
Innowacyjny interfejs stron WWW oraz technologia Text Intelligence oparte o przetwarzanie języka naturalnego

Michał Wroczyński

Fido intelligence
michal.wroczynski@fidointelligence.pl

Tomasz Krupa

Fido intelligence
tkrupa@fidointelligence.pl

Nina Daliva

Fido intelligence / PJWSTK
nina.daliva@fidointelligence.pl

Małgorzata Sulęta

Fido intelligence
malgorzata.suleta@fidointelligence.pl

Anna Kowalczyk

Fido intelligence
anna.kowalczyk@fidointelligenc.pl

Wojciech Mościbrodzki

PJWSTK
wojmos@pjawstkw.edu.pl

Abstrakt

Artykuł przedstawia autorskie opracowanie innowacyjnej koncepcji Wirtualnego Asystenta. Przedstawione podejście umożliwia implementację interakcji człowiek-komputer w oparciu o rozmowę w języku naturalnym i osadzenie jej w środowisku typowej strony internetowej. Warte podkreślenia jest wykorzystanie w komunikacji języka polskiego, którego specyfika (m.in. fleksyjność) jest czynnikiem znacznie komplikującym zbudowanie efektywnego i efektywnego interfejsu.

W artykule autorzy przedstawiają swoje doświadczenia z procesu budowy systemu - od założeń koncepcyjnych, przez realizację (w tym - współpracę z aktorami), po etap testów ergonomicznych i wdrożenie rozwiązania. Na szczególną uwagę zasługują praktyczne wnioski dla przyszłości budowy innowacyjnych interfejsów, wynikające z implementacji w przemyśle - prezentowana technologia została zastosowana w kilku produktach rynkowych.

Słowa kluczowe

Text intelligence, język naturalny, baza wiedzy, warstwa językowa, słowa kluczowe, wirtualny asystent, email processor

Fido intelligence

Fido intelligence jest spółką, której strategia rozwoju oparta została na ciągłym inwestowaniu w rozwój działu R&D firmy. Firma tworzy i rozwija technologię przetwarzania języka naturalnego (NLP – natural language processing), co zaowocowało stworzeniem nowatorskiej na skalę światową technologii Text Intelligence.

Celem Firmy jest tworzenie systemów, które automatyzują procesy komunikacji firm czy instytucji z klientami. Dlatego też w portfolio firmy znajdują się zarówno wirtualni asystenci (Virtual Assistant), jak i Email Processor.

Działalność naukowo-badawcza Fido intelligence

Fido intelligence współpracuje z jednostkami naukowymi – Politechniką Gdańską, Akademią Górniczo-Hutniczą przy rozwoju technologii przetwarzania języka naturalnego oraz z Polsko Japońską Wyższą Szkołą Technik Komputerowych. Wspólnie z Politechniką Gdańską (Wydział Elektroniki, Telekomunikacji i Informatyki, Katedra Inżynierii Wiedzy) firma bierze udział w realizacji grantu badawczo naukowego – ISPAD.

ISPAD (Inteligentny System Pozyskiwania i Analizy Danych) – zadaniem opracowywanego prototypu będzie inteligentne pozyskiwanie i analiza językowa publicznie dostępnych informacji z Internetu (fora dyskusyjne, listy dyskusyjne, blogi) w zakresie treści uznawanych za podejrzane w kontekście naruszeń bezpieczeństwa informacyjnego państwa, w tym treści rasistowskich, faszystowskich oraz nawoływania do przemocy. Innowacyjność projektu budowy systemu ISPAD

dotycząca przetwarzania informacji polega na włączeniu i zastosowaniu w procesie tworzenia prototypu algorytmów związanych z przetwarzaniem języka naturalnego. Docelowo ISPAD będzie dedykowanym narzędziem dla polskich służb mundurowych służącym monitorowaniu, analizie podejrzanych treści, wspomagającym pracę dochodzeniową – śledczą i wspomagającym procesy podejmowania decyzji.

Innym z projektów badawczych, współtworzonych przez Fido Intelligence, jest prototyp systemu Cerber. To autorski system tworzony we współpracy z Komendą Główną Policji i Ministerstwem Nauki i Szkolnictwa Wyższego. Jego zadaniem jest ochrona dzieci przed szkodliwymi i groźnymi treściami oraz zachowaniami, które mogą znaleźć w Internecie.

Projekty komercyjne Fido intelligence – Virtual Assistant

Fido intelligence jest firmą o najdłuższych na rynku tradycjach komercyjnego wdrażania wirtualnych asystentów, zwanych również botami, chatbotami, chatterbotami czy avatarami. Virtual Assistant to autorskie rozwiązanie pozwalające na prowadzenie swobodnego dialogu dowolnej liczby osób z umieszczoną w serwisie www wirtualną postacią. Użytkownik może wprowadzać dowolne zapytania za pomocą klawiatury, natomiast system odpowiada mu za pomocą tekstu odpowiedzi oraz syntezy mowy. Dodatkowo, postać przedstawiona sekwencjami video zachowuje się jak żywa osoba, reagując emocjonalnie i gestykulując.



Wśród klientów, dla których zespół Fido intelligence tworzył wirtualnych asystentów są: Telekomunikacja Polska, STU Ergo Hestia, MasterCard, Skarbiec Mennicy Polskiej, Szkoła Wyższa Psychologii Społecznej, Wyższa Szkoła Handlu i Finansów Międzynarodowych, BuildDesk Polska.

Komunikacja człowiek – komputer

Rozwiązanie Virtual Assistant znajduje swoje zastosowanie w dwóch płaszczyznach biznesowych – po pierwsze usprawnia działanie biura obsługi klienta, zmniejszając koszty prowadzenia procesów obsługi (klienci błyskawicznie otrzymują odpowiedzi na swoje pytania poprzez rozmowę z wirtualnym asystentem) oraz na płaszczyźnie marketingowo promocyjnej (obecność wirtualnych asystentów na stronie www wpływa na wzrost oglądalności strony oraz na budowanie pozytywnego wizerunku firmy).

Umieszczenie w serwisie Skarbcza Mennicy Polskiej Wirtualnego Króla Stanisława Augusta Poniatowskiego przełożyło się na czterokrotne zwiększenie oglądalności serwisu.

Wykorzystanie w serwisie www rozwiązania wirtualnego asystenta odpowiada oczekiwaniom klientów. Jak wynika z badań aż 89% konsumentów uważa, że firmy muszą rozpocząć dialog z klientami (Euro Rscg Sensors: Przyszłość Marki Korporacyjnej, 2008), a blisko 70% badanych jest skłonna czekać na połączenie z konsultantem tylko 2 minuty (<http://www.nf.pl/News/788/Czas-klienta>).

Metodologia wdrożenia

Stworzona przez Fido intelligence metodologia wdrożenia oparta jest o model, w którym ciężar przygotowania baz wiedzy i konieczny prac okołoprojektowych, jak stworzenie wizualizacji stoją po stronie Fido. Należy zwrócić uwagę na to, że każde wdrożenie rozwiązania Virtual Assistant jest dostosowane ściśle pod potrzeby klienta (custom made). Celem projektu jest stworzenie rozwiązania, które umożliwi swobodną komunikację w języku naturalnym (potocznym).

Mając na względzie powyższe założenia, w procesie tworzenia baz wiedzy biorą udział, poza programistami również psychologowie, którzy wraz z językoznawcami (autorami baz wiedzy), bazując na polityce wizerunkowej klienta tworzą tzw. wirtualną osobowość specyficzną dla każdego wirtualnego asystenta. To właśnie wirtualna osobowość pozwala na prowadzenie z wirtualnym asystentem rozmowy nie tylko na tematy ściśle związane z firmą, ale również i te dotyczące

zainteresowań wirtualnej osoby, jej ulubionych potraw, kolorów, książek, co wpływa na pozytywny odbiór wirtualnego asystenta przez użytkowników systemu.

Pamięć krótko - i długotrwała

Jednym z wyróżników Virtual Assistant, które mają kolosalny wpływ na działanie systemu i interakcję człowiek – komputer jest wyposażenie systemu w pamięć krótko- i długotrwałą.

Pamięć krótkotrwałą pozwala na zapamiętywanie przedmiotu rozmowy oraz na umiejętność kontekstowego rozwiązywania problemów użytkowników. W przypadku zbyt mało precyzyjnych pytań, system dopytuje rozmówcę celem udzielenia najtrafniejszej odpowiedzi.

Dzięki pamięci długotrwałej, system zapamiętuje użytkownika (np. jego imię, zainteresowania) oraz informacje z poprzednich rozmów (o czym rozmawiał, jakich dokonał wyborów). Pozwala to na łączenie faktów z kolejnych rozmów i wyciągania wniosków, co jest nieocenionym źródłem wiedzy dla firmy, szczególnie dla działów marketingu i sprzedaży. Na przykład, użytkownik rozmawiając z botem na stronie sklepu internetowego pyta o właściwości produktu X i go kupuje. Przy następnej rozmowie bot może zapytać użytkownika o zadowolenie z dokonanego zakupu i zaproponować komplementarny produkt.

Wizualizacja

Wizualizacja to jeden z elementów, które wspomagają interakcję użytkownika z rozwiązaniem Virtual Assistant. Obecnie, najczęściej stosowaną wizualizacją są filmy video. Wirtualny asystent obecny na stronie www może być również przedstawiony poprzez zdjęcie lub animację, jednak to właśnie nagranie video daje

najlepsze efekty komunikacyjne. Użytkownicy odnoszą wrażenie, że prowadzą dialog z żywym człowiekiem.

Fido intelligence, tworząc wizualizację video, posługuje się wykorzystuje najczęściej tzw. green box, czyli nagrywa modela na zielonym tle. W późniejszej obróbce z materiału wykluczowuje się tło, dzięki czemu postać można wstawić w dowolne tło, które może zostać później zmienione (może to być obraz lub film). Tworząc wizualizację należy pamiętać o zachowaniu odpowiedniej „wagi” filmów ze względu na przepustowość łącza użytkowników Internetu, a także na liczbę jednocześnie odtwarzanych obiektów video, które mogą spowodować zbyt duże obciążenie procesora komputera użytkownika. Tworzenie wizualizacji jest zawsze sztuką kompromisu pomiędzy efektami specjalnymi, a możliwościami łącz i maszyn.

Filmiki są przygotowane w postaci flash i wyświetlane są w specjalnie przygotowanej aplikacji końcowej, która jest flaszowym komponentem swf. Użytkownik rozmawiając z botem, w tle ściąga filmiki które wyświetla. Standardowe filmiki mają strumień od 200 do 300 kbps, ale zdążają się także filmiki o bardzo wysokiej jakości. W tym przypadku wielkość strumienia filmików dochodzi do 950kbps.

Jak wynika z doświadczeń Fido intelligence, dla właściwego odbioru i estetyki wykonania, konieczne jest stworzenie minimum 30 sekwencji filmowych, które wyrażają różne stany emocjonalne wirtualnego asystenta (przywitanie, pożegnanie, tłumaczenie, negacja, zaproszenie do rozmowy, złość). Do każdej odpowiedzi udzielonej przez Wirtualnego Asystenta dołączony jest filmik, który uruchamiany jest w aplikacji końcowej. Praca na planie filmowym wymaga od

modela dużych zdolności aktorskich, ponieważ w krótkim czasie musi on odegrać skrajne emocje, takie jak radość, złość, zniecierpliwienie.

Materiał video nagrany kamerą HD umożliwia także późniejsze wykorzystanie nagranych sekwencji do tworzenia aplikacji dialogowych i reklamowych na większe ekrany np. kioski multimedialne.



Jednym z największych projektów wirtualnego asystenta na rynku polskim było stworzenie Wirtualnego Króla Stanisława Augusta Poniatowskiego dla Skarbcza Mennicy Polskiej. Założenia projektu dotyczyły nie tylko budowy skomplikowanej bazy wiedzy, również bazy wiedzy historycznej, ale również odnosiły się do warstwy wizualnej. Na potrzeby sesji filmowej stworzono cały strój królewski, wzorując się na królewskich portretach. W trakcie dwudniowej sesji nagrano materiał, z którego powstało aż 90 sekwencji, przedstawiających emocje króla.



Fot. 1 Wirtualny Król Stanisław August Poniatowski,
www.skarbiecmennicy.pl

Interakcja z wirtualnym asystentem

Bardzo ważną cechą botów Fido intelligence jest to, iż pozwalają one na tworzenie więzi między nimi a użytkownikami. Internauci zadają botu pytania o miejsce, gdzie aktualnie się znajduje, o jego lub jej wiek, zawód, zajęcia, itd. Dlatego też w każdej bazie wiedzy rozpoznania tego typu pytań umieszczane są domyślnie. Boty odpowiadają zarówno na podstawowe pytania o swoją tożsamość, typu: „jak masz na imię?”, „ile masz lat?”, „kiedy się urodziłeś/urodziłaś?”, ale posiadają też rozpoznania szczegółowych pytań dotyczących swojego wyglądu, na przykład „czemu masz na sobie taką brzydka bluzkę?”.

Nawiązywanie więzi rozpoczyna się zazwyczaj w momencie tworzenia kontekstu. „...Internauci dążą więc do wykreowania kontekstu. Choć wirtualny, jest on niezbędny, by osadzić komunikację na odpowiednich solidnych podstawach. Kreowanie kontekstu polega na ustalaniu podstawowych parametrów związanych z odbiorcą, w związku z tym w internetowych konwersacjach pojawiają się zwykle pytania, które w tym typie komunikacji nazwałbym elementarnymi. Zanim rozpocznie się bliższy kontakt konwersacyjny,

pojawia się pytanie: „kim jesteś?”.(Grzenia Jan, Komunikacja językowa w internecie, str. 90)

Virtual Assistant zapewnia również dialogowość. Oznacza to, iż w przypadku komunikacji pomiędzy botem-nadawcą, a użytkownikiem-odbiorcą następuje zrównanie ról. Boty dają użytkownikowi możliwość interakcji. Internauta nie musi już być biernym adresatem treści, ale sam może brać czynny udział w akcie komunikacyjnym, czyli w pełni korzystać z funkcji magicznej języka. Jednocześnie boty same potrafią przejmować inicjatywę w rozmowie za pomocą losowanych propozycji zmiany tematu.

Poniżej przedstawiamy wybrane fragmenty rozmów z Virtual Assistants:

Dialog 1:

Eliza [Ekolan SA]: *Czy pozwolisz, że zapytam, czy mieszkasz w Trójmieście?*

Użytkownik: tak skarbie

Dialog 2:

Użytkownik: *kocham cie*

Wiktoria [BuildDesk Polska]: *To bardzo miłe, ale jako istota wirtualna nie znam uczucia miłości. Za to doskonale znam się na zagadnieniach dotyczących termomodernizacji.*

Dialog 3:

Użytkownik: *wiesz, ja kończę. Przepraszam Ado, może się jeszcze kiedyś spotkamy*

Ada [Fido intelligence]: *Przecież teraz się spotkaliśmy... Rozmawiamy sobie miło...*

Użytkownik: *oj tak tylko ja już muszę iść:(*

Ada [Fido intelligence]: *W takim razie do zobaczenia następnym razem.*

Warto zwrócić również uwagę, że w przypadku Wirtualnego Króla Stanisława Augusta Poniatowskiego, który posługuje się językiem stylizowanym na czasy stanisławowskie, Internauci podejmują zabawę językiem i również stylizują swoje wypowiedzi:

Dialog 4:

Użytkownik: *Dzięki stokrotne Jego Królewskiej Mości za audiencję dzisiejszą, a wkrótce o następną poproszę. Miłego dnia Królu!*

Król Stanisław [Skarbiec Mennicy Polskiej]: *Przyjemność cała po mej stronie pozostaje, tegom pewien.*



Fot. 2 Wirtualny Król Stanisław August Poniatowski, www.skarbiecmennicy.pl

Oprócz prowadzenia rozmowy poprzez wpisywanie dowolnych pytań w oknie zapytania Virtual Assistant, użytkownik może również posługiwać się proponowanymi przez bota hiperłączami w tekście jego odpowiedzi. Dzięki nim użytkownik widzi, w którym kierunku zalecane jest prowadzenie rozmowy w danym temacie, a także przyspiesza to komunikację, ponieważ linki te reprezentują określone pytania, jakie użytkownik musiałby wpisać.

Jednocześnie wybrane odpowiedzi mogą powodować automatyczne przełączanie podstron w serwisie, w którym umieszczony jest wirtualny asystent. Dzięki temu, może on ilustrować swoje wypowiedzi szerszymi opisami, grafikami i innymi materiałami, w jaki wyposażony jest serwis www.

Za komunikację z silnikiem wirtualnego asystenta, wprowadzanie tekstu przez użytkownika, wyświetlanie odpowiedzi, odgrywanie syntezy, wyświetlanie filmów, otwieranie powiązanych stron www odpowiada specjalnie zaprojektowana aplikacja końcowa, stworzona w technologii FLASH. Dzięki niej możliwa jest szybka i prosta integracja Virtual Assistant z serwisem www klienta.

Technologia Text Intelligence

Doświadczenie zdobyte przez Fido intelligence w trakcie tworzenia wirtualnych asystentów wpłynęło na stworzenie autorskiej technologii przetwarzania języka naturalnego.

Text intelligence łączy metody przetwarzania języka naturalnego (Text Analytics) z metodami stosowanymi w sztucznej inteligencji (Artificial Intelligence).

Technologię można podzielić na dwa obszary: „Text” oraz „Intelligence”.

Obszar „Text” służy do analizy dowolnego tekstu pisanego w języku naturalnym (czyli potocznym), aby nadać mu odpowiedni opis formalny. Nie ma więc znaczenia, ani stylistyka ani ortografia pisanego tekstu. Wyniki uzyskane w warstwie Text są następnie poddawane wnioskowaniu obszarze „Intelligence”. W efekcie tego procesu system potrafi dokładnie określić zagadnienia poruszane w tekście i odpowiednio na nie zareagować.

Takie podejście do przetwarzania tekstu pozwala poprawnie analizować dowolnej długości tekst (np. e-mailie), a nie tylko krótkie wypowiedzi przy rozmowie z Wirtualnym Asystentem.

Warstwa językowa – analiza płytka i głęboka

Nowatorstwo przetwarzania i analizy tekstu stosowane przez Fido Intelligence przejawia się w rezygnacji z ontologicznego podejścia do tego zagadnienia.

Stosowane przez innych dostawców podejście wykorzystujące teorie ontologii (nauka o znaczeniu), w praktyce korzysta z metod tzw. słów kluczy. W analizie tej pojawia się problem „eksplozji kombinatorycznej”, która polega na lawinowym tworzeniu dużej ilości wariantów wypowiedzi. Dla przykładu: analizując zdanie „Dzieci kupiły telefon dla mam”, przy wspomnianej metodzie słów kluczy otrzymujemy błędne interpretacje:

mama, kupować, telefony, dla, dzieci

1. *Mama kupuje telefony dla dzieci.*
2. *Mama kupowała telefony dla dzieci.*
3. *Dzieci kupowałyby telefony dla mam.*
3. *Dzieci kupują telefony dla mam.*
4. *Mam telefony kupione dla dzieci*

W takim podejściu spotykamy się również z problemem homonimii (relacja wyrażania różnych znaczeń za pomocą identycznych form językowych). Formę leksemu (wyraz rozumiany jako abstrakcyjna jednostka systemu słownikowego języka) „mam” można interpretować na wiele sposobów:

1. *Mam telefony*
2. *Telefony dla mam*

W rezultacie otrzymujemy błędną interpretację tekstu. Dlatego niezbędne jest stworzenie nowych metod technologii prowadzących do właściwej interpretacji tekstu.

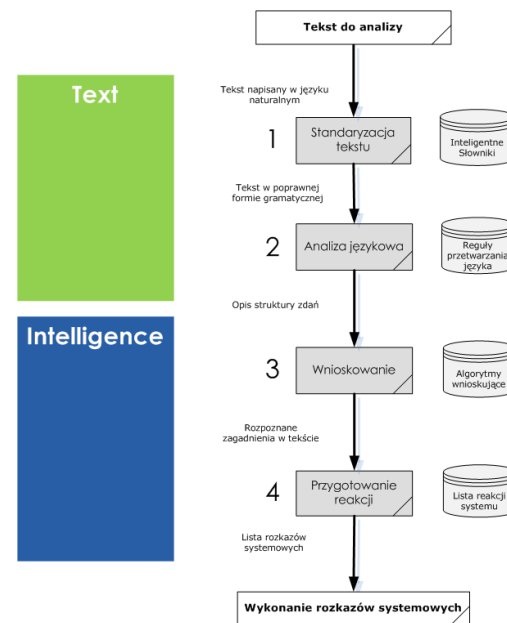
Nowe podejście, stosowane przez Fido intelligence, zakłada korzystanie z najnowszych osiągnięć inżynierii lingwistycznej. Tekst jest analizowany nie tylko semantycznie, ale w celu uzyskania właściwej interpretacji, również stosuje się analizę gramatyczną. Analiza przebiega w dwóch fazach:

- powierzchniowo składniowej
- głębinowo składniowej

Dzięki czemu wybieramy jeden właściwy wariant wypowiedzi: *Dzieci kupiły telefony dla mam.*

Text: Intelligence skutecznie rozwiązuje ten problem i w rezultacie otrzymujemy precyzyjną interpretację tekstu.

Telefony dla mam



Schemat przedstawiający działanie technologii Text Intelligence

Pierwszym etapem przetwarzania tekstu w technologii TEXT INTELLIGENCE jest standaryzacja tekstu. Operacje na tekście wykonywane są w oparciu o słowniki ortograficzne, synonimów, slangu, a także tzw. INTELLIGENTNE SŁOWNIKI stworzone przez firmę Fido intelligence. Uwzględniają one takie parametry jak rozkład liter na klawiaturze czy osadzenie danego słowa

w kontekście, a ich zadaniem jest znalezienie poprawnej formy słowa błędnie wprowadzonego przez autora tekstu (np. nadawcę e-maila).

Przygotowany w ten sposób tekst poddawany jest wnioskowaniu na płaszczyźnie INTELLIGENCE, gdzie system precyzyjnie określa poruszane zagadnienia i odpowiednio na nie reaguje. Analiza językowa zachodzi w oparciu o autorskie reguły przetwarzania języka, w wyniku czego powstaje formalny opis struktury zdań. Tak przygotowany tekst poddawany jest wnioskowaniu, na podstawie którego system nazywa i klasyfikuje poruszane w nim zagadnienia.

Email Processor

Jednym z rozwiązań Fido intelligence, opartych o technologię Text Intelligence jest rozwiązanie Email Processor, które w 2009 zostało uznane za Innowację Roku 2009.

E-mail Processor jest jedynym w Europie systemem, który potrafi rozpoznać problemy poruszane w treści e-maila oraz automatycznie na nie odpowiedzieć. Email Processor został stworzony w oparciu o autorską technologię Text Intelligence, która umożliwia skuteczną analizę dowolnego tekstu pisanego w języku naturalnym. System generuje wcześniej przygotowaną odpowiedź, którą wysyła bezpośrednio do Klienta lub przesyła szablon odpowiedzi do konsultanta. Skraca to znacznie czas reakcji na e-maila wysłanego przez Klienta do firmy oraz obniża koszty utrzymania Biura Obsługi Klienta i standaryzuje wysyłane odpowiedzi.

Zastosowane w Email Processorze algorytmy do przetwarzania tekstu umożliwiają przetworzenie nawet

kilku tysięcy maili w ciągu godziny, przy użyciu jednego średniej klasy serwera.

Email Procesor przetwarza tekst według zasad gramatycznych języka polskiego, co pozwala na jego poprawną analizę z uwzględnieniem kontekstu, jest to szczególnie ważne w przypadku języków fleksyjnych, jakim jest język polski.

Wykrywanie emocji

System jest ponadto wyposażony w dodatkową funkcjonalność potrafiącą określić emocje, z jakimi tekst był pisany przez Klienta. Emaille nacechowane negatywnym ładunkiem emocjonalnym zyskują najwyższy priorytet i są obsługiwane w pierwszej kolejności. Nie tylko maile, w których pojawiają się inwektywy, będą uznane przez system jako negatywne emocjonalnie, ale również te, w których nadawca wyraża swoje zniecierpliwienie, na przykład: „Kontaktuję się z wami już po raz trzeci w tym miesiącu, a mój telefon znowu nie działa. Kiedy to wreszcie naprawicie?”.

Szybka reakcja na maile od zniecierpliwionych klientów pozwala firmie łagodzić konfliktowe sytuacje w sposób szybki i skuteczny oraz zapobiegać utracie Klientów.

Email Processor za pomocą jednego średniej klasy serwera może w ciągu godziny przetworzyć około 5000 tysięcy e-maili, bez udziału konsultanta. Człowiek odpowiadający ręcznie jest w stanie odpowiedzieć średnio jedynie około pięć takich e-maili na godzinę.

Bibliografia

[1].Grzenia Jan: Komunikacja językowa w internecie, PWN,2008.