

Instrukcja użycia ekstraktora czynności w dokumentach prawnych na platformie CAISE – demonstratora usługi

Projekt **CAISE Cloud Artificial Intelligence Service Engineering** (caise.pl) to międzynarodowy projekt badawczo rozwojowy, w ramach którego powstaje nowoczesna platforma chmurowa umożliwiająca tworzenie i wdrażanie inteligentnych usług dopasowanych do potrzeb kluczowych sektorów publicznych, między innymi **ochrony zdrowia, szkolnictwa i wymiaru sprawiedliwości**.

Usługi na platformie chmurowej CAISE wywoływane są poprzez interfejs **REST API**. W przypadku usługi ekstraktora czynności parametrami wejściowymi są zazwyczaj **surowe dokumenty prawne** w formacie **PDF** lub **DOCX**, a także **częściowo przetworzone dane w formacie JSON**, natomiast wynikiem jest ustrukturyzowana analiza NLP obejmująca m.in. identyfikację czynności, ich stanów, ról podmiotów oraz klastrów podobnych czynności.

Udostępniony demonstrator wykorzystuje interfejs REST (w tle) i stanowi **dodatkowe oprogramowanie demonstracyjne**, prezentujące w praktyce, w jaki sposób instytucja, firma lub organizacja może rozwijać własne rozwiązania z wykorzystaniem usług i platformy CAISE. Usługi platformy CAISE nie są bezpośrednio przeznaczone dla użytkowników końcowych.

Demonstrator bazuje na prostych krokach:

- **wybór dokumentu do analizy** (PDF) lub alternatywnie wczytanie pliku JSON zawierającego wcześniej przygotowane dane,
- **uruchomienie analizy** (w tle realizowane jest żądanie REST do usług CAISE),
- **prezentacja wyników analizy**, obejmująca wizualizację klastrów czynności oraz wyróżnione czynności i ich kluczowe parametry w treści dokumentu,
- **opcjonalne pobranie wyników analizy** w formacie JSON, zawierającym pełną, szczegółową analizę NLP.

Demonstrator umożliwia zapoznanie się z potencjałem usług CAISE w zakresie automatycznej analizy treści prawnych i pokazuje, jak teksty aktów prawnych mogą zostać przekształcone w dane gotowe do dalszego przetwarzania, porównywania i integracji z innymi systemami.

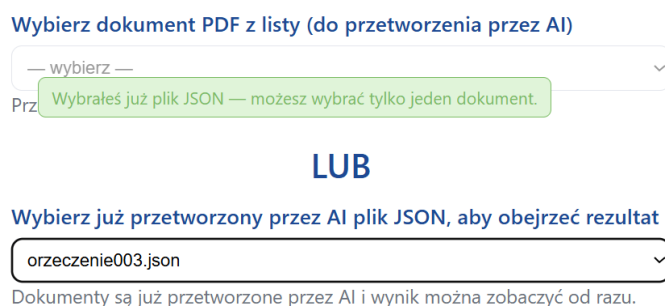
Zobacz działanie demonstratora krok po kroku na kolejnych stronach instrukcji:

1. Wybór dokumentu

Demonstrator umożliwia wybranie dokumentu PDF do ponownego przetworzenia przez algorytmy NLP lub dokumentu JSON, który został wcześniej przetworzony i można obejrzeć gotowe rozwiązanie. Pierwsza opcja zajmuje trochę czasu, nawet kilkadziesiąt sekund, dlatego została dodana możliwość obejrzenia już przeanalizowanego dokumentu JSON.



Demonstrator pozwala wskazać tylko jedną opcję. Chcąc wybrać inną, należy odznaczyć wybrany dokument (ponownie wskazać opcję – wybierz --)



Po dokonaniu wyboru, należy kliknąć przycisk „Aktualizuj”, by uruchomić analizę NLP i przejść do ekranu z wynikiem.

2. Wyniki przetwarzania

Ekran z wynikiem podzielony jest na dwa panele. Lewy prezentuje otrzymane wyniki, zaś prawy wyświetla oryginalny dokument PDF, który był poddany analizie.

Wyniki

Nazwa pliku: ustawa_o_aktach_normatywnych_d20000718lj.json

Liczba zdań: 237

Czas przetwarzania: dokument został już przetworzony wcześniej

Podgląd pliku wynikowego JSON: [Kliknij tutaj](#)

Legenda klastrów:

Kliknij klastr w legendzie – to wskaże jego reprezentację w tekście. Zwróć uwagę, że klastry uwzględniają nie tylko czynność, ale też kontekst (najczęściej Object). Gdy najedziesz myszką na czynność w tekście, zobaczysz jej szczegóły

Noise (-1)
Klastr 0
Klastr 1
Klastr 2
Klastr 3
Klastr 4
Klastr 5
Klastr 6
Klastr 7
Klastr 8
Klastr 9
Klastr 10
Klastr 11
Klastr 12
Klastr 13
Klastr 14
Klastr 15
Klastr 16
Klastr 17
Klastr 18
Klastr 19
Klastr 20
Klastr 21
Klastr 22
Klastr 23
Klastr 24
Klastr 25
Klastr 26
Klastr 27
Klastr 28
Klastr 29
Klastr 30
Klastr 31
Klastr 32
Klastr 33
Klastr 34
Klastr 35

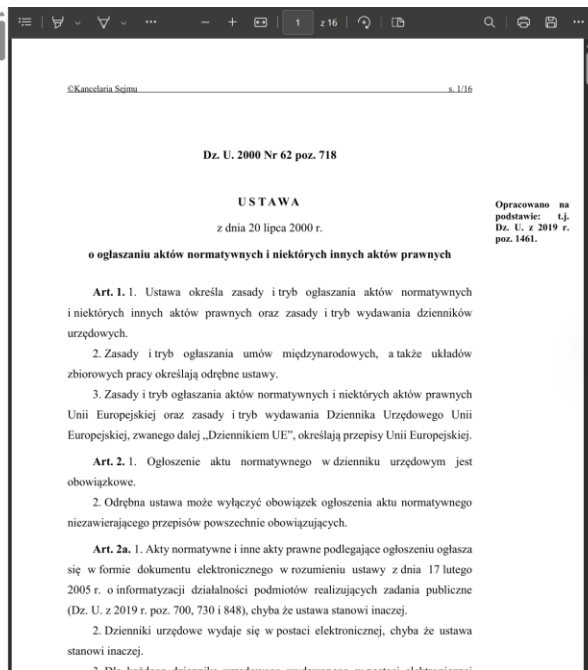
[Wróć](#)

Treść dokumentu:

©

Kancelaria Sejmu s. 1/16 20.08.2019 Dz. U. 2000 Nr 62 poz. 718 U S T A W A z dnia 20 lipca 2000 r. o ogłaszaniu aktów normatywnych i niektórych innych aktów prawnych Art. 1.

1. Ustawa określa zasady i tryb ogłaszania aktów normatywnych i niektórych innych aktów prawnych oraz zasady i tryb wydawania dzienników urzędowych.
2. Zasady i tryb ogłaszania umów międzynarodowych, a także układów zbiorowych pracy określają odrębne ustawy.



Lewy panel w szczególności zawiera:

- podsumowanie analizy NLP z możliwością podglądu pliku JSON z pełną analizą:

 Nazwa pliku: ustawa_o_aktach_normatywnych_d20000718lj.json

 Liczba zdań: 237

 Czas przetwarzania: dokument został już przetworzony wcześniej

 Podgląd pliku wynikowego JSON: [Kliknij tutaj](#)

- interaktywną legendę dla wyodrębnionych klastrów:

Legenda klastrów:

Kliknij klaster w legendzie - to wskaże jego reprezentację w tekście. Zwróć uwagę, że klastry uwzględniają nie tylko czynność, ale też kontekst (najczęściej Object). Gdy najedziesz myszką na czynność w tekście, zobaczysz jej szczegóły



Kliknięcie myszką na wybrany klaster uwidoczni w tekście czynności do niego należące.

- wizualizację treści dokumentu po analizie NLP:

Treść dokumentu:



Kancelaria Sejmu s . 1/16 20.08.2019 Dz .

U . 2000 Nr 62 poz . 718 U S T AWA z dnia 20 lipca 2000 r . o **ogłaszaniu** aktów normatywnych i niektórych innych aktów prawnych Art . 1 .

1. Ustawa **określa** zasady i tryb **ogłaszania** aktów normatywnych i niektórych innych aktów prawnych oraz zasady i tryb **wydawania** dzienników urzędowych.

2. Zasady i tryb c Activity: ['określają'] zynarodowych , a także układów zbiorowych pracy
określają odrębne State: present

3. Zasady i tryb wykonania przepisów normatywnych i niektórych aktów prawnych Unii Europejskiej oraz zasady trybu wydawania Dziennika Urzędowego Unii Europejskiej, zwanego dalej "Dziennikiem UE", określają przepisy Unii Europejskiej.

Activity: ['określają']
State: present
Cluster: 32
Performer: Zasady
Object: przepisy

Wybranie klastra w legendzie (w tym przypadku nr 32) zaznacza wszystkie czynności do niego należące, zaś najechanie myszką na wybraną czynność, wyświetla jej szczegóły.