

Komunikacja Człowiek-Komputer

Wojciech Miksa



Laboratorium 5: Interfejs głosowy.

Laboratorium/projekt



SPOŁECZNA AKADEMIA NAUK

INTERFEJS GŁOSOWY

- 1 generacja: W 1952 Bell Labs zbudowało system Audrey, który rozpoznawał wypowiedziane (w US English, ofcourse) numery od 0 do 9.
- 2 generacja: opracowany w 1984 Interactive Voice Response firmy Speechworks umożliwiał zamówienie pokoju hotelowego i sprawdzenie statusu lotu. W Polsko-Japońskiej Akademii Technik Komputerowych (PJATK) jest automat rozpoznający mowę, łączący z wydziałami uczelni.
- 3 generacja: Apple Siri, Google Assistant, Microsoft Cortana potrafią udzielać sensownych odpowiedzi na różnorodne zadawane pytania.



DVI W EF2000

- Pierwszy system głosowy (Direct Voice Input – DVI) w samolocie bojowym.
- Moduł rozpoznawania mowy rozpoznaje 26 nie krytycznych komend, (bo jakie by były efekty: eject, fire, gear up - na płycie lotniska), np. związanych z namierzaniem celów, odciążających pilota.
- Każdy pilot tworzy własny szablon.
- Słownik: 600 słów, potem zmniejszony do 100.



ZALETY

- Nie wymaga użycia rąk.
- Nie wymaga fizycznego kontaktu ze sprzętem.
- Dostępny dla niewidomych.
- Może uzupełniać inne interfejsy.



WADY

- Ograniczona długość sekwencji skomplikowanych komend.
- Jest podatny na przypadkowe zakłócenia, „uruchomienie” komend przez osoby trzecie lub użytkownika jednocześnie rozmawiającego w innym kontekście.
- Podatny na niejednoznaczność związaną z wymową (o co chodzi-ossochozi?, trzy-czy (wymowa krakowska), włączać-włanczać), homofonami (masz-maż, mieć-mieść, kod-kot), wieloznacznością znaczeń (zamek, koło, salami),
- Bardziej złożony w implementacji niż interfejsy tekstowy czy nawet GUI.



PROJEKTOWANIE KONWERSACJI

- Conversional design jest obszarem wiedzy dotyczącym projektowania dialogów sztucznej inteligencji.

- Zasady:
 1. Kooperatywność, uprzejmość
 2. Zorientowanie na cel
 3. Uwzględnienie kontekstu
 4. Zwięzłość
 5. Kolejność (zakładamy, że raz mówi jedna, raz druga strona).



OSOLOWOŚĆ BOTA

Sposób postrzegania bota wpływa na zaufanie użytkownika oraz nastawienie emocjonalne:

- Ton głosu (powinien być uprzejmy i wyraźny).
- Sposób mówienia: formalny/nieformalny.
- Stereotypy wpływają na nastawienie do bota. Inaczej postrzegany jest bot z głosem kobiety/mężczyzny, dziecka/dorosłego/starca, głos ludzki/elektroniczny.
- Osobowość podobna do użytkownika wzbudza zaufanie.



TYPOWE BŁĘDY IRYTUJĄCE UŻYTKOWNIKÓW

- Niejasność, wieloznaczność wypowiedzi bota.
- Brak zwiezłości.
- Niewyraźne mówienie.
- Powtarzanie pytań o informacje, które były podane wcześniej.
- Pytania o potwierdzenie poprawności zrozumienia przez bota.



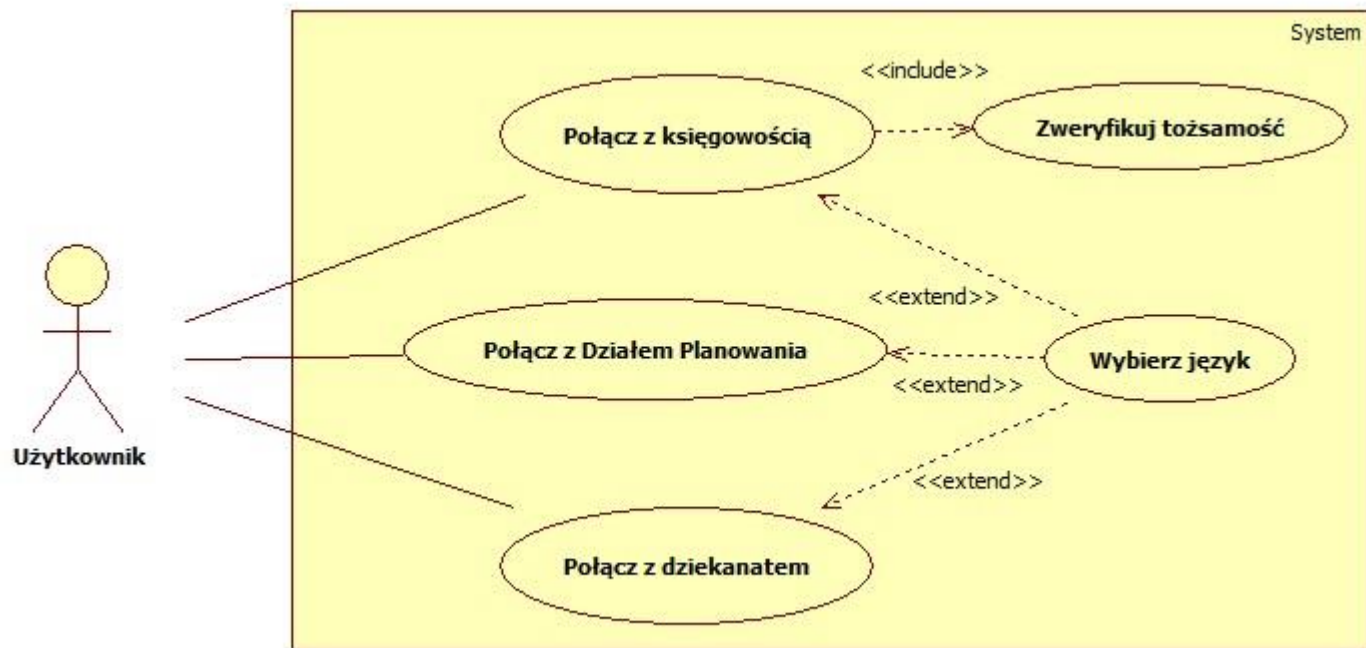
ZADANIE 5

- Celem ćwiczenia jest stworzenie diagramów sekwencji dla przypadków użycia automatycznej centrali telefonicznej obsługującej uczelnię.
- Artefakt wynikowe
 - 3 diagramy sekwencji dla każdego przypadku użycia.



DIAGRAM PRZYPADKÓW UŻYCIA

- Weryfikacja tożsamości może się odbyć za pomocą podania numeru PESEL lub daty urodzenia.
- Można przyjąć język polski i angielski.



Automatyczna centrala telefoniczna



KRYTERIA OCENY

- Na ocenę 5
 - 3 poprawne diagramy sekwencji (dla każdego bazowego przypadku użycia).
 - Uwzględnienie przypadków rozszerzających (np. weryfikacja tożsamości dla łączenia z księgowością (weryfikacja jest wymagana z powodu wrażliwości danych dotyczących płatności)).



DZIĘKUJĘ ZA UWAGĘ

Wojciech Miksa

wmiksa@san.edu.pl



SPOŁECZNA AKADEMIA NAUK

