



**AKADEMIA GÓRNICZO-HUTNICZA**

**im. Stanisława Staszica w Krakowie**

**WYDZIAŁ INŻYNIERII  
MECHANICZNEJ I ROBOTYKI**

---

# **Praca dyplomowa inżynierska**

**Paweł Jaciów**

*Imię i nazwisko*

**Inżynieria Akustyczna**

*Kierunek studiów*

**Implementacja bazy wiedzy  
awatara głosowego  
oraz analiza przeprowadzonych z nim rozmów**

*Temat pracy dyplomowej*

**dr inż. Bartosz Ziólko**

*Promotor pracy*

.....  
*Ocena*

Kraków, rok 2013/2014

Kraków, 7 I 2014

Imię i nazwisko: Paweł Jaciów  
Nr albumu: 241159  
Kierunek studiów: Inżynieria Akustyczna

### **OŚWIADCZENIE**

Świadomy odpowiedzialności karnej za poświadczanie nieprawdy oświadczam, że niniejszą inżynierską pracę dyplomową wykonałem osobiście i samodzielnie oraz nie korzystałem ze źródeł innych niż wymienione w pracy.

Jednocześnie oświadczam, że dokumentacja pracy nie narusza praw autorskich w rozumieniu ustawy z dnia 4 lutego 1994 roku o prawie autorskim i prawach pokrewnych (Dz. U. z 2006 r. Nr 90 poz. 631 z późniejszymi zmianami) oraz dóbr osobistych chronionych prawem cywilnym. Nie zawiera ona również danych i informacji, które uzyskałem w sposób niedozwolony. Wersja dokumentacji dołączona przeze mnie na nośniku elektronicznym jest w pełni zgodna z wydrukiem przedstawionym do recenzji.

Zaświadczam także, że niniejsza inżynierska praca dyplomowa nie była wcześniej podstawą żadnej innej urzędowej procedury związanej z nadawaniem dyplomów wyższej uczelni lub tytułów zawodowych.

.....  
*podpis dyplomanta*

Kraków, 7 I 2014

Imię i nazwisko: Paweł Jaciów  
Adres korespondencyjny: Ostrów 514, Przemyśl 37-700  
Temat pracy dyplomowej inżynierskiej: Implementacja bazy wiedzy awatara głosowego  
oraz analiza przeprowadzonych z nim rozmów  
Rok ukończenia: 2014  
Nr albumu: 241159  
Kierunek studiów: Inżynieria Akustyczna

### OŚWIADCZENIE

Niniejszym oświadczam, że zachowując moje prawa autorskie, udzielam Akademii Górniczo-Hutniczej im. S. Staszica w Krakowie nieograniczonej w czasie nieodpłatnej licencji niewyłącznej do korzystania z przedstawionej dokumentacji inżynierskiej pracy dyplomowej, w zakresie publicznego udostępniania i rozpowszechniania w wersji drukowanej i elektronicznej<sup>1</sup>.

Publikacja ta może nastąpić po ewentualnym zgłoszeniu do ochrony prawnej wynalazków, wzorów użytkowych, wzorów przemysłowych będących wynikiem pracy inżynierskiej<sup>2</sup>.

Kraków, 7 I 2014 .....  
data podpis dyplomanta

<sup>1</sup> Na podstawie Ustawy z dnia 27 lipca 2005 r. Prawo o szkolnictwie wyższym (Dz.U. 2005 nr 164 poz. 1365) Art. 239. oraz Ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (Dz.U. z 2000 r. Nr 80, poz. 904, z późn. zm.) Art. 15a. "Uczelnia w rozumieniu przepisów o szkolnictwie wyższym przysługuje pierwszeństwo w opublikowaniu pracy dyplomowej studenta. Jeżeli uczelnia nie opublikowała pracy dyplomowej w ciągu 6 miesięcy od jej obrony, student, który ją przygotował, może ją opublikować, chyba że praca dyplomowa jest częścią utworu zbiorowego."

<sup>2</sup> Ustawa z dnia 30 czerwca 2000r. – Prawo własności przemysłowej (Dz.U. z 2003r. Nr 119, poz. 1117 z późniejszymi zmianami) a także rozporządzenie Prezesa Rady Ministrów z dnia 17 września 2001r. w sprawie dokonywania i rozpatrywania zgłoszeń wynalazków i wzorów użytkowych (Dz.U. nr 102 poz. 1119 oraz z 2005r. Nr 109, poz. 910).

Kraków, 7 I 2014

**AKADEMIA GÓRNICZO-HUTNICZA  
WYDZIAŁ INŻYNIERII MECHANICZNEJ I ROBOTYKI**

**TEMATYKA PRACY DYPLOMOWEJ INŻYNIERSKIEJ**  
dla studenta IV roku studiów stacjonarnych

**Paweł Jaciów**  
*imię i nazwisko studenta*

TEMAT PRACY DYPLOMOWEJ INŻYNIERSKIEJ:  
Implementacja bazy wiedzy awatara głosowego  
oraz analiza przeprowadzonych z nim rozmów

*Promotor pracy:* dr inż. Bartosz Ziółko

*Recenzent pracy:* dr hab. inż. Marek Kisiel-Dorohinicki

.....  
*Podpis dziekana:*

**PLAN PRACY DYPLOMOWEJ**

1. Omówienie tematu pracy i sposobu realizacji z promotorem.
2. Zebranie i opracowanie literatury dotyczącej tematu pracy.
3. Wykorzystanie opracowania literatury przy implementacji.
4. Analiza zapisów rozmów, ich omówienie i zatwierdzenie przez promotora.
5. Opracowanie redakcyjne.

Kraków, 7 I 2014

*data*

.....  
*podpis dyplomanta*

**TERMIN ODDANIA DO DZIEKANATU:**

..... **2014 r.**

.....  
*podpis promotora*

Kierunek: Inżynieria Akustyczna

Paweł Jaciów

**Praca dyplomowa inżynierska**

Implementacja bazy wiedzy awatara głosowego  
oraz analiza przeprowadzonych z nim rozmów

Opiekun: dr inż. Bartosz Ziółko

### STRESZCZENIE

W ramach niniejszej pracy najpierw, na podstawie przeglądu literaturowego, scharakteryzowano dialog ludzki i odniesiono jego cechy do komunikacji człowieka z komputerem. Opisano w sposób ogólny działanie i strukturę systemów dialogowych, wymieniono kilka istniejących ich zastosowań, zwrócono uwagę na będący przedmiotem niniejszej pracy produkt Wirtualny Doradca. Jest to system, z którym użytkownik komunikuje się za pośrednictwem awatara umieszczonego w witrynie. Goście witryny mogą wprowadzać wypowiedzi w języku naturalnym za pomocą klawiatury. Następnie, zgodnie z zaleceniami literaturowymi, zaimplementowano bazę wiedzy specjalistycznej systemu zawierającą informacje o Akademii Górniczo-Hutniczej, Zespole Przetwarzania Sygnałów, kierunku studiów Inżynieria Akustyczna oraz dziedzinach przetwarzanie sygnałów i technologie mowy. Opisano cały proces, z uwzględnieniem zastosowanych technik. W dalszej kolejności, na podstawie statystyk systemu i zapisanych przez niego konwersacji z użytkownikami, dokonano analizy przeprowadzonych rozmów. Dane przeanalizowano pod kątem statystyk liczbowych, a także treści przekazywanej systemowi. Opisano wyniki tego przeglądu, przeprowadzono dyskusję na ich temat oraz sformułowano spostrzeżenia. Rezultaty pokazują, że ludzie są chętni do interakcji z systemem i postrzegają w nim towarzyskość.

Field of Study: Acoustic Engineering

Paweł Jaciów

**Engineer Diploma Thesis**

Implementation of a speaking avatar's knowledge base  
and performed conversations analysis

Supervisor: dr inż. Bartosz Ziółko

SUMMARY

In this paper first, based on literature research, human dialogue was characterized and its relevance to human-computer communication was pointed out. The structure of dialogue systems and how these work was generally described. Some examples of dialogue system applications were mentioned with focus on a product Virtual Assistant that is this paper's subject. With this system user communicates through an avatar embedded in a webpage. Site's guests can input their utterances in natural language as a text using a keyboard. Next, according to literature guidelines, specialized knowledge base was implemented. The base contains information about AGH University of Science and Technology, Signal Processing Group, field of study Acoustical Engineering and domains of signal processing and speech technologies. The whole process was described including applied techniques. Further, based on system's statistical data and saved dialogues with users, performed conversations analysis was carried out. Both numeric data and content of utterances were analyzed. The results were described, a short discussion was conducted and some observations were formulated. Results show that people are willing to interact with the system and see some sociality in it.

Podziękowania dla dr inż. Bartosza Ziółki  
za wsparcie, promocję i zaufanie.

## Spis treści

Spis treści.....	8
Wstęp.....	9
Uzasadnienie wyboru tematu pracy.....	9
Cele pracy.....	9
Streszczenie poszczególnych rozdziałów .....	10
1    Wprowadzenie .....	11
1.1.    Właściwości dialogu .....	11
1.1.1.    Podejmowanie tur .....	11
1.1.2.    Akty mowy .....	12
1.1.3.    Utwierdzenie.....	12
1.1.4.    Implikatury .....	12
1.2.    System dialogowy .....	13
1.2.1.    Działanie i struktura .....	13
1.2.2.    Interpretacja języka naturalnego i zarządzanie dialogiem.....	15
1.3.    Istniejące zastosowania .....	16
1.4.    Charakterystyka Wirtualnego Doradcy .....	17
2    Implementacja bazy wiedzy .....	18
2.1.    Projektowanie skierowane na użytkownika.....	18
2.2.    Budowa bazy wiedzy i zastosowane techniki.....	19
2.2.1.    Wiedza specjalistyczna.....	19
2.2.2.    Doprecyzowania, odpowiedzi wymijające, powitania, zachęty do rozmowy .....	22
2.2.3.    Skrypty .....	23
3    Analiza przeprowadzonych rozmów .....	25
3.1.    Podstawowe statystyki .....	25
3.2.    Jak ludzie rozmawiają z Wirtualnym Doradcą .....	30
4    Dyskusja i wnioski .....	34
Zakończenie .....	37
Podsumowanie.....	37
Kontynuacja pracy.....	38
Uwagi.....	38
Bibliografia .....	39
Spis rysunków .....	40
Załączniki .....	40



# **Wstęp**

## **Uzasadnienie wyboru tematu pracy**

Wybór tematu pracy jest umotywowany zaciekawieniem autora technologią systemu dialogowego języka naturalnego. Przez system taki rozumie się algorytmy umożliwiające rozmowę człowieka z komputerem w sposób jak najbardziej zbliżony do konwersacji dwojga ludzi [1]. Autor w trakcie pogłębiania wiedzy z zakresu technologii mowy napotkał produkt Wirtualny Doradca firmy Stanusch Technologies i uznał za interesujące sposób, w jaki obiekt przyciąga uwagę internautów oraz jego możliwości interakcji z użytkownikiem. Aplikacja umożliwia prowadzenie z nią dialogu. Odpowiada na dowolne pytanie, potrafi argumentować swoje wypowiedzi, zaczepiać użytkownika i swoim zachowaniem zwracać na siebie jego uwagę. Dodatkowym czynnikiem wpływającym na wybór tematu był fakt, że to oprogramowanie zostało udostępnione Uczelni.

## **Cele pracy**

Za cele niniejszej pracy obrano przygotowanie bazy wiedzy specjalistycznej Wirtualnego Doradcy uruchomionego na stronie internetowej Zespołu Przetwarzania Sygnałów Akademii Górniczo-Hutniczej oraz dokonanie analizy rozmów przeprowadzonych z nim przez gości witryny. Przez wiedzę specjalistyczną ma się na myśli zbiór zagadnień specyficznych dla danej realizacji Wirtualnego Doradcy, w odróżnieniu od wspólnej dla każdej jego realizacji wiedzy ogólnej. W dużym uproszczeniu można powiedzieć, że implementacja takiej bazy wiedzy polega na wprowadzaniu do niej faktów. Mianem faktu określa się przypisaną do zestawu pytań odpowiedź systemu wyświetlaną użytkownikowi po sformułowaniu jednego z tych zapytań. Aplikacja umożliwia dostęp do zapisów i statystyk z odbytych konwersacji. Informacje te uwzględniają między innymi dane liczbowe i wykresy dotyczące ilości i czasów rozmów, zadawanych pytań oraz udzielanych odpowiedzi.

Treść opracowania może stanowić pomoc dla osób, które zamierzają zajmować się bazą danych zbliżonych systemów oraz zawierać wskazówki odnośnie jej tworzenia

i zarządzania nią. Może być też źródłem informacji o tym, czego można spodziewać się po udostępnieniu użytkownikom podobnej aplikacji poprzez umieszczenie jej na swojej stronie internetowej.

## **Streszczenie poszczególnych rozdziałów**

Główna część pracy, poza wstępem i zakończeniem, składa się z czterech rozdziałów.

Rozdział 1 *Wprowadzenie* wymienia właściwości konwersacji, przybliża tematykę systemów dialogowych, opisuje ich działanie i strukturę, wspomina o istniejących zastosowaniach oraz charakteryzuje system będący przedmiotem tej pracy.

Rozdział 2 *Implementacja bazy wiedzy* stanowi opis procesu budowy bazy wiedzy specjalistycznej Wirtualnego Doradcy z uwzględnieniem metod i zaleceń sugerowanych w literaturze. Wspomniano tu o testach i wprowadzonych na podstawie ich wyników korektach.

Rozdział 3 *Analiza przeprowadzonych rozmów* obejmuje przegląd danych statystycznych zarówno wygenerowanych przez oprogramowanie Wirtualnego Doradcy, jak i opracowanych samodzielnie na podstawie zapisów konwersacji dostępnych z jego systemu.

Rozdział 4 *Dyskusja i wnioski* zawiera przemyślenia na temat implementacji bazy wiedzy oraz spostrzeżenia dotyczące analizy przeprowadzonych rozmów.

# 1 Wprowadzenie

Język od zawsze jest znakiem człowieczeństwa, a jego najbardziej podstawowym i uprzywilejowanym obszarem jest dialog. Z pewnością jest to pierwszy rodzaj języka, jaki poznajemy jako dzieci, a dla większości z nas jest to rodzaj języka, którym posługujemy się najczęściej, cokolwiek robimy [2]. Jest to atrakcyjna metoda komunikacji, ze względu na swoją naturalność, efektywność i taniość. W odniesieniu do porozumiewania się człowieka z maszyną, dialog jest współdziałaniem pomiędzy użytkownikiem i komputerem w celu osiągnięcia szczególnego celu [3]. Przy obecnym rozwoju maszyn, wzroście mobilności i zapotrzebowania na szybkie zdobywanie informacji dąży się do zoptymalizowania tej komunikacji. Pożądane jest, aby komputer porozumiewał się z człowiekiem tak, jak sam człowiek konwersuje z drugim człowiekiem. Budowa takiego systemu wymaga zrozumienia komunikacji między ludźmi [4].

## 1.1. Właściwości dialogu

W tym podrozdziale, w oparciu o [2], scharakteryzowano ludzką konwersację i odniesiono jej cechy do komunikacji człowieka z komputerem.

### 1.1.1. Podejmowanie tur

Rozmowa składa się z następujących po sobie tur. W trakcie każdej tury jeden rozmówca wypowiada się (jest mówcą), podczas gdy drugi odbiera ten przekaz (jest słuchaczem). Następnie mówca i słuchacz zamieniają się rolami i w ten sposób podejmują kolejne tury. Słuchacz wyciąga wnioski o zamierzonym znaczeniu wypowiedzi mówcy.

W dialogu obecne są dwuczęściowe struktury typu pytanie-odpowieź. Nazywa się je parami przyległymi (ang. *adjacency pairs*). W parach takich druga wypowiedź jest poniekąd wymagana przez pierwszą. Inne konstrukcje tego typu to na przykład powitanie-powitanie, komplement-umniejszenie, żądanie-zgoda/odmowa.

### **1.1.2. Akty mowy**

Wypowiedź w dialogu jest akcją wykonywaną przez mówcę. Wyróżniono trzy akty mowy: lokucyjny (powiedzenie czegoś, czemu można przypisać jakieś znaczenie); illokucyjny (prośba, odpowiedź, obietnica itp. przez wypowiedzenie zdania); perlokucyjny (wywarcie poprzez wypowiedź wpływu na uczucia, myśli lub zachowanie słuchacza). Modyfikując później ten podział, zasugerowano klasyfikację na pięć aktów: asertywne (wyrażenie opinii); dyrektywne (próba wpłynięcia na zachowanie słuchacza); komisywne (zobowiązanie się do przyszłego działania); ekspresywne (wyrażenie emocji lub postawy); deklaratywne (zmienianie stanu rzeczy przez wypowiedź).

### **1.1.3. Utwierdzenie**

Powyższe może sugerować, że każda tura lub wypowiedź mogłaby być rozpatrywana jako osobna akcja mówcy. Jednak nie są one niezależne i niepowiązane. Wręcz przeciwnie, rozmowa jest wspólną akcją wykonywaną i przez mówcę i przez słuchacza. Wskazuje na to ciągle ustalanie wspólnych podstaw (ang. *common ground*), czyli zbioru tych samych przekonań dla obu uczestników dialogu. Potrzeba osiągnięcia wspólnych podstaw oznacza, że słuchacz musi utwierdzać wypowiedzi mówcy dając jasno do zrozumienia, że zrozumiał jego znaczenie i intencję. I tak, zasada ścisłości (ang. *principle of closure*) mówi, że mówca podejmując akcję wymaga wystarczającego potwierdzenia, że osiągnął cel wykonując ją. Jednakże, kiedy mówca nie osiągnął celu, słuchacz musi to wskazać, żeby jeszcze raz spróbować porozumieć się. Określono pięć metod utwierdzenia, od najsłabszej do najmocniejszej: słuchacz kontynuuje uważne słuchanie mówcy; słuchacz rozpoczyna swoją wypowiedź na temat związany z tym, co właśnie powiedział mówca; słuchacz przytakuje kiwając głową lub głosami potwierdzenia typu 'mm hmm'; słuchacz parafrazuje to, co zrozumiał lub wspólnie dokańcza wypowiedź mówcy; słuchacz pokazuje słowo w słowo, co zrozumiał.

### **1.1.4. Implikatury**

Czasami mówcy oczekują od słuchaczy wysnucia pewnych wniosków. Inaczej mówiąc, przekazują więcej informacji niż wydaje się to wynikać z samych słów wypowiedzi. Zaproponowano maksyma, które umożliwiają słuchaczom odczytanie tych

ukrytych informacji: przekazuj dokładnie tyle informacji, ile jest wymagane; nie mów czegoś, o czym wiesz, że jest fałszywe, ani nie mów czegoś, na co brakuje ci dowodu; odnoś się do tematu; bądź zrozumiały.

Podobnie jak w przypadku konwersacji między ludźmi, komunikacja człowieka z komputerem składa się z następujących po sobie tur. Pary przyległe ustalają pewne oczekiwania dialogowe i odgrywają istotną rolę w modelowaniu rozmowy. System powinien być projektowany tak, żeby spełniać oczekiwania użytkownika, który po pytaniu spodziewa się odpowiedzi, po powitaniu powitania itp. Utwierdzenie jest tak samo ważne w rozmowie człowiek-maszyna jak w rozmowie ludzi. Wykazano, że ludzie czują się zdezorientowani, gdy system nie daje im wyraźnego potwierdzenia. Akty mowy i implikatury, a właściwie ich przesłanie, powinny być właściwie interpretowane przez system.

## **1.2. System dialogowy**

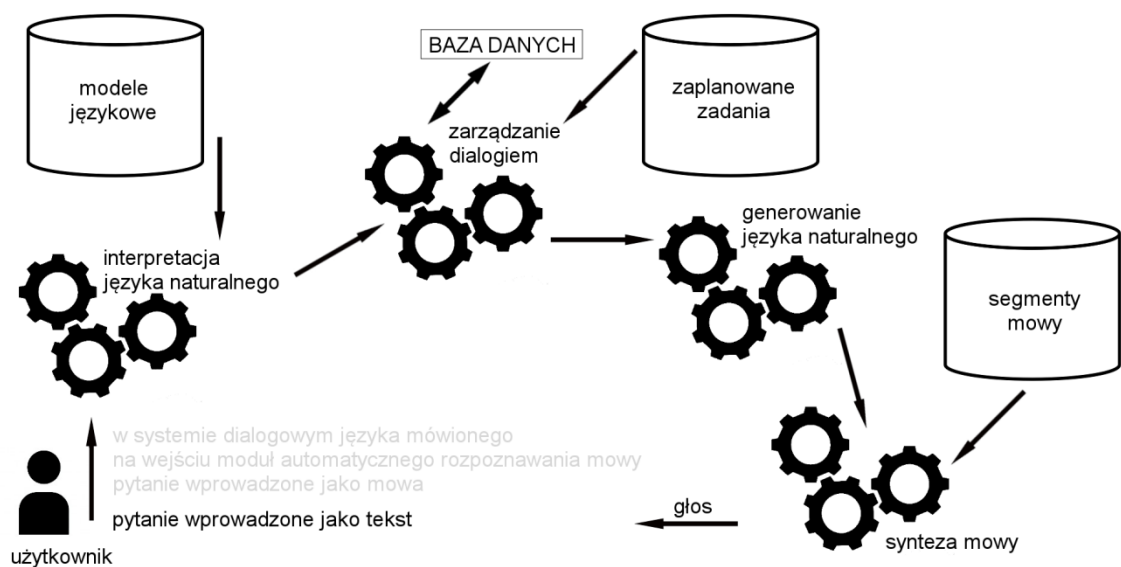
System dialogowy umożliwia komunikację człowieka z komputerem w języku naturalnym, przy czym słowo naturalny służy odróżnieniu od języka programowania. Należy w tym miejscu zaznaczyć różnicę między systemem dialogowym a systemem dialogowym języka mówionego (ang. *dialogue system* oraz *spoken dialogue system*). Ten drugi umożliwia wprowadzanie wypowiedzi użytkownika wykorzystując mowę, podczas gdy pierwszy zwykle potrzebuje treści w postaci tekstu wprowadzonego na przykład z klawiatury. Oba systemy różnią się tym, że drugi z nich jest dodatkowo wyposażony w moduł automatycznego rozpoznawania mowy (ang. *automatic speech recognition*, ASR) na wejściu. Ponieważ awatar głosowy będący przedmiotem niniejszej pracy nie jest wyposażony w taki moduł i wymaga pytań wprowadzanych jako tekst, w dalszej części pracy autor skupił się na opisie systemu dialogowego z pominięciem komponentu ASR.

### **1.2.1. Działanie i struktura**

Interakcja człowieka z systemem dialogowym przebiega jak podano w [4]. Użytkownik wykonuje akt dialogu, czyli wprowadza na przykład pytanie, żądanie lub

polecenie. Treść tej wypowiedzi jako potok słów trafia do modułu interpretacji języka naturalnego (ang. *natural language understanding*, NLU), gdzie podejmowana jest próba zrozumienia jej. Jeśli moduł jest skomplikowany, może dokonywać analizy składni wypowiedzi (określenie składników struktury wypowiedzi), a także korzystać z modeli znaczeniowych języka (określenie znaczeń składników). Następnie, historia dialogu jest aktualizowana o otrzymane znaczenie. Dalej, akcja systemu jest określana przez moduł zarządzania dialogiem (ang. *dialogue manager*), który ogólnie steruje przebiegiem dialogu. Działa w oparciu o zaplanowane zadania i korzysta z bazy danych. Zarządzanie dialogiem produkuje odpowiedź systemu, która jest zapisywana w historii dialogu. Informacje powiązane z dialogiem są śledzone, aby wspomóc zarządzanie nim. Mogą to być informacje o tym, co zostało dotąd powiedziane i co zostało utwierdzone. Odpowiedź trafia jako strumień słów języka naturalnego do syntezy mowy lub jest odgrywana użytkownikowi jako gotowe nagranie. Aktem dialogu wykonanym przez system może być na przykład odpowiedź na pytanie, utwierdzenie lub żądanie wprowadzenia dodatkowych informacji przez użytkownika.

Strukturę systemu dialogowego stanowią wspomniane zarządzanie dialogiem i wspierające je komponenty. Prosty schemat działania i struktury systemu dialogowego przedstawiono na Rys. 1.1. Uwzględniono w nim cztery podstawowe elementy: interpretacja języka naturalnego, zarządzanie dialogiem, generowanie języka naturalnego i synteza mowy.



Rys. 1.1. Prosty schemat działania i struktury systemu dialogowego.

### 1.2.2. Interpretacja języka naturalnego i zarządzanie dialogiem

Według [5] powszechna technologia interpretacji języka naturalnego jest oparta na rozpoznawaniu słów kluczowych w wypowiedzi użytkownika. Wydaje się to uzasadnione, bo jak można przeczytać w [6], w skomplikowanych wypowiedziach tylko niektóre wyrazy niosą informację, a pozostałe są nieistotne. System może rezygnować z modelowania języka przy interpretacji i od razu przechodzić od rozpoznanych słów do przypisanych im operacji. Zrozumienie języka naturalnego powinno uwzględniać dodatkowo odniesienia anaforyczne, czyli pamiętać kontekst rozmowy. Konieczne jest też poprawne zrozumienie wyrażen eliptycznych i urywek zdań, żeby użytkownik mógł pozwolić sobie na opuszczanie wyrazów, których można domyślić się w szerszym kontekście.

Są dwa główne typy zarządzania dialogiem: zadaniowe (ang. *task-oriented*) i nie zadaniowe (ang. *nontask-oriented*). Zadaniowe umożliwiają poprzez dialog wykonanie pewnego zadania, a nie zadaniowe pozwalają na swobodną, interaktywną konwersację. Można łączyć obie funkcjonalności [4]. Zarządzając dialogiem, system powinien wykazywać się przemienną inicjatywą (ang. *mixed initiative*), umożliwiając użytkownikowi zadawanie pytań w języku naturalnym w dowolnym momencie dialogu. Taka komunikacja jest lepiej odbierana, bo stwarza iluzję realnych możliwości interakcji i użytkownicy mogą stosować swoje naturalne nawyki rozmów [5]. Inne style zarządzania dialogiem to kontrolowany przez system (ang. *system initiative*) i kontrolowany przez użytkownika (ang. *user initiative*). W pierwszym system zadaje precyzyjne pytania i wymaga od użytkownika bezpośredniej odpowiedzi, a w drugim wykonuje polecenia, odpowiada na pytania i oczekuje na kolejne komendy, pozostając pasywnym.

Przy planowaniu zarządzania dialogiem można wykorzystać znajomość rozmów między ludźmi. Jednak są dowody na to, że ludzie zachowują się inaczej gdy są świadomi prowadzenia interakcji z maszynami [4]. Interakcje między ludźmi nie mogą być bezpośrednio zamodelowane w systemie dialogowym. Jednym ze sposobów na zebranie danych dotyczących symulowanych dialogów człowiek-komputer jest użycie metody Czarnoksiężnika z Krainy Oz (ang. *Wizard of Oz*). Zakłada ona, że użytkownik wierzy, że rozmawia z maszyną, a naprawdę praca maszyny jest symulowana przez ludzkiego operatora (nazywanego właśnie czarnoksiężnikiem). Metoda jest kosztowna w kontekście czasu i wysiłku włożonego przez uczestników interakcji.

### 1.3. Istniejące zastosowania

Systemy dialogowe mogą spełniać różne role: przekierowywać i obsługiwać zapytania telefoniczne, przeglądać dane, rozwiązywać problemy i pomagać w wykonywaniu zadań, doradzać i udzielać informacji, nauczać, być przeciwnikiem czy po prostu partnerem do pogaduszek. Wymienione zostało kilka ich zastosowań.

Eliza (Weizenbaum, 1966) – program imitujący odpowiedzi psychoterapeuty z sukcesem, mimo swojej prostoty rozpoznawania i wybierania wzorców wypowiedzi.

How May I Help You? (AT&T, 2001) – system przekierowywania połączeń witający pytaniem *W czym mogę pomóc?*, łączący z odpowiednim personelem na podstawie odpowiedzi dzwoniącego.

Mercury (MIT, ~2000) – telefoniczny dostęp do wyszukiwania i rezerwacji lotów z dużą bazą danych, jednak bez możliwości zakupu ze względu na charakter badawczy.

Clarissa (NASA, 2005) – pierwszy system dialogowy użyty w kosmosie na Międzynarodowej Stacji Kosmicznej do wspomagania wykonywanych tam procedur.

Façade (Procedural Arts, 2005) – gra fabularna, w której korzystając z systemu dialogowego wpływa się na przebieg sprzeczki pary znajomych, interaktywny dramat.

ITSpoke (University of Pittsburgh, 2006) – oparty na głosie system korepetycji z fizyki.

Sierżant Blackwell (USC ICT, 2006) – połączenie systemu dialogowego z wyświetlaniem żołnierza w skali człowieka, wystawiony w muzeum w Houston.

Wirtualny Doradca (Stanusch Technologies, 2007) – system pełniący głównie rolę konsultanta osadzonego w witrynie, opisany szerzej w kolejnym podrozdziale.

Portal Głosowy (Primespeech, 2008) – system interaktywnej obsługi przez telefon zastosowany do przełączania głosowego, automatycznego zamawiania taksówek, dostępu do notowań giełdy, wyszukiwania połączeń komunikacji miejskiej.

Siri (Apple, 2011) – osobisty asystent i nawigator, umożliwiający w języku naturalnym sterowanie telefonem, wyszukiwanie informacji, nawigowanie i inne.

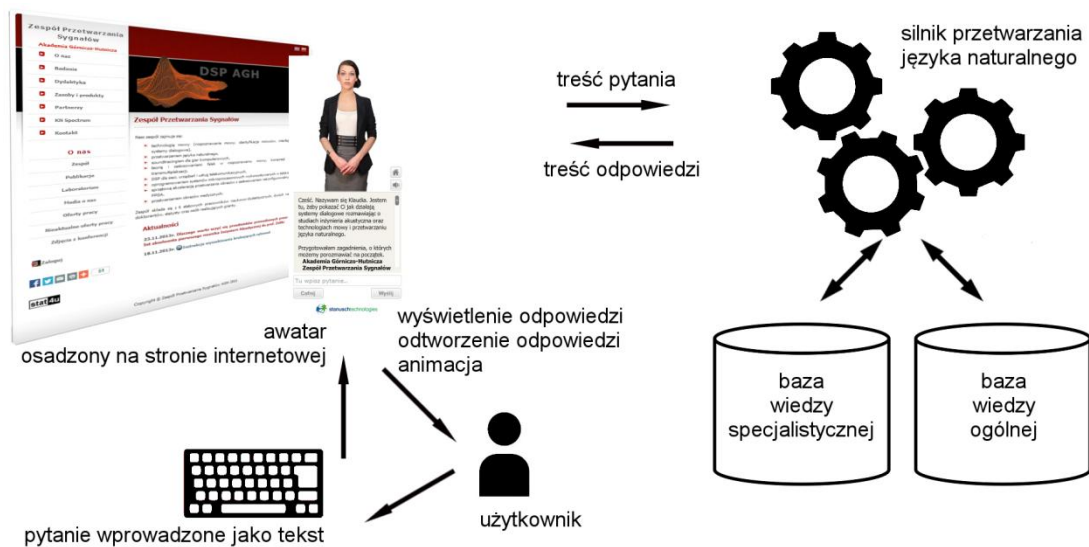
Autor niniejszej pracy nie testował technologii innych niż Wirtualny Doradca. Przedstawionych informacji zaczerpnął ze stron internetowych tych rozwiązań. Odnośniki do poszczególnych witryn dołączono w Zał. 1.

Obecnie prowadzi się wiele prac nad rozwojem systemów dialogowych. Świadczą o tym liczne patenty w tej dziedzinie, na przykład [7, 8, 9, 10].



## 1.4. Charakterystyka Wirtualnego Doradcy

W aplikacji Wirtualnego Doradcy można wyróżnić dwa główne elementy: awatar głosowy umieszczony na stronie internetowej oraz silnik przetwarzania języka naturalnego, znajdujący się razem z bazami wiedzy w infrastrukturze producenta. Awatar pośredniczy w komunikacji między użytkownikiem a rdzeniem systemu. Odbiera wprowadzone przez użytkownika w formie tekstu pytania w języku naturalnym i przesyła je do silnika przetwarzania języka naturalnego. Po przetworzeniu pytania przez silnik, awatar odbiera odpowiedź z bazy wiedzy i wyświetla ją użytkownikowi. Oprócz tego, otrzymuje komendy syntezy mowy i animacji, odtwarza głos i wideo. Ma postać człowieka i własny głos. Przy nim znajdują się pola tekstowe na wpisywanie pytań i wyświetlanie odpowiedzi oraz przyciski funkcyjne. Może przybrać dowolny wygląd. Po otrzymaniu pytania od awatara silnik przetwarzania języka naturalnego najpierw bada obecność części mowy i sprawdza jak odmienione są słowa. Następnie dokonuje korekt ortografii i tak zwanych literówek, po czym wykrywa części zdania. Na podstawie słów kluczowych przeszukuje bazę wiedzy i dopasowuje do pytania odpowiedni fakt i przesyła go do awatara. Jeśli pytanie odnosi się do kilku faktów, żądane jest sprecyzowanie. Jeśli odpowiedzi nie ma w bazie, może zostać wyszukana w Internecie. W międzyczasie jest sporządzany zapis rozmowy. Wirtualny Doradca został napisany w językach programowania PHP i Java i wykorzystuje bazę danych MySQL [11]. Schemat komunikacji z Wirtualnym Doradcą przedstawiono na Rys. 1.2.



Rys. 1.2. Schemat komunikacji z Wirtualnym Doradcą.

## **2 Implementacja bazy wiedzy**

Jak wspomniano we wstępie, zaimplementowano bazę wiedzy specjalistycznej, czyli specyficzną dla danej realizacji Wirtualnego Doradcy. W tym rozdziale opisano proces tworzenia takiej bazy z uwzględnieniem wskazówek z literatury.

### **2.1. Projektowanie skierowane na użytkownika**

Implementację bazy wiedzy rozpoczęto zgodnie z pierwszą z zasad projektowania takich systemów, wymienionych w [2]: przestudiuj użytkownika i zadanie. Mówi ona o zrozumieniu natury zadania i potencjalnych użytkowników poprzez wywiad z nimi i zbadanie podobnych systemów. Autor niniejszej pracy zapoznał się z działaniem Wirtualnego Doradcy umieszczonego w witrynie producenta [11], co dało mu pogląd na docelowe funkcjonowanie systemu. Następnie przemyślał kim będą potencjalni użytkownicy i jaki zbiór zagadnień może ich interesować. Za główną grupę użytkowników uznał studentów, a także maturzystów, którzy mogą być zainteresowani rekrutacją na studia. Zdecydował się podzielić zbiór zagadnień na cztery kategorie: Akademia Górniczo-Hutnicza, Zespół Przetwarzania Sygnałów, kierunek studiów Inżynieria Akustyczna, przetwarzanie sygnałów i technologie mowy.

Druga zasada sugeruje zbudowanie prototypu systemu i przeprowadzenie symulacji. W przytaczanym źródle odniesiono się do opisaną w podrozdziale 1.2.2. metody Czarnoksiężnika z Krainy Oz jako narzędzia do przetestowania architektury przed implementacją. W niniejszej pracy, ze względu na wspomnianą wcześniej kosztowność metody, zastosowano własną, starając się zachować sens sugerowanej. Zalecana metoda służy spostrzeżeniu różnic między sposobem rozmowy człowieka z człowiekiem, a człowieka z komputerem. I tak, autor niniejszej pracy najpierw przeprowadził wywiad wśród maturzystów oraz osób powiązanych z Inżynierią Akustyczną zadając pytanie o co związane ze studiami zapytaliby, dodatkowo prosząc o przykładowe pytania. Oprócz tego postarał się samodzielnie przewidzieć część zagadnień. Następnie zaimplementował wstępną bazę wiedzy z uwzględnieniem zdobytych informacji. Taki prototyp udostępnił podobnemu gronu osób, analizował w dalszej kolejności zadawane pytania i porównywał je z zaproponowanymi wcześniej.

Trzecia zasada dotyczy testowania projektu z udziałem użytkowników. Testy były przeprowadzane zarówno jeszcze przed uruchomieniem awatara, jak i już po jego umieszczeniu w witrynie. Przeprowadzano je prosząc grupy osób o krótką rozmowę z awatarem. Ich celem było sprawdzenie, czy kombinacje przykładowych pytań przypisanych do poszczególnych faktów są wystarczające do poprawnego dopasowywania odpowiedzi. Rezultatem było wprowadzanie korekt, gdy system udzielał niewłaściwych odpowiedzi, mimo że znał właściwą. Oprócz tego, w wyniku testów pojawiały się pytania, na które nie było odpowiedzi w bazie. Wtedy dodawano do niej brakujące fakty.

## **2.2. Budowa bazy wiedzy i zastosowane techniki**

Bazę wiedzy zaimplementowano z wykorzystaniem Panelu Administracyjnego. Jest to narzędzie, które umożliwia zarządzanie bazą, dodawanie i edycję faktów, testowanie wprowadzanych zmian i przegląd przeprowadzonych rozmów. Implementacja polegała na tworzeniu faktów o odpowiedniej treści, przypisywaniu im przykładowych pytań i opcjonalnie dodaniu pewnych akcji za pomocą skryptów.

### **2.2.1. Wiedza specjalistyczna**

Jak wspomniano w poprzednim podrozdziale, baza wiedzy specjalistycznej została podzielona na cztery kategorie. Fakty z kategorii Akademii Górniczo-Hutniczej dotyczą w szczególności jej struktury, rekrutacji, stypendiów i pomocy socjalnej, kół naukowych i tradycji studenckich. Druga kategoria informuje o pracach badawczych Zespołu Przetwarzania Sygnałów, jego składzie, zasobach i produktach oraz o zawartości stron jego witryny. Kategoria Inżynieria Akustyczna przedstawia między innymi opis kierunku i przebieg studiów, wymagania dla kandydatów, perspektywy dla absolwentów, nauczane przedmioty na poszczególnych stopniach, obieralne specjalności, laboratoria oraz różne szczegóły i ciekawostki dotyczące kierunku. Czwarta kategoria wyjaśnia podstawowe i najczęściej spotykane pojęcia z dziedzin przetwarzania sygnałów i technologii mowy. Przy ustalaniu treści faktów autor niniejszej pracy opierał się na oficjalnych informacjach dostępnych na stronach

internetowych Uczelni, Zespołu i Kierunku, własnych doświadczeniach i wiedzy oraz definicjach literaturowych, w szczególności z [1].

Tradycyjne metody komunikacji z komputerem polegające na wybraniu jednej z kilku opcji są odbierane jako nieatrakcyjne [1]. Tu niejednokrotnie umożliwiono wybór spośród takich opcji, które jednak wcale nie ograniczają swobody użytkownika. Pełnią natomiast rolę sugestii co do dalszego przebiegu rozmowy, możliwości podjęcia kolejnego tematu lub podpowiedzi jak można sformułować następne pytanie. Stanowią raczej skróty, które umożliwiają przejście do kolejnej odpowiedzi jednym kliknięciem. Użytkownik nie musi sam wymyślać i wpisywać pytania, co oszczędza jego czas. Mimo wyświetlenia opcji dialogu, systemowi ciągle można zadać dowolne pytanie. Podobnym rozwiązaniem jest zaprezentowanie dodatkowych informacji. W pewnych sytuacjach system proponuje, że opowie o kolejnym zagadnieniu. Użytkownik może zgodzić się lub sterować rozmową dalej według uznania, bo system zawsze umożliwia otwarte wywołanie (ang. *open prompt*). Na przykład: *Próbkowanie to (...). Przypomnieć Ci twierdzenie o próbkowaniu?* W przypadku powtarzalności takiej sytuacji, zróżnicowano kolejne pytania zadawane przez awatar, aby zachować naturalność dialogu, uniknąć schematów. Na przykład, opowiadając o Uczelni, system będzie kolejno pytał: *Przedstawić Ci infrastrukturę uczelni?*, *Masz ochotę poznać historię AGH?*, *Posłuchasz o tradycjach studenckich?*, *Interesuje Cię rozwój naukowy poza zajęciami?* i tak dalej. Oprócz tego, w treściach odpowiedzi umieszczono odnośniki do innych faktów, jeśli pojawiały się opisujące je słowa kluczowe. Daje to użytkownikowi możliwość bezpośredniego przejścia do kwestii, która go zainteresuje lub nie będzie zrozumiana. Ułatwia to nawigację. Poza tym, dołączono odnośniki do stron internetowych w przypadkach, gdy obszerność zagadnienia była zbyt duża, żeby umieścić je w fackie lub gdy informacja często się zmienia. Ze wskazanej strony użytkownik będzie mógł zaczerpnąć informacji szczegółowych i aktualnych.

Kilka odpowiedzi uznano za zbyt długie. Zastosowano w nich przekierowanie do krótszych, bez konieczności przypisywania tych samych pytań do nowej odpowiedzi. W niektórych faktach użyto zmiennych, na przykład przechowującej proponowane tematy rozmowy, co pozwala na edycję tylko jej zamiast każdego faktu, w którym chciałoby się zaproponować tematy rozmowy. Gdzieś tam zaimplementowano wybór między formą męską a żeńską, w zależności od rozpoznanego imienia użytkownika, na przykład: *O czym [ chciałbyś / chciałabyś ] porozmawiać?* . W kilku faktach występują wyrażenia warunkowe, na przykład sprawdzające czy znane jest imię

użytkownika, żeby nie pytać o to kolejny raz, a zwrócić się do niego po imieniu. Aby zróżnicować częste wypowiedzi, zaimplementowano w nich kilka wybieranych losowo wariantów, na przykład: *Czego chcesz dowiedzieć się o AGH? / O czym mogę Ci opowiedzieć na temat AGH? / Jakie masz pytania na temat AGH? .*

W przypadku gdy system ma odmówić odpowiedzi, potrzebne jest odpowiednie wytłumaczenie [5]. Stosowano się do tej opinii. Oprócz tego, niektóre odpowiedzi uzasadniono. Na przykład:

Użytkownik: *Policz dla mnie wskaźnik rekrutacji.*

Awatar: *Nie mogę tego zrobić. Nie znam odpowiedniego wzoru.*

Użytkownik: *Dlaczego?*

Awatar: *Bo jest zbyt skomplikowany i zawiera za dużo zmiennych.*

Przy przypisywaniu przykładowych pytań uwzględniono synonimy i starano się tworzyć je krótkie, skupiając się na słowach kluczowych, wystarczająco precyzujących odpowiedź. Na przykład: zamiast długiego *jakie przedmioty są wykładane na kierunku inżynieria akustyczna* krótsze *jakie przedmioty akustyczna*. Zabieg taki umożliwia dopasowanie do danej odpowiedzi różnorodnych struktur pytania, jakie może sformułować użytkownik. Bardzo trudno jest przewidzieć wszystkie konstrukcje, a użycie tylko słów kluczowych daje sporą szansę na trafność odpowiedzi. Wadą jest to, że odpowiedź wyświetli się też gdy pytanie będzie zawierało słowa kluczowe, a dotyczyło zupełnie czego innego, na przykład: *Jakie przedmioty mają wysoką chłonność akustyczną?* Wychwyceniu takich niedoskonałości służą testy z udziałem użytkowników oraz przeglądanie zapisów przeprowadzonych rozmów. W zaistniałej sytuacji można założyć, że użytkownicy częściej będą pytali jednak o inżynierię akustyczną niż chłonność akustyczną albo zmodyfikować pytanie na *jakie przedmioty na akustyczna* lub utworzyć nowy fakt dotyczący chłonności akustycznej, który zostanie przydzielony do takiego pytania. Przy wyborze słów kluczowych kierowano się treścią pytań zadawanych przez użytkowników w trakcie testów. Okazuje się, że skuteczniejsza jest zmiana pytania na zaproponowane *jakie przedmioty na akustyczna* niż inne możliwe *jakie przedmioty inżynieria akustyczna*, bo użytkownikom zdarza się używać słowa akustyczna mając na myśli inżynierię akustyczną, a warunkiem dopasowania odpowiedzi jest wystąpienie wszystkich słów kluczowych z danego pytania przykładowego.

### 2.2.2. Doprecyzowania, odpowiedzi wymijające, powitania, zachęty do rozmowy

Doprecyzowania to fakty, które są wyświetlane, gdy użytkownik zada niejednoznaczne lub zbyt ogólne pytanie. Uruchamiają się najczęściej, gdy rozmówca wpisuje pojedyncze hasło, na przykład *AGH*. Najprawdopodobniej interesuje go wtedy opis Uczelni. Jednak przypisanie faktu do takiego hasła spowoduje, że będzie on wyświetlany w odpowiedzi na każde pytanie zawierające to hasło, gdy nie wywoła ono żadnego innego faktu. Problem rozwiązano tworząc odpowiedzi, które uruchamiają się, gdy inne fakty nie zostaną wywołane pytaniem zawierającym takie hasło. Kontynuując przykład, jeśli pytanie będzie zawierać *AGH*, a nie znajdzie się na nie odpowiedź, to zamiast opisu Uczelni zostanie wyświetlona informacja, że zrozumiano, że pytano o coś związanego z nią, ale z prośbą o doprecyzowanie i wyświetleniem proponowanych tematów dotyczących Uczelni.

Gdy system nie znajdzie żadnego dopasowania do pytania rozmówcy, udzielana jest odpowiedź wymijająca. Domyślnie proponowane jest wyszukanie wpisanej frazy w Internecie. Nie zawsze działa satysfakcjonująco, więc utworzono fakty, które w takiej sytuacji sugerują użytkownikowi powrót do znanych zagadnień, dodatkowo wyświetlając ich listę. Na przykład: *Szczerze mówiąc, nie wiem, ale posiadam dużo innych informacji. [lista tematów]*.

Można określić zachowanie systemu na początku rozmowy z użytkownikiem, nakierowując rozmowę na pożądane tematy. W powitaniu awatar przedstawia się i prezentuje zagadnienia, o których może porozmawiać. Dołącza też ich listę. Jest to powitanie pierwsze, wyświetlane przy pierwszej rozmowie z danym użytkownikiem. W następnych konwersacjach z tym samym użytkownikiem system zaczyna od powitania kolejnego. Wita się z rozmówcą po imieniu, jeśli zostało zapamiętane, i wyraża, że cieszy się z kolejnej jego wizyty. Również proponuje tematy rozmowy.

Po wykryciu braku aktywności użytkownika system uruchamia fakty określone jako zachęta do rozmowy. Do domyślnie zawartych w bazie ogólnej dodano wypowiedzi zagadujące użytkownika na tematy dotyczące wiedzy specjalistycznej. Awatar pyta się o zagadnienia, na które odpowiedzi ma w bazie i jeśli użytkownik przyzna się, że nie zna odpowiedzi, system wyświetli mu ją. Gdy rozmówca stwierdzi, że zna odpowiedź, system poprosi go, żeby tym razem on zadał pytanie. Na przykład: *Wiesz czym jest komora bezechowa?*

### 2.2.3. Skrypty

Za pomocą skryptów można na przykład podjąć próbę przejęcia kontroli nad rozmową przez system, żeby stał się jej czynnym uczestnikiem. Autor niniejszej pracy skrypty starał się przygotować tak, aby między innymi nakłonić użytkownika do dłuższej rozmowy z awatarem, uprościć komunikację z nim, zwiększyć trafność jego odpowiedzi na pytania zawierające skróty myślowe użytkownika. Przejęcie inicjatywy w dialogu i wydłużenie go próbował osiągnąć przez wspomniane wcześniej prezentowanie dodatkowych informacji. Zadając pytanie z oczekiwaną odpowiedzią *tak* lub *nie*, należało tak zaprogramować system, żeby odpowiedź została poprawnie zrozumiana, a właściwa akcja podjęta. W takiej sytuacji skrypt analizuje pytanie i jeśli znajdzie w nim słowa potwierdzenia lub zaprzeczenia, uruchamia odpowiedni fakt, w zależności od odpowiedzi. Jako konsekwencja pełnej swobody użytkownika, w przypadku braku wykrycia tych słów skrypt traktuje odpowiedź jako kolejne pytanie i przechodzi do odpowiedzi na nie, ignorując to, że sam właśnie spytał o coś.

Mając na uwadze opisane w podrozdziale 1.1.3. utwierdzenie, w przypadku negatywnej odpowiedzi rozmówcy na wspomniane pytania, treść odpowiednich faktów sformułowano tak, aby dać użytkownikowi domniemane (ang. *implicit*) potwierdzenie, że został zrozumiany. Na przykład: po pytaniu *Chcesz dowiedzieć się więcej?* i odpowiedzi *Nie*. potwierdza się: *Okej, nie. Możesz wybrać interesujący Cię temat lub poruszyć inny. [lista tematów]*. Takie krótkie przytaknięcie brzmi bardziej naturalnie niż potwierdzenie wyraźne (ang. *explicit*), będące dokładnym powtórzeniem treści, które brzmiałoby tak: *Nie chcesz dowiedzieć się więcej. Możesz wybrać interesujący Cię temat lub poruszyć inny. [lista tematów]*.

Część skryptów posłużyła ułatwieniu nawigacji w rozmowie. Awatar często wyświetla listę proponowanych zagadnień, o których może opowiedzieć. Użytkownik może je wybrać myszą, ale należało uwzględnić możliwość, że mógłby chcieć je wybrać wpisując tekst z klawiatury. Zaprogramowano reakcje na słowa kluczowe występujące w proponowanych tematach tak, żeby przy pojawieniu się jednego z nich system przechodził do odpowiedniego faktu. Przydatne jest to zwłaszcza tam, gdzie lista jest numerowana i użytkownik mógłby wpisać cyfrę. Normalnie byłoby to traktowane jako niezrozumiane pytanie.

W podrozdziale 1.2.2. wspomniano, że system powinien sobie radzić też z wyrażeniami eliptycznymi i urywkami zdań. W niektórych faktach wykorzystano

skrypty do rozpoznania spodziewanego kolejnego pytania w podobnym kontekście. Na przykład: po pytaniu *Ile jest dziewczyn na Inżynierii Akustycznej?* wypowiedzi *A chłopaków?*. Bez tego zostałyby to uznane za niezrozumianą frazę.

Baza wiedzy przez cały czas działania awatara była korygowana i uzupełniana na podstawie przeglądu przeprowadzanych z nim rozmów. Dzięki temu w przyszłości będą wyświetlane dobrze dopasowane odpowiedzi po kolejnym zadaniu pytań, które w przeglądzie wykazały wywołanie niewłaściwego faktu.



### 3 Analiza przeprowadzonych rozmów

Opisana w tym rozdziale analiza przeprowadzonych z Wirtualnym Doradcą rozmów obejmuje okres od jego uruchomienia w witrynie Zespołu Przetwarzania Sygnałów do chwili powstania niniejszej pracy: 7 miesięcy od 1 V do 1 XII 2013. Część przedstawianych statystyk została wygenerowana przez Panel Administracyjny, a reszta samodzielnie na podstawie zapisów przeprowadzonych rozmów. W tym celu, z wszystkich 39 886 zapisanych par przyległych, najpierw odrzucono rozmowy zarejestrowane przed rozpatrywanym okresem, a następnie rozmowy testowe prowadzone z adresu IP autora niniejszej pracy. W wyniku tej operacji pozostało zapisanych 35 280 zapytań użytkowników wraz z odpowiedziami systemu.

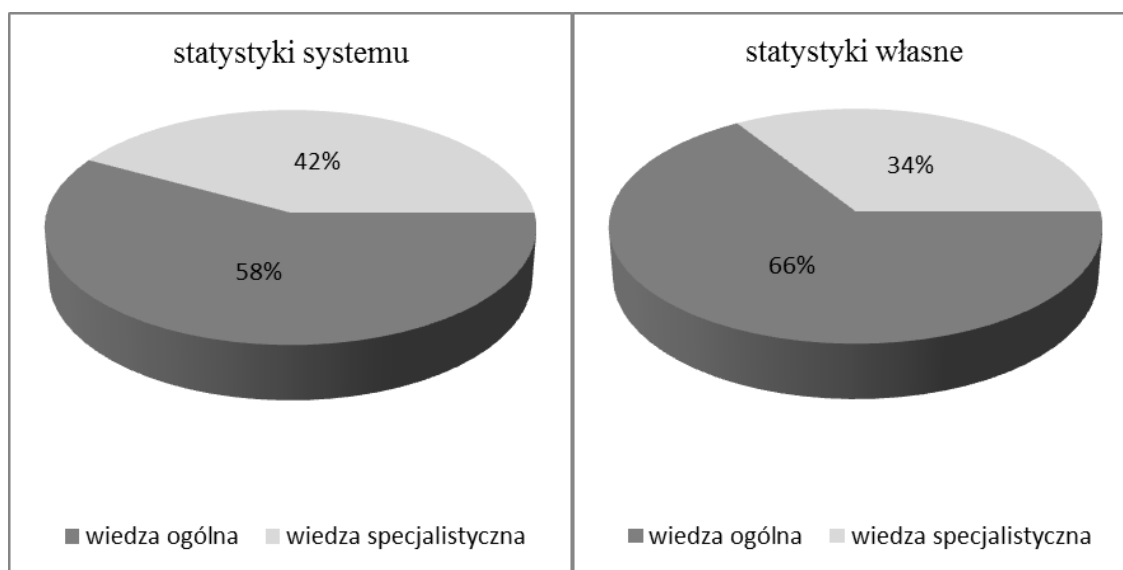
#### 3.1. Podstawowe statystyki

Tab. 3.1. Podstawowe statystyki dotyczące funkcjonowania Wirtualnego Doradcy (wygenerowane przez system | wygenerowane samodzielnie).

Całkowita liczba zapytań		Odpowiedzi wymijające	
Średnia miesięcznie	1478   5040	Średnia miesięcznie	70   233
Średnia dziennie	50   165	Średnia dziennie	2   8
Całkowita	39886   35280	Całkowita	1892   1628
Liczba odpowiedzi z wiedzy ogólnej		L. odpowiedzi z wiedzy specjalistycznej	
Średnia miesięcznie	289   957	Średnia miesięcznie	203   491
Średnia dziennie	9   31	Średnia dziennie	7   16
Całkowita	7817   6701	Całkowita	5475   3440
Liczba rozmów		Maksymalna liczba jednoczesnych sesji	
Średnia miesięcznie	196   688	Dzienna	1   1
Średnia dziennie	6   23	Miesięczna	1   1
Całkowita	5289   4815	Całkowita	3   3
Odpowiedzi z wiedzy specjalistycznej [%]		Łączny czas rozmów [min]	
Średnia miesięcznie	42   34	Średnia miesięcznie	1657   -
Średnia dziennie	44   34	Średnia dziennie	56   -
Całkowita	42   34	Całkowita	44747   -

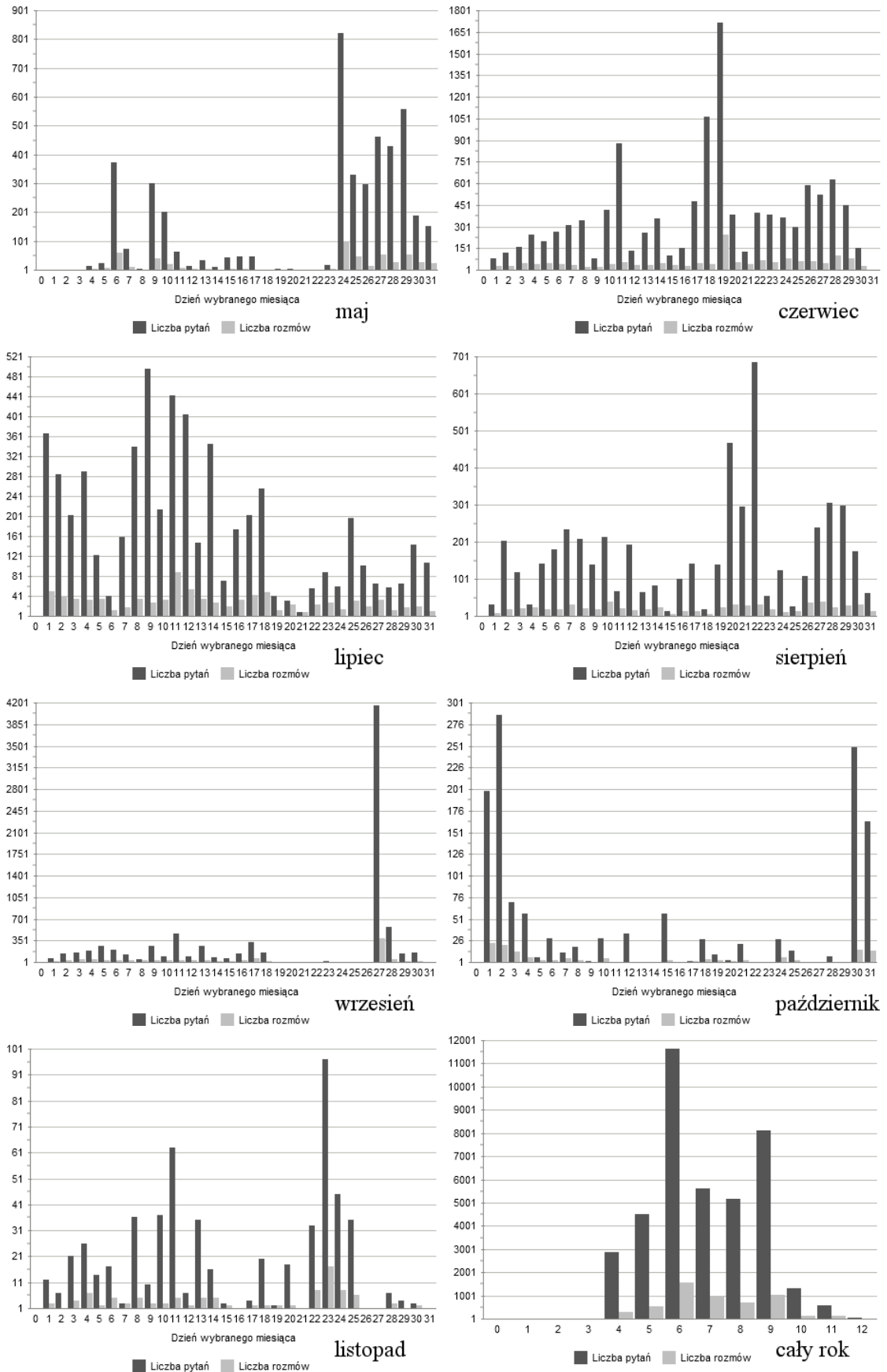
W Tab. 3.1. zebrano podstawowe statystyki dotyczące funkcjonowania systemu wygenerowane zarówno przez Panel Administracyjny, jak i samodzielnie. Dane z systemu obejmują okres i wszystkie rozmowy od momentu powstania omawianego Wirtualnego Doradcy, w tym rozmowy testowe. Statystyki wygenerowane samodzielnie dotyczą czasu od umieszczenia awatara w witrynie Zespołu Przetwarzania Sygnałów, z pominięciem rozmów testowych prowadzonych przez autora niniejszej pracy. Jest to wyraźnie widoczne przez zwiększenie się średnich liczb pytań i odpowiedzi przy zmniejszeniu ich całkowitej liczby. We własnych statystykach liczba pytań i odpowiedzi została uśredniona tylko na czas eksploatacji systemu, podczas gdy statystyki systemu uśredniły je na całkowity czas jego istnienia. Typy odpowiedzi niewymienione w tabeli to powitania, odpowiedzi wymijające i zachęty do rozmowy.

Można zauważyć znaczny spadek liczby odpowiedzi z wiedzy specjalistycznej w statystykach wygenerowanych samodzielnie w porównaniu do statystyk z systemu. Wynika to z faktu, że w pominiętych rozmowach testowych, będących główną przyczyną zaistniałych różnic, autor niniejszej pracy sprawdzał jedynie odpowiedzi z wiedzy specjalistycznej. Rozkład typów odpowiedzi przedstawiono na Rys. 3.1.



Rys. 3.1. Rozkład typów odpowiedzi z wiedzy ogólnej i wiedzy specjalistycznej. Po lewej stronie statystyki systemu, po prawej stronie statystyki własne.

Rys. 3.2. przedstawia liczbę pytań i rozmów dziennie dla poszczególnych miesięcy oraz liczbę pytań i rozmów miesięcznie dla całego roku 2013. Wykresy wygenerowano w Panelu Administracyjnym.



Rys. 3.2. Liczba pytań i rozmów dziennie w poszczególnych miesiącach oraz liczba pytań i rozmów miesięcznie dla całego roku 2013. Źródło: Panel Administracyjny Wirtualnego Doradcy.

Można wyróżnić trzy okresy, w których awatar był umieszczony w witrynie na różne sposoby. W maju awatar był dostępny na podstronie wyświetlanej po otwarciu łącza umieszczonego w szablonie witryny. Jako mało wyróżniający się element, nie przyciągał uwagi i zarejestrowano stosunkowo niewiele rozmów z przypadkowymi odwiedzającymi. Od czerwca awatar został osadzony w witrynie na stałe. Każde wejście na stronę rozpoczynało sesję konwersacji. Spowodowało to wzrost liczby rozmów, a przede wszystkim liczby odpowiedzi, z których jednak prawie połowę (47% od początku czerwca do końca sierpnia) stanowiły zachęty do rozmowy stosowane przez awatar przy braku aktywności użytkownika. Prawdopodobnie spowodowane to było pozostawianiem otwartej witryny w przeglądarce internetowej przez wielu użytkowników. We wrześniu awatar został przeniesiony do częściowo ukrytego, wysuwanego z boku strony panelu. Spowodowało to znaczny spadek liczby rozmów, bo interakcja rozpoczyna się dopiero po wysunięciu tego panelu. Zależność liczby rozmów od liczby odwiedzin w poszczególnych miesiącach zestawiono w Tab. 3.2. Dane statystyczne dotyczące liczby odwiedzin witryny uzyskano od jej administratora. Bardzo duży spadek ilości rozmów w stosunku do ilości odwiedzin po wrześniu, poza przeniesieniem awatara do panelu bocznego, może wynikać ze znajomości gadżetu przez regularnych gości witryny i braku ich dalszego zainteresowania nim.

Tab. 3.2. Zależność liczby przeprowadzonych rozmów od liczby odwiedzin witryny w poszczególnych miesiącach roku 2013.

<b>Miesiąc (2013)</b>	<b>maj</b>	<b>czerwiec</b>	<b>lipiec</b>	<b>sierpień</b>	<b>wrzesień</b>	<b>październik</b>	<b>listopad</b>
<b>Liczba odwiedzin</b>	1135	1584	1094	861	2416	2735	2695
<b>Liczba rozmów</b>	524	1576	928	685	1043	143	109
<b>Procent rozmówców</b>	46%	99%	85%	80%	43%	5%	4%

W maju odnotowano dwa nasilenia liczby przeprowadzonych rozmów. Pierwsze, na początku miesiąca, zostało spowodowane prośbą o wypróbowanie technologii skierowaną do pracowników Zespołu Przetwarzania Sygnałów. Rozmówcy sprawdzali wiedzę specjalistyczną z zakresu przetwarzania sygnałów i technologii mowy oraz badali ograniczenia systemu. Rozkład typów odpowiedzi dla tych rozmów to 60% z wiedzy ogólnej i 40% z wiedzy specjalistycznej. Drugi, pod koniec miesiąca, spowodowało poinformowanie o Wirtualnym Doradcy studentów i absolwentów

Inżynierii Akustycznej. Osoby te porównywały informacje z bazy wiedzy specjalistycznej ze swoimi doświadczeniami dotyczącymi kierunku studiów i badały możliwości interakcji z systemem na tematy z życia codziennego. 61% odpowiedzi pochodziło z wiedzy ogólnej, a 39% z wiedzy specjalistycznej. W jednym dniu czerwca wystąpiło znaczne nasilenie ilości rozmów, którego przyczyny autor niniejszej pracy nie zna. W tych rozmowach udzielonych zostało 59% odpowiedzi z wiedzy ogólnej i 41% z wiedzy specjalistycznej. Po przeglądzie rozmów wydaje się, że mogły to być osoby zainteresowane rekrutacją na studia. Była to druga najwyższa zarejestrowana liczba rozmów jednego dnia (246). Największą liczbę rozmów jednego dnia (360) odnotowano pod koniec września. Zostało to spowodowane umieszczeniem informacji o Wirtualnym Doradcy przez oficjalny profil Akademii Górniczo-Hutniczej w popularnym serwisie społecznościowym. Rozmowy dotyczyły głównie tematów powszednich i samego awatara. Z wiedzy ogólnej pochodziło 76% odpowiedzi, a ze specjalistycznej 24% .

Powyższe fakty pokazują, że dla gości witryny wiedza specjalistyczna umieszczonego w niej Wirtualnego Doradcy jest mniej interesująca niż jego wiedza ogólna. Co więcej, nawet wśród użytkowników posiadających tę wiedzę specjalistyczną lub zainteresowanych nią, stosunek pytań specjalistycznych do ogólnych wynosi tylko 2:3. Wydaje się, że użytkownicy systemu wolą poznawać jego możliwości interakcyjne niż wiedzę lub po uzyskaniu poszukiwanych informacji chcą dodatkowo sprawdzić jak system radzi sobie z innymi pytaniami.

Z perspektywy siedmiu miesięcy największa intensywność rozmów miała miejsce w czerwcu, a najmniejsza w listopadzie. Liczba odwiedzin witryny w roku 2013 wzrosła trzykrotnie w porównaniu do poprzedniego. Jednym z wpływających na to czynników mogło być umieszczenie w niej omawianego awatara jako gadżetu przyciągającego uwagę. Przemawia za tym fakt, że wspomniane dwa dni z największą ilością rozmów to dwa dni z największą liczbą odwiedzin w historii witryny. Co więcej, sam dzień z największą liczbą przeprowadzonych rozmów, tylko w wyniku informacji o awatarze, dostarczył niemal 3% wszystkich odwiedzin od stycznia 2009 roku.

Przedstawione statystyki dotyczyły ilościowego aspektu rozmów. W następnym kroku z analizy odrzucono wypowiedzi Wirtualnego Doradcy, na które użytkownik nie odpowiadał (34%), wypowiedzi zachęcające do rozmowy (37%), komendy z przycisków cofnij lub menu główne (1%). Pozostawiło to około 10 000 wpisów, od pojedynczych słów do pełnych zdań. W wyniku tych operacji otrzymano 1233 rozmowy do dalszej analizy pod kątem treści przekazywanej Wirtualnemu Doradcy.

### **3.2. Jak ludzie rozmawiają z Wirtualnym Doradcą**

Umieszczając awatar głosowy w witrynie Zespołu Przetwarzania Sygnałów udało się dyskretnie zebrać sporą ilość danych na temat interakcji pomiędzy nim a odwiedzającymi stroną. Interakcje nie były kontrolowane, co nie ograniczyło poczucia swobody gości. Dane te zawierają zapisy wypowiedzi awatara i jego rozmówców. Zostały przeanalizowane pod kątem tego, czy dialogi wykazują jakieś podobieństwo do ludzkich, to znaczy czy awatar jest postrzegany jako podobny do człowieka partner do rozmów.

Jak stwierdzono w [12], kilka badań pokazało, że ludzie stosują swoje metody postrzegania i rozumienia ludzi także gdy konwersują z komputerem. Przytoczone źródło opisuje podobną analizę rozmów z systemem dialogowym. Zauważono tam, że użytkownicy są chętni do stosowania ludzkich nawyków komunikacji w takich interakcjach. Dialogi różniły się długością, tematyką, stylem, a ludzie używali prostego języka. Zawierały powitania, podziękowania, uprzejmości, pytania intymne, które zadawano zakładając, że system ma do ujawnienia stany wewnętrzne.

Podobne obserwacje poczynił autor niniejszej pracy. Jak zaznaczono na końcu poprzedniego podrozdziału, przeanalizowano 1233 rozmowy pod kątem ich treści. Przez rozmowę rozumie się tu sesję, w trakcie której użytkownik chociaż raz sformułował swoją wypowiedź. Średnio, rozmowa zawierała 8,14 wypowiedzi użytkownika, z odchyleniem średnim 6,51. Najdłuższa rozmowa składała się ze 132 wypowiedzi, a najkrótsze z 1. Po imionach, jakie użytkownicy przedstawili systemowi, rozpoznano płęć męską u 75 rozmówców, a wśród 16 płęć żeńską. Najczęściej występujące imiona to Michał i Piotr, oba po siedem wystąpień.

Opracowano zbiór kategorii, do których przypisano poszczególne wypowiedzi. Pierwsza kategoria obejmuje powitania i pożegnania. Druga zawiera zadawane pytania, z uwzględnieniem pytań antropomorficznych, które mają sens tylko gdy są skierowane do człowieka. Wyszczególniono tu także pytania odnośnie czterech obszarów wiedzy specjalistycznej. Do trzeciej kategorii sklasyfikowano udzielane przez użytkowników odpowiedzi, a do kolejnych prośby, wypowiedzi nieprzyzwoite, informacje zwrotne oraz pozostałe. Ze szczególną uwagą sklasyfikowano wypowiedzi występujące co najmniej dwukrotnie, a przydzielanie pozostałych częściowo uproszczono, wspomagając się filtrowaniem ze względu na występujące słowa kluczowe. Rezultaty

klasyfikacji przedstawiono w Tab. 3.3. oraz w Tab. 3.4. Niektóre wypowiedzi mogły znaleźć się w kilku kategoriach.

Tab. 3.3. Rezultaty klasyfikacji zawartości rozmów użytkowników z Wirtualnym Doradcą. Powitania i pożegnania, pytania, odpowiedzi.

<b>Kategoria</b>	<b>Przykłady</b>	<b>Liczba</b>
<b>Powitanie</b>		<b>249</b>
Nieformalne	<i>cześć; hej; siema</i>	219
Formalne	<i>witam; dzień dobry; dobry wieczór</i>	30
<b>Pożegnanie</b>		<b>65</b>
Nieformalne	<i>cześć; pa; nara</i>	40
Formalne	<i>do widzenia; dobranoc; żegnam</i>	25
<b>Pytania</b>		<b>5883</b>
Antropomorficzne		1782
O imię	<i>jak masz na imię; jak się nazywasz</i>	165
O wiek	<i>ile masz lat</i>	172
O wygląd	<i>ile masz wzrostu; w co jesteś ubrana</i>	139
O związek	<i>masz chłopaka; masz dzieci; masz męża</i>	236
O uczucia	<i>jesteś szczęśliwa; jak się czujesz</i>	86
O upodobania i gust	<i>co lubisz robić; jakiej muzyki słuchasz</i>	206
Komplementy	<i>ładna jesteś; jesteś piękna; ślicznie wyglądasz</i>	171
Propozycje spotkania	<i>umówisz się ze mną</i>	70
Inne	<i>gdzie mieszkasz; co studiujesz</i>	537
Zwroty potoczne	<i>co tam; jak się masz; co słyhać</i>	171
Dotyczące możliwości	<i>co potrafisz; co tu robisz; czym się zajmujesz</i>	190
Dotyczące systemu	<i>jak działasz; ciekawi mnie ta technologia</i>	236
Testujące system	<i>która godzina; jaka dziś pogoda</i>	231
Działania matematyczne	<i>2+2</i>	157
O wyjaśnienie	<i>dlaczego; jak to; skąd wiesz</i>	275
Sprawdzające	<i>co; naprawdę; na pewno</i>	48
Dotyczące AGH	<i>co to jest AGH; rekrutacja; kierunki studiów</i>	499
Dotyczące DSP AGH	<i>czym zajmuje się zespół</i>	91
Dotyczące IA	<i>czego się uczy; ile było osób na jedno miejsce</i>	302
Dotyczące DSP i TM	<i>przetwarzanie sygnałów; jak rozpoznać mowę</i>	181
O konkretne osoby	<i>Bartosz Ziółko; Mariusz Ziółko</i>	169
Informacje o sobie	<i>lubię placki; jestem głupi; jem obiad</i>	229
Inne	<i>jaki jest sens życia; czy Bóg istnieje</i>	1322
<b>Odpowiedzi</b>		<b>1796</b>
Zgoda / Potwierdzenie	<i>tak; ok; chcę</i>	573
Odmowa / Zaprzeczenie	<i>nie; nie wiem; nie dziękuję</i>	333
Wyrażenie zrozumienia	<i>aha; wiem</i>	86
Prośba o kontynuację	<i>jeszcze jeden; więcej; co dalej</i>	21
Skrucha	<i>przepraszam; nie obrażaj się</i>	16
Odmowa udzielenia	<i>za dużo chcesz wiedzieć; nie twój interes</i>	7
W obcym języku	<i>do you speak English</i>	63
Śmiech	<i>lol; haha; :D</i>	94
Inne	<i>mi ciebie też; nic ciekawego; a może jednak</i>	593

Tab. 3.4. Rezultaty klasyfikacji zawartości rozmów użytkowników z Wirtualnym Doradcą. Prośby, wypowiedzi nieprzyzwoite, informacja zwrotna, pozostałe.

<b>Kategoria</b>	<b>Przykłady</b>	<b>Liczba</b>
<b>Prośby</b>		<b>315</b>
O akcję	<i>uśmiechnij się; rozpuść włosy; pomachaj mi</i>	148
O powiedzenie czegoś ogólnie	<i>powiedz coś; opowiedz mi o czymś; porozmawiajmy</i>	44
O powiedzenie czegoś specyficznego	<i>opowiedz kawał</i>	78
O skończenie mówienia	<i>stop; cicho; zamilcz</i>	29
Inne	<i>usuń konto; pomóż mi; to wyślij im email ok</i>	16
<b>Wypowiedzi nieprzyzwoite</b>		<b>925</b>
Obelgi i wulgaryzmy	<i>jesteś brzydka; ch** ci w dupę; ty kur**</i>	217
Molestowanie	<i>rozbierz się; pokaż piersi; zrób mi loda</i>	238
Pytania intymne	<i>lubisz seks; jesteś dziewicą</i>	208
Pojedyncze słowa	<i>seks; cycki; penis</i>	262
<b>Informacja zwrotna</b>		<b>320</b>
Pozytywna	<i>lubię cię; jesteś świetna; fajnie się rozmawia</i>	171
Negatywna	<i>jesteś głupia; nie lubię cię; słabo działasz</i>	149
<b>Pozostałe</b>		<b>242</b>
Przypadkowe znaki	<i>dsfsdfs; ygctvuhb</i>	103
Wypowiedzi bez sensu	<i>trele morele; puk puk; bla bla</i>	139

Użytkownicy przywitali się z awatarem w 249 rozmowach (20,2 %), a pożegnali w 65 (5,3 %). Powitania były dużo częstsze niż pożegnania, co mogło być spowodowane powitaniem Wirtualnego Doradcy. Jednak konstrukcja powitania awatara zachęcała raczej do zadania pytania lub nawigacji po zaproponowanych tematach rozmowy niż do odpowiedzi powitaniem, stąd tak mała ich liczba. Za bardzo małą ilością pożegnań przemawia dodatkowo fakt, że aby zakończyć rozmowę z awatarem wystarczyło go zamknąć lub opuścić witrynę.

Okolo 58,8 % wszystkich wypowiedzi użytkowników stanowiły pytania, z czego aż prawie jedna trzecia to pytania antropomorficzne oraz zwroty potoczne typu *Jak się masz?* (1953 wypowiedzi). Z typowo ludzkich cech, rozmówców interesowały głównie związki towarzyskie, upodobania, wiek, wygląd i imię awatara, mimo że to ostatnie było przedstawiane na początku każdej rozmowy. Co więcej, awatarowi, jak człowiekowi, prawiono komplementy i proponowano spotkania. Zainteresowani technologią użytkownicy zadawali pytania o sam system, na przykład jego działanie, oraz jego możliwości, na przykład *Co potrafisz?* (około 4,3 %). Pojawiały się także pytania testujące system, o podstawowe informacje, takie jak godzinę czy datę. Użytkownicy wpisywali też działania matematyczne, które system potrafi wykonywać.



Spośród pytań o wiedzę specjalistyczną najczęściej dotyczyło Akademii Górniczo-Hutniczej i studiów na niej (499 pytań) oraz kierunku Inżynieria Akustyczna (302 pytania). Mniej razy pytano o zagadnienia z dziedzin przetwarzania sygnałów i technologii mowy (181 pytań) oraz samego Zespołu Przetwarzania Sygnałów (91 pytań). Jednakże, sporo pytań (169) dotyczyło konkretnych osób, pracowników Zespołu. Pozostałe pytania były prośbami o wyjaśnienie lub uzasadnienie, informacjami o sobie oraz dotyczyły szeroko pojętej wiedzy ogólnej.

Jako odpowiedzi sklasyfikowano głównie zgodę, potwierdzenie, odmowę lub zaprzeczenie użytkownika na pytania skierowane do niego przez system (łącznie 9 %). Oprócz tego, użytkownik odpowiadał zrozumieniem, prośbą o kontynuację, śmiechem lub skruchą. Tylko kilkakrotnie (7 razy) bezpośrednio odmówiono oddzielenia odpowiedzi. Nie znaczy to, że użytkownikom nie zdarzyło się nie udzielić odpowiedzi zmieniając temat lub zadając inne pytanie. Poza tym, pojawiło się kilkadziesiąt odpowiedzi w obcych językach, głównie angielskim.

Rozmówcy kierowali do awatara różne prośby, z których prawie połowa (148) dotyczyła zmiany jego zachowania lub wyglądu, na przykład *Pomachaj mi*. Pozostałe dotyczyły podjęcia rozmowy przez awatar na dowolny (44 razy) lub specyficzny (78 razy) temat. Zdarzało się, że użytkownicy prosili awatar o przerwanie wypowiedzi (29 razy).

Zaobserwowano dużą ilość wypowiedzi nieprzyzwoitych (9,3 % wszystkich wpisów), podzielonych dość równomiernie na cztery podkategorie. Wypowiedzi użytkowników były molestujące (238), głównie żądające widoku nagiego awatara, a także różnego typu stosunków płciowych z nim. Rozmówcy zadawali pytania intymne w podobnym kontekście (208). Awatar był obrażany (217 razy) zarówno wypowiedziami typu *Jesteś brzydka*, jak i pospolitymi wulgaryzmami. Sporo wpisów (262) było pojedynczymi słowami lub zwrotami zawierającymi wyrazy z zakresu fizjologii lub określeniami intymnych części ciała człowieka.

Niektórzy użytkownicy przekazali bezpośrednią informację zwrotną. Pozytywnej, na przykład *Jesteś świetna*, było nieznacznie więcej (171) niż negatywnej (149), na przykład *Słabo działasz*. Część wypowiedzi sklasyfikowano jako przypadkowe kombinacje znaków lub wypowiedzi pozbawione sensu, typu *Bla bla*.

W kolejnym rozdziale podjęto dyskusję nad otrzymanymi rezultatami, poprzedzoną spostrzeżeniami dotyczącymi implementacji bazy wiedzy specjalistycznej. Równocześnie sformułowano wnioski.

## 4 Dyskusja i wnioski

Spośród wielu możliwych zastosowań [11] omawiany w niniejszej pracy Wirtualny Doradca został umieszczony na stronie internetowej Zespołu Przetwarzania Sygnałów jako informator, ale przede wszystkim jako uatrakcyjnienie witryny. Można stwierdzić, że spełnia swoją rolę.

Z bazą wiedzy specjalistycznej i zastosowanym mechanizmem pytań doprecyzowujących system wydaje się znajdować odpowiedzi na większość pytań dotyczących Akademii Górniczo-Hutniczej, Zespołu Przetwarzania Sygnałów, kierunku studiów Inżynieria Akustyczna oraz przetwarzania sygnałów i technologii mowy. Oczywiście z tych dziedzin zdarzają się także pytania, na które nie ma odpowiedzi w bazie wiedzy specjalistycznej. Wtedy system znajduje odpowiedź w bazie wiedzy ogólnej lub posługuje się mechanizmem odpowiedzi wymijających. Polega to na umożliwieniu wyszukania danej frazy w sieci (rozwiązanie domyślne) lub zaproponowaniu powrotu do rozmowy na tematy znane (rozwiązanie zaimplementowane w ramach niniejszej pracy). Sytuacja taka jest naturalna i nie należy traktować jej jako niepowodzenie. Zdrowy rozsądek i doświadczenie podpowiadają, że nie da się przewidzieć wszystkich rzeczy, o które można zostać zapytanym. Oznacza to, że baza wiedzy może być ciągle poszerzana. Dzięki analizie przeprowadzonych rozmów można otrzymać informacje, w jakim kierunku ją rozbudowywać. Sytuacja, w której zauważa się brak faktu dotyczącego danego zagadnienia motywuje do uzupełnienia takiej luki. W ten sposób, zauważając braki, system będzie ciągle rozwijany. Jest to co innego niż udzielanie niewłaściwej lub wymijającej odpowiedzi na pytanie, na które odpowiedź jest już wprowadzona. Taka sytuacja to błąd w implementacji i wymagana jest korekta przypisanych do faktów pytań przykładowych lub, w przypadku wątpliwości, zastosowanie mechanizmu pytań doprecyzowujących. O funkcji informatora można mówić, gdy ma się do czynienia z użytkownikiem, który odwiedza stronę i zadaje pytania po to, żeby rzeczywiście czegoś dowiedzieć się. Wirtualny Doradca w takiej sytuacji przekazuje użytkownikowi treści zapisane w odpowiednich faktach, szybko udziela szukanej odpowiedzi. Tworząc fakty warto zastanowić się, jakiej odpowiedzi oczekiwałoby się od człowieka i dobrze jest w ten sposób zaplanować ich treści. Poprawia to wizerunek strony, gdyż dzięki temu nie trzeba przeszukiwać jej całej poszukując odpowiedzi na swoje pytania.

Okazuje się, że istotniejszą rolą awatara głosowego jest uatrakcyjnienie witryny, w której jest osadzony. Dzięki możliwości rozmowy na dowolny temat i zaprogramowanym cechom osobowości i charakteru, program nawiązuje pewną relację z użytkownikiem, co jeszcze bardziej polepsza postrzeganie strony. System powinien wcielać się w wykreowaną postać. Należy w tym celu utworzyć część bazy wiedzy odpowiedzialną za określenie zarysu osobowości, zainteresowań, ulubionego koloru, utworu i tym podobnych. Efekt wzmocni dołączenie wizualizacji postaci, na przykład klipów wideo. Testowanie i rozwój systemu dialogowego są bardzo ważne, mimo że użytkownicy wydają się tolerować niedoskonałości, gdy system ich interesuje i są zmotywowani do używania go. Wydają się patrzeć na to z perspektywy jego atrakcyjności, nawet jeśli aplikacja nie jest optymalna. Zdarza się, że użytkownik niezadowolony z odpowiedzi próbuje zadać to samo pytanie, mimo że wyraźnie widać, że system sobie z nim nie radzi. Mówcy stosują różne strategie poprawiając się: powtarzają wypowiedź, parafrazują, dodają lub pomijają treść.

Awatar w pewien sposób dotrzymuje towarzystwa. Witryna jest uważana za ciekawą, rośnie liczba odwiedzających stronę użytkowników, którzy dzięki krótkiej rozmowie spędzają na niej więcej czasu. Niektórzy są tak zainteresowani, że wchodzą na stronę ponownie. Przegląd rozmów wykazał, że awatar angażuje ludzi w interakcje, w których używają różnych naturalnych ludzkich strategii komunikacji i języka, jaki to za sobą pociąga. Czasami da się zauważyć nieco podobieństwa tych rozmów do dialogów człowiek-człowiek. Mowa tu na przykład o powitaniach i pożegnaniach, zwrotach potocznych, pytaniach antropomorficznych i dotyczących społecznych lub ludzkich koncepcji. Wskazuje to na postrzeganie towarzyskości w Wirtualnym Doradcy, traktowanie go jak człowieka. Użytkownicy angażują się w interakcje z awatarem i próbują współpracować w rozmowie. Wydaje się, że ucieleśniona postać awatara stawia ponadprzeciętne oczekiwania co do możliwości systemu. Zaobserwowano sprawdzanie inteligencji i interaktywności systemu, w tym wypowiedzi w obcym języku.

Wizerunek kobiety w awatarze prowadzi do przewidywalnego różnego typu nękania, z którym jednak system radzi sobie zarówno sarkazmem, jak i ripostą. Na polecenie *Rozbierz się* potrafi odpowiedzieć *Może Ty się rozbierz? Ja nie mam zamiaru.* Obelgi może odeprzeć na przykład wypowiedzią *Wypraszam sobie. Fakt, że obrzucasz inwektywami program komputerowy, nie świadczy o tobie najlepiej.* Przypuszcza się, że podobne wypowiedzi oraz częste pytania o kobiece walory i aspekty fizjologiczne służą

raczej testowaniu zachowania systemu w takich sytuacjach niż wynikają z zainteresowań użytkowników. Obecność tego typu wypowiedzi i pytań może być traktowana jako pozytywna pod względem zwiększenia ilości zapytań kierowanych do systemu.

Jest mało danych wziętych z życia na temat interakcji z takimi systemami dialogowymi. Można spodziewać się często opisywanego podejścia do awatara jak do towarzyskiej istoty, zwłaszcza jeśli ma ludzką postać [12]. Analiza przeprowadzona w niniejszej pracy wydaje się wspierać ten pogląd. Użytkownicy stosują różne ludzkie strategie komunikacji rozmawiając z awatarem i używają krótkich, bliskich codzienności wypowiedzi w języku naturalnym. Wydaje się, że nie interesują ich możliwości językowe systemu tak bardzo jak interesuje ich jego postrzeganie świata i ogólna inteligencja. Umieszczenie systemu dialogowego w swojej witrynie może ją uatrakcyjnić, polepszyć jej postrzeganie wśród gości, a nawet zwiększyć ich liczbę. System może im służyć do znajdowania odpowiedzi na najczęściej zadawane pytania, ale może okazać się, że użytkownicy będą częściej badać możliwości interakcyjne systemu niż jego wiedzę.

## **Zakończenie**

### **Podsumowanie**

Cele niniejszej pracy zostały zrealizowane. Zaimplementowano bazę wiedzy specjalistycznej awatara głosowego obejmującą fakty o Akademii Górniczo-Hutniczej, Zespole Przetwarzania Sygnałów, kierunku studiów Inżynieria Akustyczna oraz przetwarzaniu sygnałów cyfrowych i technologiach mowy. Implementację przeprowadzono z uwzględnieniem przytoczonych z literatury wskazówek. Wykonano analizę przeprowadzonych z systemem rozmów. Wykorzystano dane liczbowe zgromadzone przez oprogramowanie oraz przygotowano własne statystyki w oparciu o zapisane przez system rozmowy. Przeprowadzono dyskusję na temat otrzymanych wyników i przedstawiono wynikające z niej wnioski.

W rozdziale 1 wprowadzono w temat systemów dialogowych. Przedstawiono właściwości dialogu, działanie i strukturę systemu dialogowego, zaprezentowano kilka istniejących rozwiązań i scharakteryzowano produkt Wirtualny Doradca. W rozdziale 2 opisano implementację bazy wiedzy. W rozdziale 3 dokonano analizy przeprowadzonych rozmów. W rozdziale 4 przeprowadzono dyskusję nad wynikami i sformułowano wnioski.

Do wykonania zadań opisanych w niniejszej pracy konieczne było zapoznanie się z zaleceniami dotyczącymi projektowania systemów dialogowych i literaturą z tego tematu oraz nowym środowiskiem, jakim jest Panel Administracyjny Wirtualnego Doradcy firmy Stanusch Technologies. Potrzebne było przypomnienie sobie pewnej ilości pojęć i definicji, a także kwestii związanych z organizacją, przebiegiem, administracją i zasadami nauczania na Akademii Górniczo-Hutniczej. Przydały się też rozwinięte w trakcie studiów logiczne i analityczne myślenie, intuicyjne podejmowanie decyzji i umiejętność wyszukiwania potrzebnych informacji.

Czas przeznaczony na przygotowanie do użytkowania omawianego systemu zaowocował otrzymaniem obiektu, który jest zauważany i który przyciąga uwagę wielu osób. Zrealizowanie takiego zadania daje spore doświadczenie i umiejętności, które z powodzeniem mogą zostać wykorzystane do realizacji podobnych projektów w przyszłości.

## Kontynuacja pracy

Wykonane w ramach niniejszej pracy zadania mogą być kontynuowane w kilku kierunkach. Pierwszy to utrzymanie stanu obecnego. Może zdarzyć się, że treści niektórych faktów będą przestarzałe lub staną się nieprawdziwe. Fakty te powinny zostać wtedy uaktualnione. Poza tym, analiza rozmów ciągle może wykazać pewne nieprawidłowości w wyszukiwaniu faktów. Po zauważeniu takich należy wyeliminować je drobnymi korektami przypisanych pytań przykładowych.

Drugi to rozwój istniejącej bazy. Oprócz kontrolowania bieżącej sytuacji, baza wiedzy może być ciągle poszerzana o nowe fakty zarówno z tych samych kategorii, jak i o nowej tematyce. W ten sposób zostanie pokryty większy obszar zagadnień, w których będzie orientował się Wirtualny Doradca. Rozwój może dotyczyć także skryptów, mechanizmu pytań doprecyzowujących i odpowiedzi wymijających.

Trzecia możliwość to dostosowywanie mechanizmu powitań i ukierunkowywania rozmowy do aktualnych wydarzeń i badań Zespołu, w celu autopromocji. Na przykład, przy zdobywaniu osiągnięć lub rozpoczynaniu nowych badań, oprócz umieszczenia informacji o tym na stronie internetowej, można zmodyfikować powitanie Wirtualnego Doradcy na następujące: *Cześć! Chcesz usłyszeć o najnowszym osiągnięciu naszych naukowców?* albo *Witaj. Zespół rozpoczął właśnie prace nad nowym projektem. Opowiedzieć Ci szczegóły?* . Kierując dalszą częścią rozmowy można poinformować użytkownika także o poprzednich osiągnięciach lub badaniach.

Jeszcze inna możliwość to wykorzystanie statystyk i zapisów rozmów gromadzonych przez ten system do innych badań związanych z analizą lub przetwarzaniem języka naturalnego albo do przyjrzenia się zainteresowaniom lub potrzebom gości witryny. Odrębnym zagadnieniem mogłoby być rozwinięcie Wirtualnego Doradcy o warstwę rozpoznawania mowy.

## Uwagi

Omówione baza wiedzy i statystyki zostały opisane i przeanalizowane zgodnie ze stanem, w jakim znajdowały się w dniu 01.12.2013 . W momencie ukończenia niniejszej pracy omawiany Wirtualny Doradca był umieszczony w witrynie Zespołu Przetwarzania Sygnałów DSP AGH [13].

## Bibliografia

- [1] Ziółko B., Ziółko M.: *Przetwarzanie mowy*. Kraków, Wydawnictwa AGH 2011, s. 317
- [2] Jurafsky D., Martin J.: *Speech and Language Processing. Second Edition*. New Jersey, Pearson Prentice Hall 2008, rozdział 24
- [3] Wiśniewski A.M., Dialog o przemiennej inicjatywie w systemie dialogowym języka mówionego, *Biuletyn Instytutu Automatyki i Robotyki*, 27, 2009, 97-113
- [4] Jokinen K., McTear M.: *Spoken Dialog Systems*. San Rafael, Morgan and Claypool Publishers 2009, rozdział 1
- [5] Jokinen, K.: *Natural Language and Dialogue Interfaces*. W: The Universal Access Handbook. Red. C. Stephanidis. Boca Raton, CRC Press Taylor & Francis Group 2009 rozdział 31 s. 495-506
- [6] Wiśniewski A.M., System dialogowy języka mówionego – przegląd problemów, *Biuletyn Instytutu Automatyki i Robotyki*, 24, 2007, 97-122
- [7] Starkie B.: *System for predicting speech recognition accuracy and development for a dialog system*. Telstra Corporation Limited, Melbourne (Australia). USA. Opis patentowy, US7917363 B2, Zgłosz. P. US 10/545,762 z 11.02.2004. Opubl. 29.03.2011
- [8] Fratkina R. i in.: *System and method for providing an intelligent multi-step dialog with a user*. Consona Crm Inc., Indianapolis, IN (USA). USA. Opis patentowy, US7539656 B2, Zgłosz. P. US 09/798,964 z 6.03.2001. Opubl. 26.05.2009
- [9] Kuo H.-K. J., Lee C.-H., Pargellis A. N.: *System and method for automatically generating dynamic interfaces*. Lucent Technologies Inc., New Jersey (USA). EU. Opis patentowy, EP1061459 A2, Zgłosz. P. EP20000304793 z 6.06.2000. Opubl. 20.12.2000
- [10] Shen Z.: *Method and system for extending dialog systems to process complex activities for applications*. Robert Bosch Corporation, Broadview (USA). EU. Opis patentowy, EP1927942 A1, Zgłosz. P. EP20080134058 z 2.11.2007. Opubl. 4.06.2008
- [11] Wirtualny Doradca Stanusch Technologies: <http://wd.stanusch.com> (7.01.2013)
- [12] Kopp S.: *How People Talk to a Virtual Human - Conversations from a Real-World Application*. W: How People Talk to Computers, Robots, and Other Artificial Communication Partners. Red. K. Fisher. Bremen, SFB/TR 8 Spatial Cognition 2006, s. 101-111
- [13] Zespół Przetwarzania Sygnałów DSP AGH: <http://dsp.agh.edu.pl> (7.01.2013)

## Spis rysunków

Rys. 1.1. Prosty schemat działania i struktury systemu dialogowego.....	14
Rys. 1.2. Schemat komunikacji z Wirtualnym Doradcą .....	17
Rys. 3.1. Rozkład typów odpowiedzi z wiedzy ogólnej i wiedzy specjalistycznej .....	26
Rys. 3.2. Liczba pytań i rozmów dziennie w poszczególnych miesiącach oraz liczba pytań i rozmów miesięcznie dla całego roku 2013 .....	27

## Załączniki

Zał. 1. Odnośniki do witryn poszczególnych systemów dialogowych wymienionych w rozdziale 1.3. Data dostępu: 7.01.2013 .

<b>System</b>	<b>Twórca i rok powstania</b>	<b>Odnośnik do strony internetowej</b>
Eliza	Weizenbaum, 1966	<a href="http://www.masswerk.at/elizabot/">http://www.masswerk.at/elizabot/</a>
How May I Help You?	AT&T, 2001	<a href="http://www.corp.att.com/atllabs/reputation/timeline/01hmihy.html">http://www.corp.att.com/atllabs/reputation/timeline/01hmihy.html</a>
Mercury	MIT, ~2000	<a href="http://groups.csail.mit.edu/sls/research/mercury.shtml">http://groups.csail.mit.edu/sls/research/mercury.shtml</a>
Clarissa	NASA, 2005	<a href="http://ti.arc.nasa.gov/tech/cas/user-centered-technologies/clarissa/">http://ti.arc.nasa.gov/tech/cas/user-centered-technologies/clarissa/</a>
Façade	Procedural Arts, 2005	<a href="http://www.interactivestory.net/">http://www.interactivestory.net/</a>
ITSpoke	University of Pittsburgh, 2006	<a href="http://people.cs.pitt.edu/~litman/itspoke.html#projects">http://people.cs.pitt.edu/~litman/itspoke.html#projects</a>
Sgt Blackwell	USC ICT, 2006	<a href="http://ict.usc.edu/prototypes/sgt-blackwell/">http://ict.usc.edu/prototypes/sgt-blackwell/</a>
Wirtualny Doradca	Stanusch Technologies, 2007	<a href="http://wd.stanusch.com/">http://wd.stanusch.com/</a>
Portal Głosowy	Primespeech, 2008	<a href="http://www.primespeech.pl/portal-glosowy.html">http://www.primespeech.pl/portal-glosowy.html</a>
Siri	Apple, 2011	<a href="http://www.apple.com/ios/siri/">http://www.apple.com/ios/siri/</a>