

**Specyfikacja modelu pokazowego sposobu użycia rozpoznawania mowy i systemów dialogowych  
w reklamach i punktach informacyjnych**

Realizacja oprogramowania implementującego funkcjonalność punktu informacyjnego udzielającego informacji o Systemie rozpoznawania mowy i sposobie wykorzystania dynamicznych modeli sieci Bayesa, zgłoszonym do ochrony patentowej pod nr P.403724 oraz o technologiach mowy oferowanych przez AGH w Krakowie związanych z tym systemem.

Realizacja usługi będzie polegać na:

1. Zaprojektowaniu scenariuszy rozmów informacyjnych o Systemie rozpoznawania mowy i sposobie wykorzystania dynamicznych modeli sieci Bayesa, zgłoszonym do ochrony patentowej pod nr P.403724 oraz o funkcjonalności i użyteczności rozpoznawania mowy i systemów dialogowych w reklamach i punktach informacyjnych

Opracowane scenariusze mają za zadanie opisywać dialog człowieka z automatycznym systemem dialogowym zgłoszonym do ochrony patentowej, w formule umożliwiającej swobodne wypowiedzi rozmówcy. System dialogowy powinien rozumieć zarówno słowa kluczowe (komendy) jak i pełne zapytania dotyczące wyszczególnionych tematów rozmowy. Liczba klas tematycznych uwzględnionych w dialogu wynosi 15. Związane one są wynalazkiem zgłoszonym do ochrony patentowej pod nr P.403724 oraz z podstawowymi zagadnieniami rozpoznawania mowy i systemów dialogowych w reklamach i punktach informacyjnych.

2. Implementacja wspomnianych scenariuszy dialogowych w VoiceXML, SRGS lub innej funkcjonalnie podobnej technologii.
3. Implementacja programu do obsługi stworzonych scenariuszy w wybranej technologii, realizującej przetwarzanie strumienia audio, integracja tej technologii z dostarczonym przez zamawiającego Systemem rozpoznawania mowy AGH Sarmata oraz wybranie najbardziej prawdopodobnego scenariusza na podstawie dostarczonych przez system rozpoznawania mowy AGH Sarmata hipotez rozpoznawania słów na dowolnym fragmencie nagrania, określenie pewności dopasowania strumienia audio do scenariuszy dialogowych na podstawie scenariusza oraz wartości wiarygodności testowanej hipotezy słowa.

Wspomniany program powinien realizować funkcjonalność *parsera* strumienia audio w oparciu o zdefiniowane w pkt. 2 gramatyki (opis dialogu, opis akceptowanych wypowiedzi) oraz zwracaną przez system rozpoznawania AGH – Sarmata wartość prawdopodobieństwa (wiarygodności) dowolnego (sprawdzanego przez *parser*) słowa, listy słów lub listy kombinacji słów, dla wybranego zakresu czasowego nagrania audio dla którego.

4. Integracja systemu rozpoznawania mowy AGH - Sarmata z serwerem dialogowym polegająca na stworzeniu odpowiedniego API programowego, współpracującego z aplikacjami stworzonymi w

języku C#, w technologii .NET oraz Mono. Stworzenie usługi typu Web Service (SOAP) do komunikacji między aplikacjami jako modelu pokazowego.

5. Implementacja interfejsu programowego (API) dla serwera usług pełniącego rolę automatycznej infolinii. Integracja rozwiązania z infolinią telefoniczną w technologii VoIP (SIP, zestawienie łącza), realizacja podłączenia linii telefonicznej VoIP do stworzonego serwera usług oraz uruchomienie systemu dialogowego na serwerze AGH w Krakowie pracującym pod kontrolą systemu operacyjnego Linux w wersji 64 bitowej (Debian/Ubuntu x64), wyposażonego w 2 procesory Intel Xeon 2 GHz oraz 12 GB pamięci operacyjnej RAM.
6. Sporządzenie dokumentacji technicznej oraz instalacyjnej zrealizowanego oprogramowania.
9. Scenariusz rozmowy ma także zawierać odniesienie do Programu Operacyjnego Innowacyjna Gospodarka oraz Europejskiego Funduszu Rozwoju Regionalnego oraz do projektu UDA-POIG.01.03.02-12-009/11

Szacowana wartość zamówienia 10 000 zł brutto