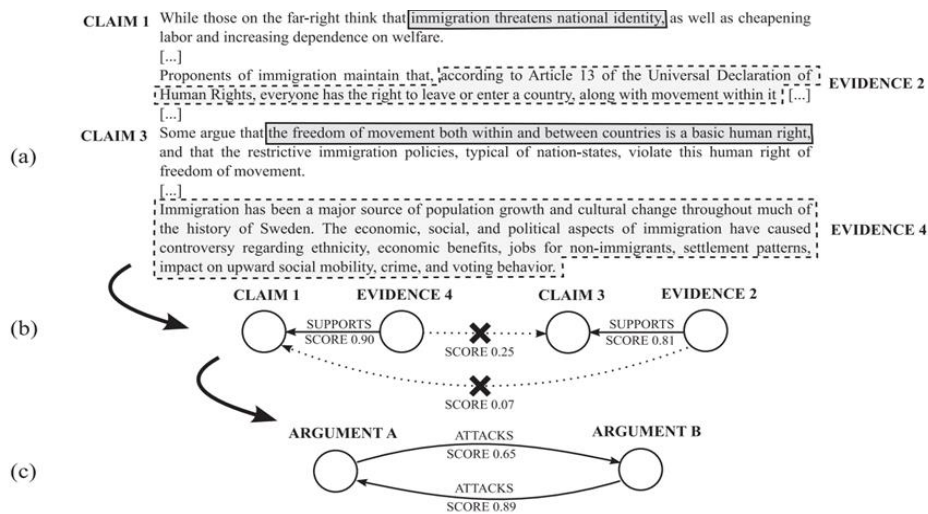


Komunikacja online, zwłaszcza media społecznościowe, to ogromne kopalnie danych, pochodzących z wielu różnych źródeł. Niestety jest to również obszar szczególnie podatny na manipulacje, fake newsy, skrajne emocje i polaryzację. Podczas gdy najczęściej stosowane techniki ekstrakcji danych opierają się na analizie statystycznej i sieciowej, mniej konwencjonalne podejście Argument Mining (AM) daje możliwość kompleksowego badania jakościowego debat publikowanych na platformach i portalach społecznościowych. Takie podejście oferuje bezprecedensowy zestaw narzędzi dla decydentów i badaczy. Jako dziedzina przetwarzania języka naturalnego (NLP) i lingwistyki komputerowej, Argument Mining polega na automatycznej identyfikacji, ekstrakcji i analizie struktur argumentacyjnych w tekstach języka naturalnego, co obejmuje rozpoznawanie podstawowych elementów argumentów, takich jak twierdzenia i dowody.

Pojawia się jednak problem: te nowe techniki są często używane w ograniczony sposób. Badania z wykorzystaniem AM skupiają się głównie na aspektach technicznych lub językowych i pomijają szerszy kontekst społeczny. Mają tendencję do skupiania się na tym, jak przekonujący jest pojedynczy argument, bez rozważania jego miejsca w całej debacie. A przecież debaty publiczne odgrywają ogromną rolę w kształtowaniu społeczeństwa, angażowaniu ludzi w działania społeczne i tworzeniu więzi. Dlatego w tym projekcie chcemy wykorzystać techniki oparte na sztucznej inteligencji do analizy i zrozumienia internetowych debat na tematy publiczne w całej ich złożoności, łącząc metody z nauk społecznych i informatyki. Badanie partycypacji obywatelskiej online w sferze publicznej, debat i argumentacji staje się jednym z najważniejszych wyzwań stojących obecnie przed naukami społecznymi, a zwłaszcza badaniem polityki publicznej. O ile posiadanie ogromnej ilości informacji jest bezsprzecznie wartościowe, o tyle stają się mniej przydatne, a nawet bezużyteczne, jeśli nie możemy przetwarzać danych wystarczająco szybko i wydajnie. Zmniejszenie szumu informacyjnego we współczesnych mediach społecznościowych i wydobywanie wartościowych treści z internetowej debaty to prawdziwe wyzwanie, które staje się coraz trudniejsze zarówno ze względu na eksplozję treści, jak i szerzenie się dezinformacji. Z pomocą przychodzi Argument Mining.



Rys. 1. Proces automatycznego wydobywania argumentów z tekstu (źródło: Lippi & Torroni, 2016).

Bardzo powszechna jest opinia, że najlepszym sposobem modelowania internetowych debat publicznych jest wykorzystanie teorii demokracji deliberatywnej wywodzącej się od J. Habermasa i J. Rawlsa. Uważamy jednak, że można na to spojrzeć na inny sposób. W projekcie dokonamy empirycznej analizy modelu zwanego ‘agonistycznym’, opartego na pracach takich myślicielami jak C. Mouffe i H. Arendt. Chcemy znaleźć sposoby na pomiar skuteczności tych różnych stylów debat online. Przyjrzymy się również, co się dzieje, gdy debaty przybierają zły obrót – kiedy ludzie przestają się debatować rozsądnie i zaczynają szerzyć dezinformację lub atakować innych.

Z technicznego punktu widzenia opracujemy nowe podejście do eksploracji argumentów, które nazywamy „hybrydowym modelem AM”. Będzie to wykorzystywać mieszankę zaawansowanych modeli językowych do przewidywania argumentów, wraz z technikami badania znaczeń zdań. Długoterminowym celem tego projektu jest zbudowanie nowego, wielojęzycznego korpusu tekstowego dotyczącego większości aktualnych tematów debaty publicznej, w tym anglojęzycznych i polskich baz danych. Ponadto korpus tekstów powstały w wyniku projektu zostanie udostępniony na licencji Creative Commons innym badaczom, co będzie podstawą do dalszych projektów Argument Mining w języku polskim.