



Elektroniczny korpus polskiej łaciny średniowiecznej

Warsztaty

*Reprezentatywność:
funkcje i gatunki piśmiennictwa*

Instytut Języka Polskiego PAN
Kraków, 12–13 grudnia 2019 r.



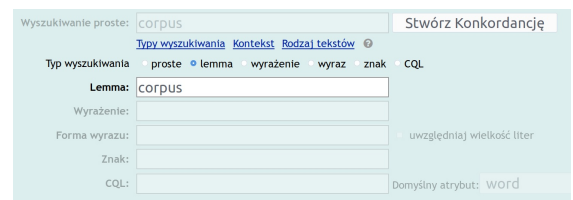
NARODOWY PROGRAM
ROZWOJU HUMANISTYKI

PAN **IJP** Instytut Polskiej
Języka Akademii
Polskiego Nauk

O projekcie

Projekt Elektronicznego korpusu polskiej łaciny średniowiecznej realizowany jest w latach 2018-2023 i finansowany ze środków Narodowego Programu Rozwoju Humanistyki. Oprócz rozbudowy pilotażowej wersji korpusu powstałej w latach 2012-2017 zakłada powstanie rozprawy doktorskiej, popularyzację technik korpusowych w badaniach nad tekstami średniowiecznymi oraz organizację cyklicznych warsztatów poświęconych zróżnicowaniu polskiej łaciny średniowiecznej.

Już w obecnej wersji korpusu wyszukiwać można pojedyncze słowa czy wyrażenia, jak i formułować zapytania w składni CQL.



Interfejs NoSketchEngine pozwala także wyświetlać konkordancję, generować listy wyrazów i kolokacji, tworzyć podkorpusy czy sortować wyniki.

et hora eduxit de carcere **corporis** , cum nouissima tuba audientes ex fama , quod **corpus** sue Ptis iam fuisset in iristi , recepturi , prout in **corpore** gessimus , siue bonum sed sani mente pariter et **corpore** existentes , et animi , n in Prossouicz in crastino **Corporis** Christi (d. 2 Junii 138 portatur supra dominicum **Corpus** temporibus Corporis C ieri . His eciam diebus rex **corpus** magistri inter cadaver

Zespół

Kierownik projektu

Krzysztof Nowak

Opieka nad doktorantem

Michał Rzepiela

Doktorant

Łukasz Halida

Anotacja tekstów

Michał Kulpa / Hanna Lesner / Piotr Małek /
Monika Mansfeld / Katarzyna Wojtalik / Łukasz
Tomanek

Wsparcie IT

Bartłomiej Borek

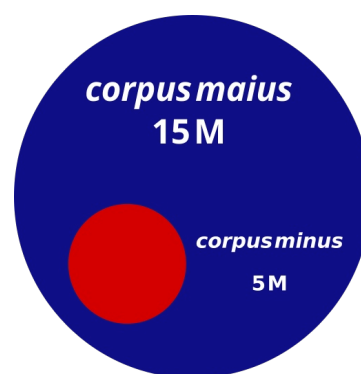
Administracja i księgowość

Barbara Fular / Mariola Kalinowska /
Szymon Seweryn

<https://scriptores.pl/efontes>

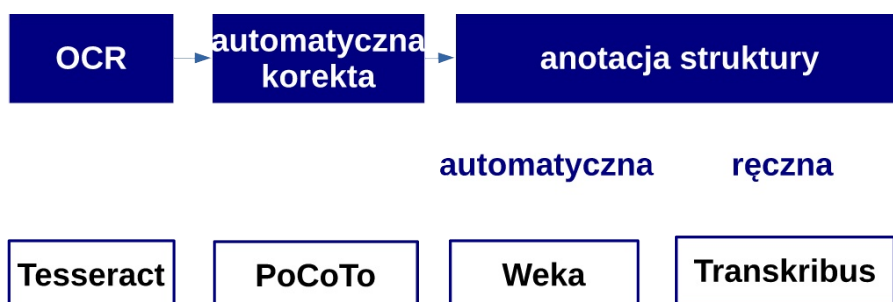
O projekcie

Opracowywana w ramach projektu nowa wersja korpusu przyniesie znaczące zmiany jego architektury. W obrębie liczącego ok. 15 milionów segmentów corpus maius wyodrębnimy ręcznie anotowane (lemmatyzacja i anotacja morfosyntaktyczna) i zrównoważone corpus minus.

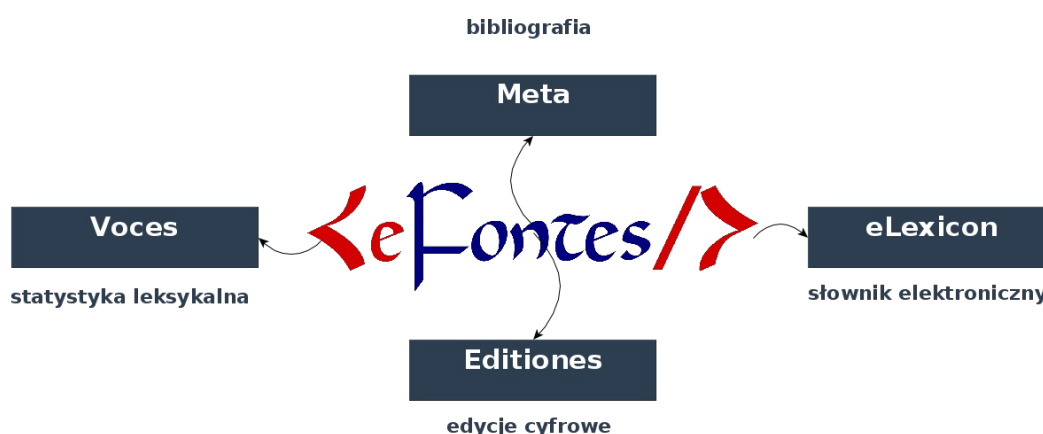


Znacząco zmienił się w nowej odsłonie projektu sam proces tworzenia korpusu. Tekst rozpoznajemy za pomocą wolnego i darmowego programu Tesseract, którego wyniki pracy poprawiamy w programie PoCoTo. Wstępne rozpoznanie struktury tekstu umożliwia zastosowanie algorytmów uczenia maszynowego,

natomiast za korektę tekstu odpowiada zespół anotatorów-filologów.



Oprócz interfejsu wyszukiwania teksty korpusu udostępnione będą na portalu edycji cyfrowych Editiones. Narzędzie Voces zbudowane w oparciu o język R pozwoli na precyzyjne badania statystyczne tekstów polskiego średniowiecza. W ramach projektu powstaje również bibliografia polskiego średniowiecza łacińskiego.



Warsztaty projektu eFontes

Organizowane co roku spotkania stanowią istotną część projektu. Mają one pogłębić naszą wiedzę o zróżnicowaniu i bogactwie łacińskiego piśmiennictwa w średniowiecznej Polsce, a zarazem poddać pod dyskusję najważniejsze aspekty tworzenia korpusu, który ma ambicję owo zróżnicowanie odzwierciedlać. Warsztaty staną się okazją do zaprezentowania projektu oraz poznania oczekiwań środowiska. Mamy również nadzieję, że przyczynią się do rozpoznania pytań badawczych, w odpowiedzi na które korpus może w przyszłości pomóc.

Elektroniczny korpus służyć ma badaczom reprezentującym różne dziedziny nauk, stąd i interdyscyplinarny charakter warsztatów, do udziału w nich zapraszamy bowiem wydawców, filologów, językoznawców, historyków, historyków sztuki czy filozofii, a także przedstawicieli humanistyki cyfrowej. Ze względu na charakter projektu szczególnie zależy nam na szerokich ujęciach tematu, w których oprócz przeglądu stanu badań i omówienia kwestii spornych, znajdą się sugestie dotyczące struktury korpusu lub propozycje jego wykorzystania.

Nasz projekt ma na celu również wsparcie badaczy na wczesnym etapie kariery. Mogą oni przedstawić wyniki swoich prac w gronie uznanych specjalistów oraz uczestniczyć (od 2020 r.) w cyklicznych szkoleniach poświęconych wykorzystaniu korpusu w badaniach. Uczestnikom naszych warsztatów wyłonionym w drodze konkursu pokrywamy koszty uczestnictwa w warsztatach.

I Warsztaty eFontes: Reprezentatywność

Na gruncie językoznawstwa reprezentatywnym nazywa się taki korpus, który ma ambicję odzwierciedlania zróżnicowania funkcji i gatunków odmiany języka czy też produkcji językowej w danej epoce. Dzięki temu, jak reprezentatywna próbka w badaniu sondażowym daje wgląd w nastroje czy poglądy populacji, tak reprezentatywny korpus powinien pozwolić na empiryczne badania nad językiem i kulturą średniowiecza.

Facile dictu... W korpusie historycznym, a takim jest Elektroniczny korpus polskiej łaciny średniowiecznej, osiągnięcie reprezentatywności okazuje się być nie lada wyzwaniem. Problemów nastręcza już choćby stan zachowania i wydania zabytków polskiego średniowiecza: wiele dzieł nie zachowało się do naszych czasów, a te zachowane – pozostają często wciąż niewydane. Osobną kwestią jest jakość edycji, na którą niebagatelny wpływ mają kompetencje językowe wydawców czy ich normalizacyjne wybory. Nawet jednak wydania dobre i znakomite bywają objęte prawami autorskimi, co uniemożliwia ich włączenie do korpusu.

Najpoważniejszym wyzwaniem dla twórców korpusu jest jednak samo bogactwo i zróżnicowanie polskiego piśmiennictwa łacińskiego. Choć ostatnie lata przyniosły fundamentalne prace poświęcone średniowiecznej piśmienności oraz pomocne syntezy stanu badań, to na pełną charakterystykę funkcjonalnego i gatunkowego zróżnicowania łacińskiej produkcji w średniowiecznej Polsce przyjdzie nam jeszcze poczekać. A przecież Elektroniczny korpus polskiej łaciny średniowiecznej ma ambicję owo bogactwo już wkrótce odzwierciedlać.

Program

W I warsztatach projektu eFontes uczestniczą tak doświadczeni badacze i wydawcy tekstów, jak i naukowcy stawiający dopiero pierwsze kroki w filologii, historii i edytorstwie. Ci pierwsi pomagają nam w nakreśleniu szerokiej perspektywy piśmiennictwa łacińskiego polskiego średniowiecza; dla tych drugich nasze warsztaty są również okazją dla zaprezentowania pomysłów oraz poddania pod dyskusję wyników badań.

I dzień obrad otworzą referaty przedstawiające problem reprezentatywności w korpusie łaciny średniowiecznej eFontes oraz w historycznym korpusie polszczyzny KorBa. Kolejne wystąpienia będą poświęcone charakterystyce średniowiecznego piśmiennictwa fachowego, w tym traktatom z zakresu teorii muzyki, astronomii, filozofii oraz mnemotechniki. Następnie pochylimy się nad stanem badań nad łacińskim kazaniem w Polsce, a młodzi badacze przedstawią swoje badania poświęcone hagiografii oraz dziełom liturgicznym. Obrady zakończy podsumowanie referatów wygłoszonych w I dniu.

II dzień warsztatów rozpocznie sesja młodych badaczy poświęcona historiografii - dziełu Gala Anonima oraz projektowi "Długosz 2.0". Doświadczeni badacze przedstawią następnie problemy edycji źródeł epigraficznych, zróżnicowanie piśmienności pragmatycznej uniwersytetu oraz miasta, a także stan wydań tekstów prawnych. Obrady zakończymy spojrzeniem ... w przyszłość, ku kolejnej edycji warsztatów, za sprawą referatów poświęconych obecności polszczyzny w średniowiecznym tekście łacińskim oraz w korpusie polszczyzny XVII i XVIII w.

I dzień, 12 grudnia

13:00-13:15

Powitanie Uczestników

dr hab. Maciej Eder, dyrektor IJP PAN

Reprezentatywność w korpusie

13:15-13:45

Prezentacja projektu. Reprezentatywność w korpusie historycznym

dr Krzysztof Nowak

13:45-14:15

Reprezentatywność i zrównoważenie – jak zrealizować te założenia podczas tworzenia korpusu historycznego? (na podstawie doświadczeń z pracy nad Korpusem Barokowym)

dr Dorota Adamiec, prof. dr hab. Włodzimierz Gruszczyński

14:15-14:45

Dyskusja

Przerwa

Piśmiennictwo fachowe: muzyka i astronomia

14:45-15:15

Teksty teoretyczno-muzyczne w zbiorach polskich. Stan prac dokumentacyjnych i edytorskich

prof. dr hab. Elżbieta Witkowska-Zaremba

15:15-15:45

Coelum est corpus simplex... O przedkopernikańskich pismach astronomicznych polskich autorów

dr hab. Agnieszka Maciąg-Fiedler

15:45-16:15

Dyskusja

Przerwa

Piśmiennictwo fachowe: filozofia i mnemotechnika

16:15-16:45

Teksty filozoficzne w Korpusie łaciny średniowiecznej na ziemiach polskich

dr Monika Mansfeld

16:45-17:15

Traktaty mnemotechniczne z XV i XVI w. powstałe w Polsce. Stan badań i kilka słów o technicznym słownictwie mnemotechnicznym

dr hab. Rafał Wójcik

17:15-17:45

Dyskusja

Przerwa

Sermonistyka, hagiografia, liturgia

17:45-18:15

Polska mediewistyczna sermonistyka. Źródła, badania, edycje, perspektywy.

prof. dr hab. Krzysztof Bracha

18:15-18:35

Polskie średniowieczne teksty hagiograficzne o św. Wojciechu. Dlaczego (prawie) nic o nich nie wiemy?

mgr Michał Kulpa

18:35-18:55

Liturgia w polskich źródłach średniowiecznych: przegląd gatunkowo-treściowy

mgr Łukasz Halida

18:55-19:15

Dyskusja

19:15-19:30

Podsumowanie I dnia obrad

dr hab. Michał Rzepiela, prof. IJP PAN

II dzień, 13 grudnia

Historiografia

9:00-9:20

Niejednoznaczność typologii źródeł średniowiecznych na przykładzie „Kroniki” Galla Anonima

mgr Norbert Gołdys

9:20-9:40

Projekt "Długosz 2.0"

dr hab. Michał Rzepiela, mgr Zdzisław Koczarski, mgr Iwona Krawczyk

9:40-10:10

Dyskusja

Przerwa

Epigrafika. Prawo

10:10-10:40

Zagadnienie korpusu epigrafiki średniowiecznej i wczesnonowożytnej w Polsce. Stan badań i edycji, uwarunkowania teoretyczne i potrzeba ich rewizji

dr hab. Marek Janicki

10:40-11:10

Korpus źródeł średniowiecznego prawa polskiego – czy istnieje uniwersalny klucz edycji elektronicznej?

dr hab. Maciej Mikuła

11:10-11:25

Dyskusja

11:25-12:45

Przerwa

Piśmiennictwo pragmatyczne: uniwersytet i miasto

12:45-13:15

Piśmiennictwo pragmatyczne w kręgu Uniwersytetu Krakowskiego (XV-XVIII w.). Teksty, stan badań, edycje
dr hab. Maciej Zdanek

13:15-13:45

Piśmiennictwo pragmatyczne w mieście na przykładzie Krakowa
dr hab. Marcin Starzyński, prof. UJ

13:45-14:15

Dyskusja
Przerwa

Granice: łacina i polszczyzna

14:15-14:45

Na styku staropolszczyzny i łaciny – badania interferencji językowych z wykorzystaniem Elektronicznego korpusu łaciny średniowiecznej?
dr Anna Ledzińska

14:45-15:15

Łacina w KorBie – co ciekawego znajdzie latynista w korpusie polskich tekstów barokowych?
mgr Katarzyna Kryńska, dr Renata Bronikowska

15:15-15:30

Dyskusja
Przerwa

15:30-16:00

Podsumowanie warsztatów
Krzysztof Nowak

Streszczenia wystąpień

Streszczenia wystąpień można znaleźć na stronie internetowej warsztatów: <https://scriptores.pl/efontes>

Sylwetki Prelegentów

Sylwetki naszych Prelegentów można znaleźć na stronie internetowej warsztatów: <https://scriptores.pl/efontes>

dr Dorota Adamiec



Instytut Języka Polskiego PAN
Pracownia Historii Języka Polskiego XVII i XVIII wieku



Warszawa



dorota.adamiec@interia.pl

prof. dr hab. Krzysztof Bracha



Uniwersytet Jana Kochanowskiego w Kielcach
Instytut Historii



Kielce



kbracha@ujk.edu.pl

dr Renata Bronikowska



Instytut Języka Polskiego PAN
Pracownia Historii Języka Polskiego XVII i XVIII wieku



Warszawa



r.bronikowska@wp.pl

dr hab. Maciej Eder



Instytut Języka Polskiego PAN
Dyrektor



Warszawa



maciej.eder@ijp.pan.pl

mgr Norbert Gołdys



Instytut Języka Polskiego PAN
Pracownia Historii Języka Polskiego XVII i XVIII wieku



Warszawa



norbert.goldys@gmail.com

prof. dr hab. Włodzimierz Gruszczyński



Instytut Języka Polskiego PAN
Pracownia Historii Języka Polskiego XVII i XVIII wieku



Warszawa



wlodekiewa@poczta.onet.pl

mgr Łukasz Halida



Instytut Języka Polskiego PAN
Pracownia Łaciny Średniowiecznej



Kraków



lukasz.halida@ijp.pan.pl

dr hab. Marek A. Janicki



Uniwersytet Warszawski
Instytut Historyczny



Warszawa



m.a.janicki@uw.edu.pl

mgr Zdzisław Koczarski



Instytut Języka Polskiego PAN
Pracownia Łaciny Średniowiecznej



Kraków



zdzislaw.koczarski@ijp.pan.pl

mgr Iwona Krawczyk



Instytut Języka Polskiego PAN
Pracownia Łaciny Średniowiecznej



Kraków



iwona.krawczyk@ijp.pan.pl

mgr Katarzyna Kryńska



Instytut Języka Polskiego PAN
Pracownia Historii Języka Polskiego XVII i XVIII wieku



Warszawa



katarzyna.krynska@ijp.pan.pl

mgr Michał Kulpa



Uniwersytet Jagielloński
Instytut Historii



Kraków



michal.kulpa@student.uj.edu.pl

dr Anna Ledzińska



Instytut Języka Polskiego PAN
Pracownia Łaciny Średniowiecznej



Kraków



anna.ledzinska@ijp.pan.pl

dr hab. Agnieszka Maciąg-Fiedler



Instytut Języka Polskiego PAN
Pracownia Łaciny Średniowiecznej



Kraków



agnieszka.maciag@ijp.pan.pl

dr Monika Mansfeld



Uniwersytet Łódzki
Katedra Etyki



Łódź



morginess@gmail.com

dr hab. Maciej Mikula



Wydział Prawa i Administracji Uniwersytetu Jagiellońskiego
Zakład Prawa Kościelnego i Wyznaniowego



Kraków



maciej.mikula@uj.edu.pl

dr Krzysztof Nowak



**Instytut Języka Polskiego PAN
Pracownia Łaciny Średniowiecznej**



Kraków



krzysztof.nowak@ijp.pan.pl

dr hab. prof. IJP PAN Michał Rzepiela



**Instytut Języka Polskiego PAN
Pracownia Łaciny Średniowiecznej**



Kraków



michal.rzepiela@ijp.pan.pl

dr hab. prof. UJ Marcin Starzyński



**Uniwersytet Jagielloński
Instytut Historii**



Kraków



starzynski@ymail.com

prof. dr hab. Elżbieta Witkowska-Zaremba



Instytut Sztuki PAN



Warszawa



elzbieta.witkowska-zaremba@ispan.pl

dr hab. Rafał Wójcik



**Uniwersytet im. Adama Mickiewicza
Biblioteka Uniwersytecka**



Poznań



rafal@amu.edu.pl

dr hab. Maciej Zdanek



Archiwum Uniwersytetu Jagiellońskiego



Kraków

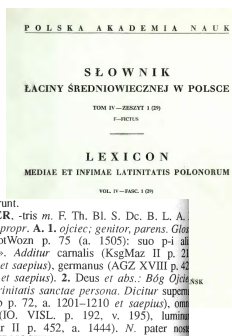


maciej.zdanek@uj.edu.pl

Pracownia Łaciny Średniowiecznej IJP PAN

Zespół Pracowni liczy obecnie 8 osób: Michał Rzepiela (kierownik) / Zdzisław Koczarski / Iwona Krawczyk / Aleksandra Kulbicka / Anna Ledzińska / Agnieszka Maciąg-Fiedler / Jagoda Marszałek / Krzysztof Nowak

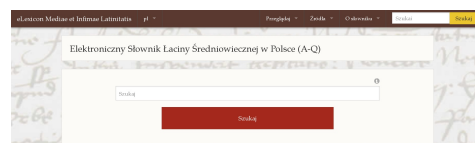
Słownik łaciny średniowiecznej w Polsce



Słownik łaciny średniowiecznej w Polsce powstaje w Pracowni od 1953 r. Jego pierwszym redaktorem był prof. Marian Plezia, którego obowiązki przejęła następnie prof. Krystyna Weyssenhoff-Brożkova. Obecnie pracom nad Słownikiem przewodzi dr hab. Michał Rzepiela.

Elektroniczny słownik łaciny średniowiecznej

W 2015 roku w Pracowni powstała elektroniczna wersja Słownika (hasła A-Q).



<https://scriptores.pl/elexicon>

Hortus
Deliciarum

Iwona Krawczyk i Anna Ledzińska z Pracowni organizują cykliczne spotkania mediewistyczne pod hasłem Hortus deliciarum. Wokół średniowiecznego tekstu. Spotkania transmitowane są na kanale YouTube cyklu.

<https://scriptores.pl/hortus>

Długosz 2.0

Od 2018 r. zespół pod kierunkiem M. Rzepli realizuje projekt Długosz 2.0. Korpus i narzędzia badań nad językiem Jana Długosza. Prace mają na celu m. in. drobiazgową anotację dzieł historyka. W ramach projektu mgr Z. Koczarski opracowuje rozprawę poświęconą językowi Długosza.

<https://scriptores.pl/dlugosz>

Instytut Języka Polskiego PAN

Instytut Języka Polskiego Polskiej Akademii Nauk

jest czołowym ośrodkiem badań językoznawczych



w Polsce. W Instytucie oprócz prac naukowych

powstają również monumentalne słowniki, bibliografie, bazy danych i korpusy. IJP PAN jest także jednym z najważniejszych ośrodków badań stylometrycznych na świecie.

Słowniki

Słowniki historyczne

- Słownik łaciny średniowiecznej w Polsce
- Słownik staropolski
- Słownik pojęciowy języka staropolskiego
- Elektroniczny Słownik Języka Polskiego XVII i XVIII wieku

Słowniki współczesnej polszczyzny

- Wielki Słownik Języka Polskiego
- Słownik gwar polskich

Słowniki onomastyczne

- Antroponimia Polski od XVI w.
- Słownik staropolskich nazw osobowych
- Słownik etymologiczno-motywacyjny staropolskich nazw osobowych

Dyrekcja IJP PAN

Dyrektor Instytutu

dr hab. Maciej Eder

Zastępca dyrektora ds. naukowych

dr hab. Michał Rzepiela

Zastępca dyrektora ds. administracyjnych

mgr Szymon Seweryn

Szkoła doktorska Anthropos

Instytut Języka Polskiego PAN prowadzi studia doktoranckie w ramach Szkoły Doktorskiej Anthropos Instytutów Polskiej Akademii Nauk. Szkoła zrzesza instytuty humanistyczne prowadzące badania z zakresu m.in. językoznawstwa, literaturoznawstwa, historii, archeologii i etnologii.

<https://ijp.pan.pl>

Notatki

Notatki

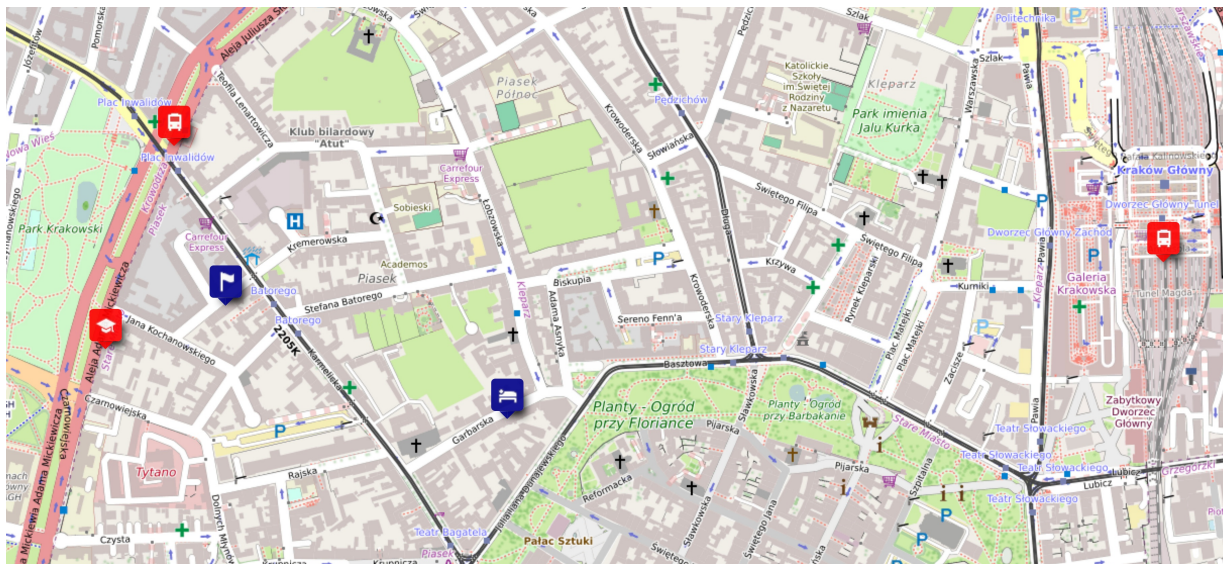
Jak do nas trafić

Instytut Języka Polskiego

Polskiej Akademii Nauk

al. Adama Mickiewicza 31

31-120 Kraków



Z Dworca PKP Kraków Główny oraz z Dworca PKS najłatwiej dojechać do IJP PAN tramwajem z przystanku Teatr Słowackiego do przystanku Plac Inwalidów (w kierunku Bronowic, Bronowic Małych). Niektóre autobusy PKS zatrzymują się na przystanku Plac Inwalidów.

Kontakt ws. warsztatów

krzysztof.nowak@ijp.pan.pl

II Warsztaty projektu eFontes (2020): Granice

Proces tworzenia korpusu wiąże się z podejmowaniem decyzji, które arbitralnie wkraczają w rzeczywistość historyczną i kulturową. Kiedy zaczyna się, a kiedy kończy polskie średniowiecze? Jakie granice geograficzne wytyczyć dla polskiej produkcji piśmiennej w XIII, a jakie w XV w.? Czy dzieła Mateusza z Krakowa wiązać ze środowiskiem Krakowa, Pragi czy może Heidelbergu? Czy szeroko recypowane teksty europejskiego średniowiecza winny znaleźć miejsce w korpusie? Jaki jest status tekstów polsko-łacińskich?

Elektroniczny korpus może także dostarczyć odpowiedzi na powyższe pytanie i przyczynić się do wyłonienia nowych hipotez badawczych. Badania nad zmiennością języka i metody humanistyki cyfrowej (ze stylometrią na czele) pozwalają nam dziś bowiem na nowo zdefiniować ważne pojęcia analityczne, a także zobiektywizować opis różnicowania geograficznego, chronologicznego, stylistycznego i społecznego łacińskich tekstów polskiego średniowiecza.

Podczas przyszłorocznych warsztatów poświęconych *Granicom, cezuirom, epokom* pochylimy się m. in. nad następującymi zagadnieniami:

- cezury, epoki, okresy w świetle analizy łacińskich tekstów polskiego średniowiecza
- granice geograficzne a różnicowanie polskiej łaciny średniowiecznej
- społeczna wariacja tekstów
- tekstowe przejawy polsko-łacińskiej dyglosji, języki wernakularne w tekście łacińskim
- polska łacina średniowieczna a europejska łacina średniowieczna

Daty

luty 2020 r.

call for papers

kwiecień 2020 r.

skrócony abstrakt wystąpienia

maj 2020 r.

informacja o przyjęciu referatu

listopad 2020 r.

nadsyłanie artykułu do tomu pokonferencyjnego

grudzień 2020 r.

warsztaty

czerwiec 2021 r.

publikacja tomu pokonferencyjnego