

Celem projektu JuDDGES (Judicial Decision Data Gathering and Encoding) jest rozwój i zaproponowanie nowych technologii przetwarzania języka naturalnego i technologii Human-In-The-Loop, aby dać naukowcom zajmującym się prawem, otwarte oprogramowanie i narzędzia, które umożliwiają rozległą, elastyczną i ciągłą możliwość dodawania meta-annotacji (zarówno zautomatyzowanej, jak i dla ekspertów dziedzinowych w pętli sprzężenia zwrotnego). Założeniem projektu jest wytworzenie podejść do annotacji oraz zasoby danych akt prawnych/wyroków wydanych przez sądy karne w różnych jurysdykcjach, o różnych konstytucjach prawnych (Polska, Anglia i Walia). W związku z tym w projekcie będzie do usunięcia wiele barier związanych z niejednorodnością zasobów, języka, danych i formatów, które obecnie utrudniają badania nad podejmowaniem decyzji sądowych. Dzięki udostępnieniu nowego podejścia do annotacji, narzędzi i zasobów danych w publicznym repozytorium, badacze będą mogli rozwijać i empirycznie testować teorie podejmowania decyzji sądowych oraz odpowiadać na pytania związane z polityką sądowniczą oraz praktyką. Ponadto badacze z instytucji publicznych (prawnych) również będą w szerszym zakresie mogli wykorzystywać dane do swoich celów. Zastosowanie podejścia Otwartej Nauki (Open Science) wypełni w trakcie realizacji projektu istotną lukę w dziedzinie prawniczych badań empirycznych z wykorzystaniem nowo opracowanych podejść uczenia aktywnego, słabej superwizji, podejścia Human-In-The-Loop oraz uczenia ciągłego w annotacji tekstów prawnych. Powstały międzynarodowy zestaw danych i zestaw narzędzi z annotacjami będzie największym i najbardziej wszechstronnym i nadającym się do wielokrotnego użytku repozytorium badań prawnych w Europie służącym do eksploracji zagadnień podejmowania decyzji sądowych.

W ramach projektu opracowane zostanie rozwiązanie (narzędzie) oparte na sztucznej inteligencji, które będzie mogło być wykorzystywane przez naukowców do analizy nieustrukturyzowanych danych tekstowych w aktach sądowych i pisemnych orzeczeniach prawnych. W ten sposób zostanie stworzony największy istniejący ogólnokrajowy zbiór danych prawnych w Europie. Narzędzie umożliwi dostęp i stopniowe wzbogacanie dużych, szczegółowych i reprezentatywnych prawnie próbek danych (począwszy od tego, co będzie efektem udostępnienia w ramach projektu) w sposób oszczędny pod względem zasobów, kosztów i czasu. Zbiory danych w ten sposób wytworzone i dostosowane do zasad otwartej nauki oraz publicznie dostępne z zaufanego repozytorium danych będą również przekazane dla innych osób do ponownego wykorzystania - zarówno w ramach jurysdykcji prawnych, jak i między nimi, w tym dla nieznanymi obecnie użytkownikami (np. z instytucji organizujących proces legislacyjny lub sądowy). Otwarte oprogramowanie i narzędzia Human-In-The-Loop (HITL), które będzie efektem projektu umożliwi naukowcom na wczesnym etapie rozwoju, którzy nie są specjalistami w dziedzinie sztucznej inteligencji, przeszukiwanie danych z decyzji sądowych i identyfikowanie nowych linii orzecznictwa. Narzędzie i uzyskane w ten sposób dane poszerzą horyzonty metodologiczne. W dłuższej perspektywie projekt przyczyni się do rozwoju naukowego podejścia do polityki i praktyki sądowej, a sposób jego realizacji aimplementuje zasady zapisane w ramach inicjatyw „Horyzont Europa” dotyczących otwartych danych, stosowania zasad otwartej nauki i wykorzystania otwartych infrastruktur.