważniejsze zagadnienia metodologiczne związane z badaniami korpusowymi;</li> </ul>	Warsztaty w formule <i>hands-on workshop</i> , obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach,

		<ul style="list-style-type: none"> <li>– posługuje się pojęciami wykorzystywanymi w lingwistyce korpusowej (m.in. korpus ujmowany jako teoria i metoda, token, lemat, parsowanie);</li> <li>– posiada podstawową wiedzę dotyczącą rodzajów korpusów, historii analiz korpusowych oraz głównych kierunków współczesnych badań w tym zakresie;</li> <li>– rozumie rozróżnienie na badania <i>corpus-driven</i> i <i>corpus-based</i>.</li> </ul>	<p>poprzedzone wprowadzeniem z wykorzystaniem technik multimedialnych.</p> <p>Udział dwóch prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektyw językoznawczej i socjologicznej oraz z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.</p>
<b>Praca z korpusami</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna najważniejsze korpusy języka polskiego i wie, w jaki sposób można je wykorzystać we własnych badaniach;</li> <li>– potrafi wymienić najważniejsze narzędzia do pracy z korpusami w języku polskim i angielskim (ich wady, zalety i zakres stosowania);</li> <li>– potrafi korzystać z wyszukiwarek korpusowych w języku polskim i angielskim;</li> <li>– zna podstawy składni języka CQL;</li> <li>– umie posługiwać się wyrażeniami regularnymi.</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z wykorzystaniem specjalistycznego oprogramowania, poprzedzone wprowadzeniem z wykorzystaniem technik multimedialnych</p> <p>Udział dwojga prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektywy <i>stricte</i> programistycznej i szerszego spojrzenia humanistycznego oraz z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących</p>



			indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.
<b>Projektowanie korpusów</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna metody określania potrzebnej wielkości korpusu w zależności od jego przeznaczenia (uczenie maszynowe, anotacja manualna itp.);</li> <li>– potrafi dobrać korpus do problemu badawczego;</li> <li>– dostrzega i rozumie wpływ budowy korpusu na rezultaty analiz;</li> <li>– zna wiodące standardy opisu tekstów metadanymi i potrafi dobrać je do korpusu.</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach</p> <p>Udział dwojga prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektywy socjologicznej i językoznawczej oraz z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.</p>
<b>Przetwarzanie tekstów</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna podstawowe etapy czyszczenia i przetwarzania danych tekstowych, konieczne do uzyskania korpusu wysokiej jakości;</li> <li>– potrafi pozyskiwać i wydobywać tekst z plików zapisanych w różnych formatach;</li> <li>– potrafi zastosować narzędzia do automatycznego czyszczenia tekstu, tokenizacji, lematyzacji, rozpoznawania części mowy i parsowania;</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z wykorzystaniem języka Python</p> <p>Udział dwojga prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektywy <i>stricte</i> programistycznej i szerszego spojrzenia humanistycznego oraz z faktu, że ćwiczenia wykonywane podczas</p>

		<ul style="list-style-type: none"> <li>– zna podstawy programowania w języku Python oraz sposoby jego wykorzystania w przygotowaniu i przetwarzaniu danych tekstowych;</li> <li>– potrafi wykorzystać język Python jako środowisko programistyczne do analizy danych tekstowych.</li> </ul>	warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.
<b>Niejednoznaczność: polisemia i homonimia</b>	wykład	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– rozumie niejednoznaczność jako zjawisko tekstowe;</li> <li>– zna zjawisko polisemii regularnej i potrafi wyróżnić jej główne typy;</li> <li>– zna i rozumie założenia oraz podstawowe metody dezambiguacji znaczeń w procesie automatycznego przetwarzania korpusów: dehomonimizację i desynkretyzację, a także ujednoznacznianie fleksyjne, morfologiczne, składniowe, przez relacje leksykalne, przez kryteria selekcji.</li> </ul>	Metody charakterystyczne dla wykładu (z wykorzystaniem technik multimedialnych)
<b>Analiza metafor w tekście na przykładzie metafor synestezyjnych</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna i rozumie założenia analizy metafory w dyskursie z wykorzystaniem teorii ram semantycznych Ch. Fillmore’a;</li> <li>– dostrzega problemy związane z definicją metafory i wyznaczeniem granic między użyciem dosłownym i przenośnym;</li> <li>– identyfikuje i rozumie właściwości metafor w tekście (kontaminacje, łańcuchy,</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z użyciem korpusu Synamet</p> <p>Udział dwóch prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektywy językoznawczej i literaturoznawczej oraz z faktu, że ćwiczenia wykonywane</p>

		<p>współwystępowanie z metonimią i hypellage), które utrudniają anotację;</p> <ul style="list-style-type: none"> <li>– potrafi posługiwać się korpusem Synamet;</li> <li>– potrafi interpretować wyniki analiz statystycznych metafor w większej bazie danych.</li> </ul>	<p>podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.</p>
<b>Podstawy statystyki w lingwistyce korpusowej</b>	wykład	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna podstawowe metody statystyczne wykorzystywane w lingwistyce korpusowej oraz zakres ich użycia;</li> <li>– dostrzega i rozumie znaczenie metod statystycznych w analizie tekstu oraz ich wpływ na sposób interpretacji wyników.</li> </ul>	<p>Metody charakterystyczne dla wykładu (z wykorzystaniem technik multimedialnych)</p>
<b>Automatyczne metody ekstrakcji treści</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna metody automatycznego rozpoznawania oraz wydobywania informacji z tekstu;</li> <li>– zna i rozumie ogólne mechanizmy działania narzędzi do automatycznej ekstrakcji treści, a także możliwości ich zastosowania w badaniach humanistycznych i społecznych;</li> <li>– potrafi wykorzystywać aplikacje webowe i desktopowe służące do rozpoznawania nazw własnych, identyfikacji wyrażen przestrzennych, wydobywania terminologii, słów kluczowych, słownictwa charakterystycznego, leksykalnych jednostek wielowyrazowych itp.</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z wykorzystaniem specjalistycznego oprogramowania</p> <p>Udział dwojga prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektywy <i>stricte</i> programistycznej i szerszego spojrzenia humanistycznego oraz z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego,</p>

			koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.
<b>Korpusy mowy, korpusy monitorujące, ekstrakcja frazeologii. Zasoby, narzędzia, zastosowania</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna dostępne dla języka polskiego korpusy mowy oraz korpusy monitorujące;</li> <li>– potrafi zastosować narzędzia służące do przeszukiwania i ekstrakcji jednostek frazeologicznych z korpusów mowy i korpusów monitorujących;</li> <li>– zna i rozumie metody ekstrakcji pozycyjnej oraz relacyjnej;</li> <li>– potrafi posługiwać się narzędziem Treelets, które umożliwia wydobywanie frazeologii z dowolnego korpusu w kilkudziesięciu obsługiwanych językach.</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z wykorzystaniem specjalistycznego oprogramowania</p> <p>Udział dwojga prowadzących wynika z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.</p>
<b>WordNety w badaniach korpusowych</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna podstawy metodologiczne tworzenia i wykorzystania relacyjnych leksykalnych baz danych typu WordNet;</li> <li>– zna i rozumie metodę tworzenia wordnetów ze szczególnym uwzględnieniem polskiej bazy SłowoSieć;</li> <li>– dostrzega i rozumie możliwości zastosowania wordnetów w badaniach humanistycznych i społecznych z wykorzystaniem metod korpusowych.</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z wykorzystaniem specjalistycznego oprogramowania, poprzedzone wprowadzeniem z wykorzystaniem technik multimedialnych.</p> <p>Udział dwóch prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektyw <i>stricte</i> programistycznej i językoznawczej oraz z faktu, że ćwiczenia wykonywane</p>

			podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.
<b>Analiza wydźwięku emocjonalnego</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna i rozumie założenia stojące u podstaw metod automatycznej analizy oparte na mechanizmach zakładających wykorzystanie słowników, reguł semantyczno-składniowych i uczenia maszynowego, a także metody hybrydowe;</li> <li>– zna najważniejsze narzędzia służące do analizy wydźwięku emocjonalnego i dostrzega ich ograniczenia;</li> <li>– potrafi zastosować polskojęzyczne narzędzia do analizy wydźwięku (Wydźwięk i Sentemo).</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z wykorzystaniem specjalistycznego oprogramowania</p> <p>Udział dwojga prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektyw <i>stricte</i> programistycznej i językoznawczej oraz z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.</p>
<b>Analiza stylometryczna</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna i rozumie teoretyczne i statystyczne podstawy stylometrii oraz jej potencjalne zastosowania badawcze;</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z wykorzystaniem specjalistycznego oprogramowania</p>

		<ul style="list-style-type: none"> <li>– zna i rozumie metody analizy podobieństwa treści, metody analizy podobieństwa języka, metody grupowania tekstów podobnych treściowo lub językowo;</li> <li>– dostrzega możliwości, ograniczenia oraz warunki brzegowe wynikające z zastosowanych metod;</li> <li>– potrafi zastosować polskojęzyczne narzędzia do analizy stylometrycznej (LEM i WebSty).</li> </ul>	<p>Udział dwóch prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektyw <i>stricte</i> programistycznej i literaturoznawczej oraz z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.</p>
<b>Modelowanie tematyczne</b>	warsztaty	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– dostrzega podobieństwa i różnice pomiędzy analizą stylometryczną a modelowaniem tematycznym;</li> <li>– zna i rozumie techniczne aspekty modelowania tematycznego, metody wyznaczania markerów tematycznych, sposoby grupowania tekstów podobnych oraz rozdzielania grup tekstów różnych;</li> <li>– potrafi posługiwać się narzędziem Topic oraz funkcją platformy WebSty: „Grupy podobieństwa treści” oraz zna mechanizmy ich działania.</li> </ul>	<p>Warsztaty w formule <i>hands-on workshop</i>, obejmujące analizę materiałów tekstowych dostarczonych przez prowadzących, pracę indywidualną oraz pracę w grupach z wykorzystaniem specjalistycznego oprogramowania</p> <p>Udział dwóch prowadzących wynika z chęci uwzględnienia w treści i sposobie prowadzenia zajęć perspektyw <i>stricte</i> programistycznej i literaturoznawczej oraz z faktu, że ćwiczenia wykonywane podczas warsztatów wymagać będą monitorowania postępów uczestników na bieżąco, udzielania przez prowadzących indywidualnego wsparcia merytorycznego, koordynacji pracy w grupach, a także rozwiązywania ewentualnych problemów technicznych.</p>

<p><b>Wykłady i dyskusje z praktykami</b></p>	<p>wykład, seminarium</p>	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– zna możliwe zastosowania poszczególnych metod analizy tekstu, w tym zwłaszcza analizy leksykalno-semantycznej, analizy korpusowej oraz text mining, w zróżnicowanych badaniach humanistycznych i społecznych;</li> <li>– jest gotów/gotowa do krytycznej oceny poszczególnych podejść badawczych;</li> <li>– dostrzega problemy związane z tworzeniem korpusów i wie, jak ich uniknąć we własnych badaniach.</li> </ul>	<p>Cykl czterech spotkań połączonych z dyskusją. W części wykładowej eksperci-praktycy przedstawią założenia, metody analizy oraz wyniki prowadzonych przez siebie badań. Uczestnicy będą mieli możliwość zadawania pytań oraz skonsultowania z wykładowcą własnych projektów badawczych. W części seminaryjnej odbędzie się dyskusja na temat zastosowania danej metody, jej zalet i ograniczeń.</p>
<p><b>Sesje pracy grupowej</b></p>	<p>ćwiczenia</p>	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– potrafi wykorzystywać wiedzę teoretyczną i metodologiczną zdobytą w trakcie zajęć do rozwiązywania konkretnych problemów badawczych;</li> <li>– dysponuje umiejętnością krytycznego doboru metody analizy adekwatnej do wybranego problemu badawczego;</li> <li>– jest gotów/gotowa do pracy zespołowej i współpracy, również międzydziedzinowej, w ramach prowadzonych badań.</li> </ul>	<p>Praca grupowa nad projektami badawczymi pod opieką dwojga prowadzących</p> <p>Udział dwojga prowadzących wynika z chęci uwzględnienia perspektywy humanistycznej i społecznej oraz z potrzeby koordynacji pracy w grupach, monitorowania postępów, udzielania wsparcia merytorycznego, a także rozwiązywania ewentualnych problemów technicznych.</p>
<p><b>Prezentacja projektów i konsultacje z prowadzącymi</b></p>	<p>seminarium</p>	<p>Doktorant/Doktorantka:</p> <ul style="list-style-type: none"> <li>– potrafi wykorzystywać wiedzę teoretyczną i metodologiczną zdobytą w trakcie zajęć we własnej pracy badawczej;</li> </ul>	<p>Członkowie i członkinie poszczególnych zespołów przedstawią opracowane procedury badawcze oraz omówią je z prowadzącymi specjalizującymi się w danej metodologii. Będą mieli także możliwość skorzystania z indywidualnych</p>

		<ul style="list-style-type: none"><li>– jest gotów/gotowa do krytycznej oceny własnego projektu badawczego w świetle poznanych metodologii analizy tekstu wypracowanych na gruncie humanistyki cyfrowej i cyfrowych nauk społecznych;</li><li>– jest gotów/gotowa do krytycznej oceny podejść badawczych wypracowanych w ramach reprezentowanej przez siebie dyscypliny naukowej;</li><li>– jest gotów/gotowa do pracy zespołowej i współpracy, również międzydziedzinowej, w ramach prowadzonych badań.</li></ul>	<p>konsultacji w zakresie prowadzonych przez siebie badań.</p> <p>W seminarium uczestniczyć będzie sześcioro prowadzących reprezentujących różne dyscypliny naukowe i patrzących na tekst z różnych perspektyw. Dzięki temu doktoranci/doktorantki będą mogli/mogły skonsultować swoje projekty badawcze z ekspertami w danym polu.</p>
--	--	--	---



## 11. LITERATURA PRZEDMIOTU OBLIGATORYJNA:

Każdy z prowadzących przygotowuje swój oddzielny sylabus, w którym uwzględni literaturę obowiązkową i zalecaną. Bibliografia zostanie przekazana uczestnikom/uczestniczkom po zakończeniu procesu rekrutacyjnego.

## 12. LITERATURA PRZEDMIOTU ZALECANA:

Patrz: punkt 11.

## 13. PLAN ZAJĘĆ

W osobnym pliku

## 14. METODY WERYFIKACJI EFEKTÓW KSZTAŁCENIA DLA SZKOŁY LETNIEJ

*(w tym sposób przeprowadzenia bilansu kompetencji)*

Weryfikacja osiągnięcia efektów kształcenia w zakresie wiedzy przeprowadzona zostanie za pomocą dwóch testów pisemnych. Weryfikacja osiągnięcia efektów kształcenia w zakresie umiejętności i kompetencji społecznych przeprowadzona zostanie za pomocą mini projektów grupowych. Uwzględniony zostanie także czynny udział w dyskusjach, które odbędą się w ramach zajęć z ekspertami-praktykami.

Pierwszego dnia szkoły przeprowadzony zostanie test składający się z pytań zamkniętych, sprawdzający wiedzę uczestników i uczestniczek na temat metod analizy tekstu. Ostatniego dnia podejną oni/one do tego samego testu, dzięki czemu możliwe będzie porównanie ich stanu wiedzy w momencie rozpoczęcia i ukończenia szkoły letniej. Do uzyskania pozytywnej oceny niezbędne będzie udzielenie co najmniej 60% poprawnych odpowiedzi.

Projekty grupowe polegać będą na opracowaniu procedury analizy wskazanych danych tekstowych przy użyciu metod poznanych w trakcie zajęć. Na pierwszych zajęciach doktoranci i doktorantki zostaną podzieleni na zespoły składające się z 3–4 osób. Przy podziale na zespoły uwzględnione zostaną zainteresowania naukowe uczestników/uczestniczek oraz rodzaj materiału, jaki wykorzystują w swoich badaniach. Każdy zespół będzie pracował na własnym zbiorze danych tekstowych. Zadanie będzie polegało na opracowaniu procedury analizy danych zgodnie z metodologią omawianą w trakcie zajęć. Ostatniego dnia szkoły letniej członkowie/członkinie poszczególnych zespołów przedstawią opracowane procedury analizy pozostałym doktorantom oraz omówią je z prowadzącymi specjalizującymi się w danej

metodologii. Pisemny konspekt z każdego z wystąpień zostanie zarchiwizowany i załączony do raportu końcowego ze szkoły letniej.

Bilans kompetencji sporządzony zostanie na podstawie: informacji podanych przez doktorantów w formularzu rekrutacyjnym, porównania wyników testów pisemnych, przebiegu sesji pracy grupowej, prezentacji projektów grupowych oraz wyników ankiet ewaluacyjnych, w których uwzględnione zostaną m.in. pytania dotyczące przyrostu wiedzy, umiejętności i kompetencji niezbędnych do dalszego prowadzenia badań naukowych oraz stopnia realizacji oczekiwań, jakie uczestnicy/uczestniczki szkoły przedstawili w liście motywacyjnym.

#### 15. SYLWETKA KANDYDATA/KOMPETENCJE WYMAGANE NA STARCIE

*(do kogo szkoła jest adresowana, jakich kompetencji oczekuje się od kandydatów)*

Szkoła adresowana jest do doktorantów i doktorantek wszystkich lat studiów, reprezentujących dyscypliny humanistyczne i społeczne (literaturoznawstwo, językoznawstwo, historia, kulturoznawstwo, antropologia, archeologia, filologie, filozofia, politologia, socjologia, pedagogika, psychologia, prawo, nauki o komunikacji), wykorzystujących w swojej pracy naukowej szeroko rozumiane dane tekstowe (od wywiadów przez media społecznościowe do powieści). Dla słuchaczy początkowych lat studiów udział w szkole może stanowić inspirację do opracowania metodologii badań prowadzonych w ramach rozprawy doktorskiej, z kolei słuchacze lat późniejszych będą mieli szansę poszerzyć swój repertuar analityczny.

Warunkiem uczestnictwa w Szkole Letniej jest posiadanie statusu doktoranta/ki Uniwersytetu Warszawskiego, zainteresowanie problematyką jakościowej i/albo ilościowej analizy tekstu i nowoczesnymi metodami badawczymi wykorzystywanymi w naukach humanistycznych i społecznych, znajomość języka polskiego w stopniu umożliwiającym aktywne uczestnictwo w zajęciach, w tym znajomość gramatyki języka polskiego na poziomie co najmniej egzaminu maturalnego, znajomość języka angielskiego na poziomie co najmniej B2, a także znajomość obsługi komputera w stopniu umożliwiającym sprawne korzystanie z aplikacji webowych i desktopowych do analizy tekstu. Podstawowa wiedza i doświadczenie badawcze w zakresie metod analizy tekstu są mile widziane. Wcześniejsze doświadczenie w pracy z narzędziami służącymi do przetwarzania języka naturalnego oraz znajomość języków programowania nie są jednak wymagane.

## 16. SZCZEGÓLNE ZASADY I SPOSÓB PRZEPROWADZENIA REKRUTACJI

(w tym wymagane dokumenty)

- i. Uczestnikiem/uczestniczką szkoły letniej może być wyłącznie doktorant/doktorantka Uniwersytetu Warszawskiego.
- ii. Doktorant/doktorantka zgłasza chęć uczestnictwa w szkole letniej poprzez wypełnienie formularza dostępnego pod adresem: <https://forms.gle/WiMta5vzd2QMkvgw8>.
- iii. Formularz obejmuje informacje na temat motywacji do udziału w szkole letniej oraz zainteresowań naukowych i doświadczeń badawczych kandydata/kandydatki, a także zgodę na przetwarzanie danych osobowych..
- iv. Minimalna liczba uczestników wynosi 15, a maksymalna – 18.
- v. Kandydaci będą przyjmowani zgodnie z kolejnością miejsc na liście rankingowej, przy czym pierwszeństwo będzie przysługiwało stypendystom programu „Zwiększanie mobilności doktorantów UW”.
- vi. Lista rankingowa zostanie utworzona na podstawie oceny zgłoszeń nadesłanych przez kandydatów. Zostanie ona przeprowadzona przez trzyosobową komisję złożoną z wykładowców szkoły letniej w stopniu co najmniej doktora wedle następujących kryteriów:
  - a. adekwatność prowadzonych badań oraz zainteresowań naukowych do celów oraz programu szkoły – od 0 do 10 punktów;
  - b. spodziewany wpływ udziału doktoranta w szkole letniej na jego rozwój naukowy, w szczególności na postępy w przygotowywaniu rozprawy doktorskiej – od 0 do 10 punktów.
- vii. Rejestracja do szkoły letniej rozpocznie się 6.05.2020 r. i zakończy 6.06.2020 r. Lista rankingowa kandydatów zostanie opublikowana na stronie: [www.zip.uw.edu.pl](http://www.zip.uw.edu.pl) najpóźniej do 20.06.2020 r.
- viii. O wynikach rekrutacji wszyscy kandydaci zostaną powiadomieni drogą mailową niezwłocznie po opublikowaniu listy rankingowej.
- ix. W przypadku niewypełnienia limitu miejsc organizatorzy przeprowadzą dodatkowe tury rekrutacji, o terminach których poinformują ogłoszeniem na stronie: [www.zip.uw.edu.pl](http://www.zip.uw.edu.pl).
- x. W przypadku rezygnacji uczestnika przed rozpoczęciem zajęć lub w pierwszym dniu zajęć do udziału w szkole dopuszcza się kandydata z pierwszego miejsca listy rezerwowej.

- xi. W przypadku nieprzyjęcia odpowiedniej liczby kandydatów Szkoła Letnia nie odbędzie się w wyznaczonym terminie, o czym osoby przyjęte zostaną poinformowane drogą mailową.
- xii. W przypadku konieczności przełożenia terminu szkoły letniej ze względu na sytuację epidemiologiczną o zmianie terminu osoby przyjęte zostaną poinformowane najpóźniej na tydzień przed inauguracją. Nowy termin szkoły letniej zostanie również ogłoszony na stronie: [www.zip.uw.edu.pl](http://www.zip.uw.edu.pl).

## 17. WARUNKI UKOŃCZENIA SZKOŁY LETNIEJ I UZYSKANIA DYPLOMU

Szkoła letnia kończy się zaliczeniem na ocenę. Warunkiem ukończenia szkoły i uzyskania dyplomu jest:

- i. obecność i aktywny udział w zajęciach (dopuszczalna jest nieobecność na 10% zajęć);
- ii. przystąpienie do testu końcowego i uzyskanie co najmniej 60% punktów;
- iii. aktywny udział w rozwiązywaniu zadań grupowych i prezentacji wyników wspólnej pracy podczas końcowych zajęć Szkoły Letniej.

## 18. WYKAZ NAUCZYCIELI AKADEMICKICH PROWADZĄCYCH ZAJĘCIA

WRAZ ZE WSKAZANIEM KOMPETENCJI NAUKOWO-DYDAKTYCZNYCH.

**dr hab. Magdalena Derwojedowa** – językoznawca, adiunkt w Instytucie Języka Polskiego UW. Kierowała projektem „Automatyczna analiza fleksyjna tekstów polskich z lat 1830–1918 z uwzględnieniem zmian w odmianie i pisowni” (OPUS NCN), w ramach którego opracowany został m.in. analizator morfologiczny „Morfeusz XIX-wieczny” oraz korpus tekstów z lat 1830–1918. Brała udział w pracach nad pierwszą wersją Słownosieci oraz pilotażowym polskim FrameNetem (RAMKI). Jej doświadczenie dydaktyczne obejmuje językoznawstwo ogólne, metodologię językoznawstwa, semantykę leksykalną, morfologię i składnię języka polskiego oraz nauczanie języka polskiego jako obcego.

**dr Agata Hącia** – językoznawca, wykładowca UW z ponad 15-letnim doświadczeniem. Specjalizuje się w zagadnieniach kultury języka, semantyki i leksykologii. Promotor licznych prac licencjackich oraz promotor pomocniczy rozprawy doktorskiej, obronionej w IJP UW z wyróżnieniem. Zastępca kierownika grantu „Postawy wobec języka” w ramach programu Obserwatorium Kultury MKiDN (2016–2017) – przy UW. Współtwórczyni (2011–2015) ogólnopolskich badań edukacyjnych (z ramienia IBE). Tworzy i wdraża programy poprawy jakości komunikacji w firmach i instytucjach (z ramienia PRO.PL). Szefowa interdyscyplinarnego zespołu przy PRO.PL, gromadzącego pracowników naukowych UW, UJ,

WUM, PAN, opracowującego kwalifikacje rynkowe i potrzebne społecznie związane z różnymi aspektami komunikacji. Założycielka i prezes Fundacji PRO.PL, prowadzącej prace naukowe (badawcze) i popularyzatorskie. Wieloletnia współprowadząca poradnię językową Uniwersytetu Warszawskiego. Członkini jury Olimpiady Literatury i Języka Polskiego. Ścisłe współpracuje m.in. z Radą Języka Polskiego przy Prezydium PAN. Popularyzatorka wiedzy o języku polskim m.in. jako autorka książek i wieloletnia stała współpracownica Polskiego Radia.

**mgr Agnieszka Karlińska** – socjolożka, literaturoznawczyni i kulturoznawczyni, absolwentka Kolegium MISH UW, doktorantka w Instytucie Socjologii oraz Instytucie Literatury Polskiej UW. Umiejętności z zakresu przetwarzania języka naturalnego i text mining doskonalila na licznych kursach i szkołach letnich, m.in. w Oxfordzie i Heidelbergu. Wykorzystuje je do analizy różnego rodzaju danych tekstowych, od XIX-wiecznych powieści kryminalnych przez tweety po teksty prawne/prawnicze i medyczne. Artykuły z zakresu analizy tekstu przy użyciu metod cyfrowych publikowała m.in. w „Digital Scholarship in the Humanities” i w „Studiach Socjologiczno-Politycznych”. Kompetencje dydaktyczne rozwijała, prowadząc zajęcia w Instytucie Socjologii, Instytucie Literatury Polskiej oraz w ramach Uniwersytetu Otwartego UW. Obecnie prowadzi zajęcia z Analizy danych ilościowych z R i Text mining dla studentów kierunku Socjologia cyfrowa. W roku akademickim 2020/2021 współprowadzić będzie autorski kurs Warsztaty z analizy komunikacji cyfrowej i danych tekstowych: narzędziownik, finansowany ze środków Funduszu Innowacji Dydaktycznych UW.

**dr hab. Katarzyna Kłosińska** – językoznawca, adiunkt w Instytucie Języka Polskiego UW. Jej zainteresowania badawcze koncentrują się wokół semantyki i leksykologii, analizy dyskursu, kultury języka polskiego, normatywistyki teoretycznej, socjolingwistyki, polityki językowej i lingwistyki kulturowej. Kierowała projektem „Postawy wobec języka” w ramach programu Obserwatorium Kultury MKiDN (2016–2017). W latach 1999–2018 pełniła funkcję sekretarza naukowej Rady Języka Polskiego przy Prezydium PAN, a od roku 2019 jest jej przewodniczącą. Zajmuje się również popularyzacją wiedzy o języku. Od 2015 roku prowadzi internetową poradnię językową. Posiada 30-letnie doświadczenie w pracy naukowej i dydaktycznej na UW. Prowadziła i prowadzi zarówno zajęcia kursowe, jak i konwersatoria oraz seminaria (licencjackie, magisterskie, doktorskie). Obecnie ma pod opieką troje doktorantów, wypromowała kilkadziesiąt licencjatów i kilkanaście prac magisterskich.

**dr inż. Łukasz Kobyliński** – adiunkt w Zespole Inżynierii Lingwistycznej Instytutu Podstaw Informatyki PAN, członek zespołu CLARIN-PL. Specjalizuje się w zastosowaniu metod uczenia maszynowego w przetwarzaniu języka naturalnego. W swojej pracy badawczej zajmował się zagadnieniami dezambiguacji znaczenia słów, znakowaniem morfosyntaktycznym, a także lingwistyką korpusową. Kierował projektem badawczym „Automatyczne wykrywanie i korekcja błędów anotacyjnych w polskich korpusach językowych” (Preludium NCN). Współautor oprogramowania Korpusomat. Prowadzi liczne warsztaty i wykłady z zakresu przetwarzania języka naturalnego.

**dr inż. Jan Kocoń** – pracownik Katedry Inteligencji Obliczeniowej PWr, członek zespołu CLARIN-PL oraz Grupy Technologii Językowych G4.19, programista i projektant odpowiedzialny za tworzenie narzędzi do analizy wydźwięku emocjonalnego. Prowadził i współorganizował szereg warsztatów, prelekcji, seminariów dotyczących maszynowego przetwarzania języka. Pracownik dydaktyczny prowadzący od kilku lat zajęcia kursowe ze studentami.

**mgr Małgorzata Krawentek** – doktorantka w Instytucie Anglistyki UŁ, Zakład Pragmatyki. Tematyka jej pracy badawczej obejmuje językoznawstwo korpusowe i komputerowe, analizę dyskursu, (nie)grzeczność językową i dyskurs filmowy. Jest asystentem w zespole CLARIN-PL w UŁ, gdzie zajmuje się m.in. korpusami mowy i korpusami równoległymi.

**dr hab. Sławomir Mandes** – socjolog, adiunkt w Instytucie Socjologii UW, współtwórca kierunku Socjologia cyfrowa. Był zaangażowany, jako kierownik i wykonawca, w międzynarodowe badania o charakterze ilościowym i jakościowym, w tym badania wykorzystujące wspomaganą komputerowo analizę treści. Członek zespołu w projekcie „Postawy wobec języka” realizowanego w ramach programu Obserwatorium Kultury MKiDN (2016–2017). Posiada wieloletnie doświadczenie dydaktyczne z zakresu metodologii badań społecznych, teorii komunikacji i socjologii języka. Prowadzi zajęcia z jakościowo-ilościowej analizy treści z wykorzystaniem programu Atlas.ti oraz zajęcia z *text mining* w Instytucie Socjologii UW.

**dr hab. Małgorzata Marciniak** – zatrudniona w Instytucie Podstaw Informatyki PAN, członkini zespołu CLARIN-PL, programistka i projektantka odpowiedzialna za powstanie systemu detekcji i ekstrakcji terminologii z tekstu TermoPL. Prowadziła szereg warsztatów, prelekcji, seminariów dotyczących maszynowego przetwarzania języka.

**dr Maciej Maryl** – literaturoznawca, socjolog, adiunkt w Instytucie Badań Literackich Polskiej Akademii Nauk, kierownik Centrum Humanistyki Cyfrowej IBL PAN, Ambasador CLARIN-ERIC. Interesuje się komunikacją literacką, nowymi mediami, piśmiennictwem multimedialnym, edytorstwem elektronicznym, humanistyką cyfrową oraz związkami technologii z kulturą. Z zespołem CLARIN-PL współpracuje przy projektach dotyczących analizy genologicznej tekstów multimedialnych (grupowanie blogów), retrokonwersji danych bibliograficznych oraz rozwoju Literackiego Eksploratora Maszynowego (LEM) w wersji polskiej i wielojęzycznej. Jest przedstawicielem IBL PAN w Radzie konsorcjum DARIAH-PL, w Core Group konsorcjum OPERAS i forum CLE. Bierze udział w pracach ALLEA E-humanities Working Group, DARIAH Digital Methods and Practices Observatory oraz kieruje grupą roboczą DARIAH-PL Filologia Cyfrowa. Zajmuje się również popularyzacją humanistyki cyfrowej poprzez organizację warsztatów i wykładów w Polsce i za granicą.

**dr hab. Piotr Pęzik** – adiunkt w Instytucie Anglistyki UŁ, autor prac naukowych i rozwiązań informatycznych z dziedziny językoznawstwa korpusowego i komputerowego. Obszarem jego zainteresowań językoznawczych jest frazeologia dystrybucyjna w ujęciu korpusowym. Był wykonawcą i kierownik krajowych i europejskich projektów badawczo-rozwojowych, m.in. NKJP, CESAR, PLEC, BootStrep. Odbył kilkuletni staż naukowy w grupie Ekstrakcji Informacji Europejskiego Instytutu Bioinformatyki w Cambridge, był konsultantem Institut für Deutsche Sprache w Mannheim w zakresie wielkoskalowych narzędzi korpusowych. Konsultant naukowy sektora prywatnego (np. Pfizer UK, Transition Technologies, Advertine) w zakresie systemów ekstrakcji informacji z danych językowych. Autor wyszukiwarki frazeo.pl, monitorcorpus.com, słowników frazeologicznych HASK EN i HASK PL, Spokes. Koordynator projektu CLARIN-PL w Uniwersytecie Łódzkim. Główny autor i opiekun specjalnego urządzenia badawczego – Wyszukiwarki PELCRA dla NKJP. Prowadzi liczne warsztaty i wykłady z zakresu badań nad frazeologią i korpusów mowy w Polsce i zagranicą.

**dr hab. inż. Maciej Piasecki, prof. PWr** – wieloletni pracownik naukowy i dydaktyczny Wydziału Informatyki i Zarządzania Politechniki Wrocławskiej, zastępca kierownika Katedry Inteligencji Obliczeniowej; promotor kilkudziesięciu prac magisterskich i licencjackich; Polski Koordynator Narodowy CLARIN z ramienia MNiSW, kierownik Grupy Technologii Językowych G4.19, specjalista w zakresie maszynowego przetwarzania języka.

**dr Ewa Rudnicka** – pracowniczka Katedry Inteligencji Obliczeniowej PWr, członkini zespołu CLARIN-PL oraz Grupy Technologii Językowych G4.19, językoznawczyni, koordynatorka prac nad rzutowaniem Słownosieci na angielski Princeton WordNet, autorka i współautorka wielu publikacji dotyczących relacyjnych baz danych leksykalnych. Prowadziła i współorganizowała szereg warsztatów, prelekcji, seminariów dotyczących maszynowego przetwarzania języka.

**dr inż. Tomasz Walkowiak** – wieloletni pracownik Wydziału Elektroniki PWr, Zakładu Systemów Komputerowych i Dyskretnych, członek Grupy Technologii Językowych G4.19, jeden z głównych wykonawców i projektantów systemów tworzonych na potrzeby Infrastruktury Badawczej CLARIN-PL, promotor kilkudziesięciu prac licencjackich i magisterskich, wieloletni wykładowca i dydaktyk.

**dr hab. Monika Zaśko-Zielińska, prof. UWr** – językoznawca, zatrudniona w Zakładzie Współczesnego Języka Polskiego Instytutu Filologii Polskiej UWr. Od wielu lat współpracuje z zespołem CLARIN-PL. Członkini Language and Law Association (ILLA). Interesuje się m.in. stylistyką języka polskiego, analizą dyskursu, lingwistyką kryminalistyczną i językoznawstwem korpusowym. Jest autorką prac na temat świadomości językowej, genologii lingwistycznej, ideologii w dyskursie oraz poszukiwania śladu językowego w ramach kryminalistycznej analizy dokumentu. Promotorka licznych prac magisterskich, prowadziła warsztaty na temat analiz korpusowych przeznaczone dla teoretyków i praktyków.

**dr hab. Magdalena Zawisławska** – adiunkt, od 1995 roku zatrudniona w Instytucie Języka Polskiego UW. Wykłada tam gramatykę opisową współczesnego języka polskiego, porównawczą semantykę leksykalną oraz metodologię nauczania w szkole wyższej dla doktorantów. Prowadzi także kursy z poprawności językowej i prostej polszczyzny dla redaktorów, pracowników banków i ministerstw. Wykładała na Uniwersytecie Dzieci (gdzie zdobyła wyróżnienie za wykład w kategorii Najlepszy Wykładowca w roku 2018/2019) oraz na Uniwersytecie Trzeciego Wieku. Jest autorką wielu publikacji w języku polskim i angielskim dotyczących takich zagadnień, jak metafora, językoznawstwo kognitywne, semantyka leksykalna, semantyka kontrastywna, językoznawstwo korpusowe. Była wykonawcą w wielu projektach naukowych, w tym współtworzyła polski WordNet (Słowociec), polski FrameNet (Ramki) oraz Polski Korpus Koreferencyjny. Kierowała projektem „Synamet – mikrokorpus metafor synestezyjnych. Formalizacja opisu i wypracowanie efektywnych metod analizy metafor w dyskursie” (OPUS NCN).



## 19. SPOSÓB PRZEPROWADZENIA EWALUACJI ZAJĘĆ I CAŁEGO PROGRAMU SZKOŁY LETNIEJ.

Ewaluacja zostanie przeprowadzona za pomocą anonimowej ankiety internetowej, dostępnej po zakończeniu zajęć. Ocenie podlegać będą: program szkoły, poszczególne bloki zajęciowe, a także strona organizacyjna projektu. Kryteria oceny obejmą: zawartość merytoryczną, słabe i mocne strony zajęć, użyteczność zdobytej wiedzy oraz ogólną satysfakcję z udziału w szkole i jej wpływ na rozwój badawczy uczestników.

20. W ZAŁĄCZNIKU – Wykaz materiałów przekazywanych studentom podczas zajęć.