



Fundusze Europejskie
dla Mazowsza



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Mazowsze.
serce Polski

Tytuł projektu:

Przeprowadzenie prac B+R nad stworzeniem modelu tłumaczeniowego do jakościowego i zgodnego z oczekiwaniami odbiorców tłumaczenia tekstów patentowych z języka angielskiego oraz innych języków obcych na język polski, opartego na sieciach neuronowych

Realizator:

Patpol Kancelaria Patentowa sp. z o.o.

Projekt współfinansowany przez Fundusze Europejskie dla Mazowsza 2021-2027, w ramach działania 1.1 Badania, Rozwój i Innowacje Przedsiębiorstw

Całkowity koszt projektu: 3 450 953,10 PLN

Kwota dofinansowania: 1 648 186,54 PLN



patpol

TYTUŁ PROJEKTU:

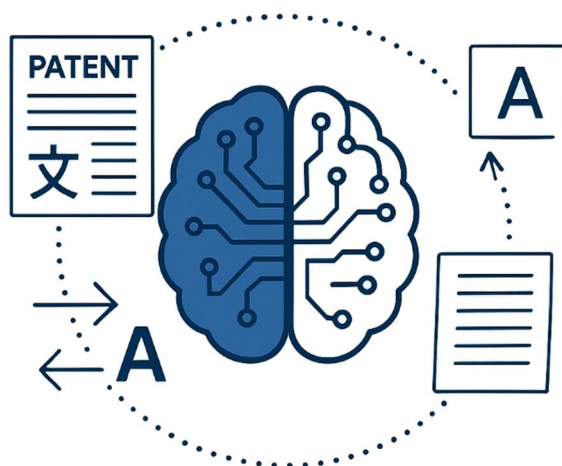
Przeprowadzenie prac B+R nad stworzeniem modelu tłumaczeniowego do jakościowego i zgodnego z oczekiwaniami odbiorców tłumaczenia tekstów patentowych z języka angielskiego oraz innych języków obcych na język polski, opartego na sieciach neuronowych.

REALIZATOR:

Patpol Kancelaria Patentowa Sp. z o.o.

CEL PROJEKTU:

Celem niniejszego projektu jest opracowanie narzędzia tłumaczeniowego dedykowanego tekstom patentowym, które zapewni wysoką jakość dostarczanych tłumaczeń opisów patentowych oraz zagwarantuje automatyzację tego procesu przy zachowaniu niezwykle pożądanej w tej dziedzinie standaryzacji terminologii technicznej. W projekcie będziemy się koncentrować na parze języków angielski – polski. Jednakże mechanizmy wypracowane przy realizacji projektu będą mogły być z powodzeniem wykorzystane do innych par językowych.



Projekt ten stanowi odpowiedź na istniejącą na rynku krajowym i wciąż wzrastającą potrzebę szybkiego dostępu do precyzyjnego, zgodnego z wymaganiami prawnymi tłumaczenia dokumentacji patentowej. W dobie ekspansji działalności polskich podmiotów na rynku światowym, w tym w szczególności coraz powszechniejszej migracji działalności krajowych przedsiębiorców na rynek unijny, dostęp do dobrej jakości tłumaczeń – czy to zgłoszeń patentowych mogących skutkować objęciem ujawnionych w nich rozwiązań patentem, czy też samych praw wyłącznych obowiązujących na tych terytoriach i ograniczających swobodę gospodarczą w danym zakresie – jest niezbędny dla zminimalizowania ryzyka naruszenia przez polskie przedsiębiorstwa praw osób trzecich, co stanowi bezpośrednie zagrożenie dla ich przyszłego funkcjonowania.

Zgodnie z założeniami projektu, dzięki zastosowaniu sztucznej inteligencji (sieci neuronowe), opracowywane narzędzie umożliwi tłumaczenie tekstów patentowych na poziomie porównywalnym z tłumaczeniem wykonanym przez człowieka, przy jednoczesnej redukcji kosztów i nakładu czasu niezbędnego na pozyskanie takiego tłumaczenia.

Rezultaty prac B+R przeprowadzonych w ramach projektu zostaną wdrożone w kancelarii Patpol w postaci innowacji procesowej. Na potrzeby realizacji projektu utworzone zostaną również dodatkowe miejsca pracy.

FINANSOWANIE:

Wartość projektu: 3 450 953,10 PLN

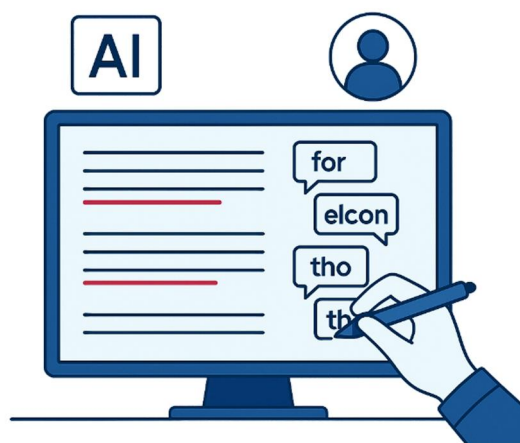
Wysokość wkładu z Funduszy Europejskich: 1 648 186,54 PLN

OKRES REALIZACJI:

01.08.2025 – 31.01.2027 (projekt realizowany zgodnie z harmonogramem).

PROBLEMY Z TŁUMACZENIAMI MASZYNOWYMI DOKUMENTÓW PATENTOWYCH

Tłumaczenia patentowe wymagają wyjątkowej precyzji i spójności terminologicznej, ponieważ każde słowo może mieć znaczenie prawne. Obecne narzędzia tłumaczenia maszynowego, takie jak Google Translate czy DeepL nie spełniają tych wymagań – generują niespójne frazy, ignorują kontekst techniczny i prawny, a także często „halucynują”, generując nieistniejące wyrażenia specjalistyczne, w szkodliwy sposób parafrazują, czy nawet pomijają fragmenty tekstów.



Zespół B+R kancelarii Patpol zidentyfikował kilka kluczowych wyzwań:

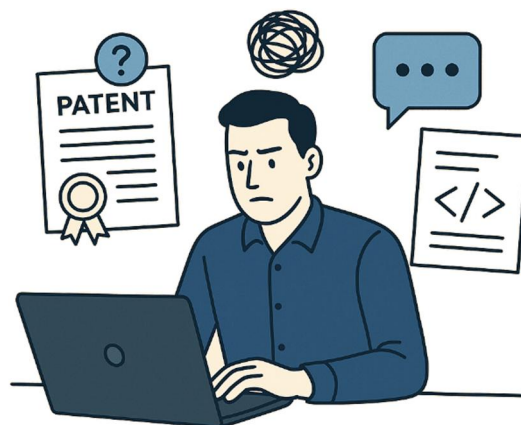
- **Niespójność danych:** dokumenty, które służą do trenowania modeli generycznych były tłumaczone przez różnych tłumaczy, a często nawet z wykorzystaniem narzędzi maszynowych; powoduje to, że tłumaczenia są niespójne, co utrudnia ich wykorzystanie do trenowania modeli;

- Brak kontekstu: tłumaczenie fraz technicznych w dokumentacji patentowej wymaga uwzględnienia dziedziny technologii, zidentyfikowanej, na przykład za pomocą klasyfikacji patentowej, jak również konkretnego kontekstu zdania; obecne modele do tłumaczeń maszynowych ignorują te aspekty co skutkuje generowaniem nieprawidłowych tłumaczeń;
- Trudności z automatyzacją: wyszukiwanie sensownych fraz i ich dopasowanie wymaga zaawansowanej analizy językowej, której nie zapewniają standardowe algorytmy (np. stosowane w programach do wspomaganie procesu tłumaczenia – CAT);
- Potrzeba narzędzi wspierających post-edycję: pożądane byłoby stworzenie systemów, które automatycznie wykrywałyby błędy i sugerowały korekty, wspierając i upraszczając pracę specjalistów z konkretnej dziedziny technologii.

W odpowiedzi na te problemy, zespół kancelarii Patpol rozpoczął – w ramach realizacji projektu - budowę własnej bazy terminologicznej oraz narzędzi NLP, które będą lepiej dopasowane do specyfiki dokumentów patentowych i realnych potrzeb tłumaczy.

DLACZEGO TO WAŻNE?

W patentach precyzja językowa jest kluczowa – błędne tłumaczenie może prowadzić do utraty ochrony prawnej. Na przykład, zawężające tłumaczenie patentu europejskiego może skutkować znaczącym ograniczeniem ochrony patentowej. Niewłaściwa interpretacja zastrzeżeń patentowych wynikająca z nieprawidłowego tłumaczenia może również prowadzić do poważnych konsekwencji biznesowych i niezamierzonego naruszenia prawa osób trzecich.



Każde słowo dokumentu patentowego, w szczególności w zastrzeżeniach patentowych, staje się istotne w trakcie sporu patentowego, czy to w sprawie o naruszenie patentu czy jego unieważnienie.

Projekt tłumaczeniowy prowadzony przez kancelarię Patpol odpowiada na te wyzwania poprzez budowę dedykowanego modelu tłumaczeniowego, trenowanego na specjalistycznych zbiorach danych, z udziałem specjalistów dysponujących wiedzą technologiczną oraz patentową.

ZAKRES PRAC B+R:

- Opracowanie algorytmu do utworzenia bazy jednostek leksykalnych (frazy, terminy techniczne).
- Przygotowanie zbioru segmentów tłumaczeniowych na podstawie rzeczywistych dokumentów patentowych.
- Wytrenowanie sieci neuronowej na bazie tego zbioru segmentów tłumaczeniowych, z uwzględnieniem specyfiki różnych dziedzin techniki.
- Testowanie i walidacja jakości tłumaczeń przez ekspertów patentowych.
- Wdrożenie aplikacji webowej umożliwiającej automatyczne tłumaczenie i post-edycję dokumentów patentowych.

EFEKTY DOTYCHCZASOWE:

- Powstanie prototypu modelu tłumaczeniowego
- Automatyzacja procesu tłumaczenia patentów
- Standaryzacja terminologii technicznej
- Usprawnienie pracy rzeczników patentowych i tłumaczy



KORZYŚCI DLA KANCELARII I KLIENTÓW:

- Obniżenie kosztów tłumaczeń
- Skrócenie czasu niezbędnego do przygotowania tłumaczenia wysokiej jakości
- Zwiększenie jakości i spójności tłumaczeń patentowych
- Możliwość automatyzacji post-edycji dokumentów

FINANSOWANIE PROJEKTU:

Projekt współfinansowany przez Fundusze Europejskie dla Mazowsza 2021-2027, w ramach działania 1.1 Badania, Rozwój i Innowacje Przedsiębiorstw.

Całkowity koszt projektu: 3 450 953,10 PLN

Kwota dofinansowania: 1 648 186,54 PLN