

Komputerowe systemy dialogowe i interfejs użytkownika

Wprowadzenie

Jolanta Bachan

2020-10-15

2020-10-22

2020-10-29

2020-11-05

Informacje do kontaktu

- email: jbachan@amu.edu.pl
 - jolabachan@gmail.com
- strona internetowa: <http://bachan.speechlabs.pl/>
- Konsultacje online na MS Teams po wcześniejszym umówieniu.

Syllabus (1)

- Wprowadzenie do systemów dialogowych - historia i współczesne systemy.
- Architektura systemów dialogowych, komunikacja multimodalna.
- Systemy syntezy mowy (wizualna - talking heads, mowa ekspresywna, gesty i mowa ciała – humanoidy).
- Rozpoznawanie mowy, rozumienie mowy, identyfikacja mówcy.

Syllabus (2)

- Projektowanie interfejsu użytkownika (podejścia skupione na użytkowniku oraz na zadaniu).
- Systemy dialogowe w badaniach psycholingwistycznych (np. ang. *alignment*).
- Projektowanie systemów dialogowych.
- Ewaluacja systemu.
- Projekt studenta – projektowanie, implementacja i ewaluacja systemu dialogowego.

Literatura

Branigan, H.P., Pickering, M.J., Pearson, J. & McLean, J.F. 2010. Linguistic alignment between humans and computers. In: *Journal of Pragmatics*, 42, pp. 2355–2368

Gibbon, D., Mertins, I. & Moore, R. 2000. *Handbook of Multimodal and Spoken Dialogue Systems: Terminology, Resources and Product Evaluation*. New York: Kluwer Academic Publishers

Mehler, A. L. Romary, D. Gibbon. 2012. *Handbook of Technical Communication*. Berlin/Boston: Mouton de Gruyter

Gibbon, D., Moore, R. & Winski, R. 1997. *Handbook of Standards and Resources for Spoken Language Systems*. Berlin: Mouton de Gruyter

Bachan, J. 2011. *Communicative Alignment of Synthetic Speech*. Praca doktorska. Instytut językoznawstwa. Uniwersytet im. Adama Mickiewicza.

Zaliczenie

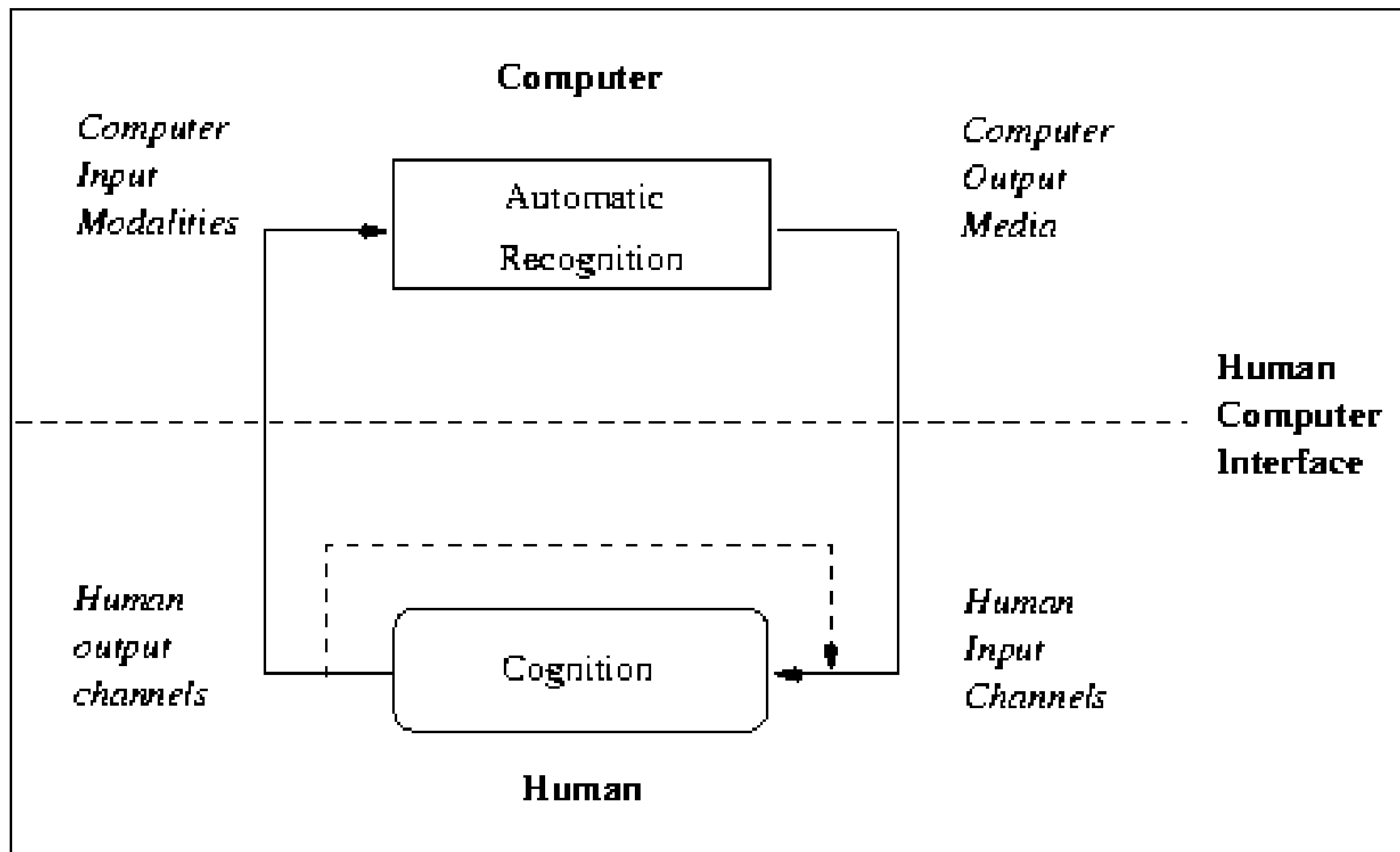
- Aktywność na zajęciach
- Obecność na zajęciach (dopuszczone 2 nieobecności)
- Uzyskanie pozytywnej oceny z kolokwium zaliczeniowego (50% oceny końcowej)
- Wykonanie projektu zaliczeniowego, którym będzie system dialogowy (50% oceny końcowej)
- Rejestracja w USOSie

POWODZENIA!

Zaliczenie

- Zaliczenie – test pisemny: przedostatnie zajęcia
- Projekty studentów: ostatnie zajęcia

Komunikacja człowiek-komputer



Gibbon, Mertins & Moore 2000, from Schomaker et al.1995

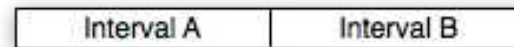
Jolanta Bachan

Relacje czasowe (Allen 1983)

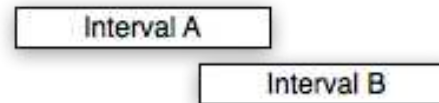
A is before B or
B is after A



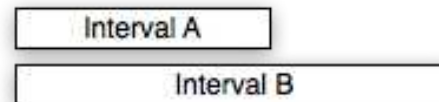
A meets B or
B is met by A



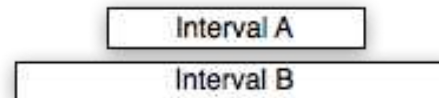
A overlaps with B or
B is overlapped by A



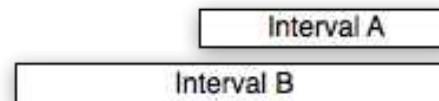
A starts B or
B is started-by A



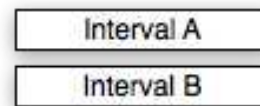
A during B or
B contains A



A finishes B or
B is finished-by A



A and B are cotemporal



Systemy dialogowe, które już znacie

- Text-In-Text-Out

```
>>> import nltk
```

```
>>> nltk.chat.chatbots()
```

System dialogowy, który już znacie

Hello. How are you feeling today?

>Great! And you?

Why don't you tell me?

>I am sure you are feeling great too.

Did you come to me because you are sure me am feeling great too.?

>Yes.

OK, but can you elaborate a bit?

>You are feeling great, because you came to my class.

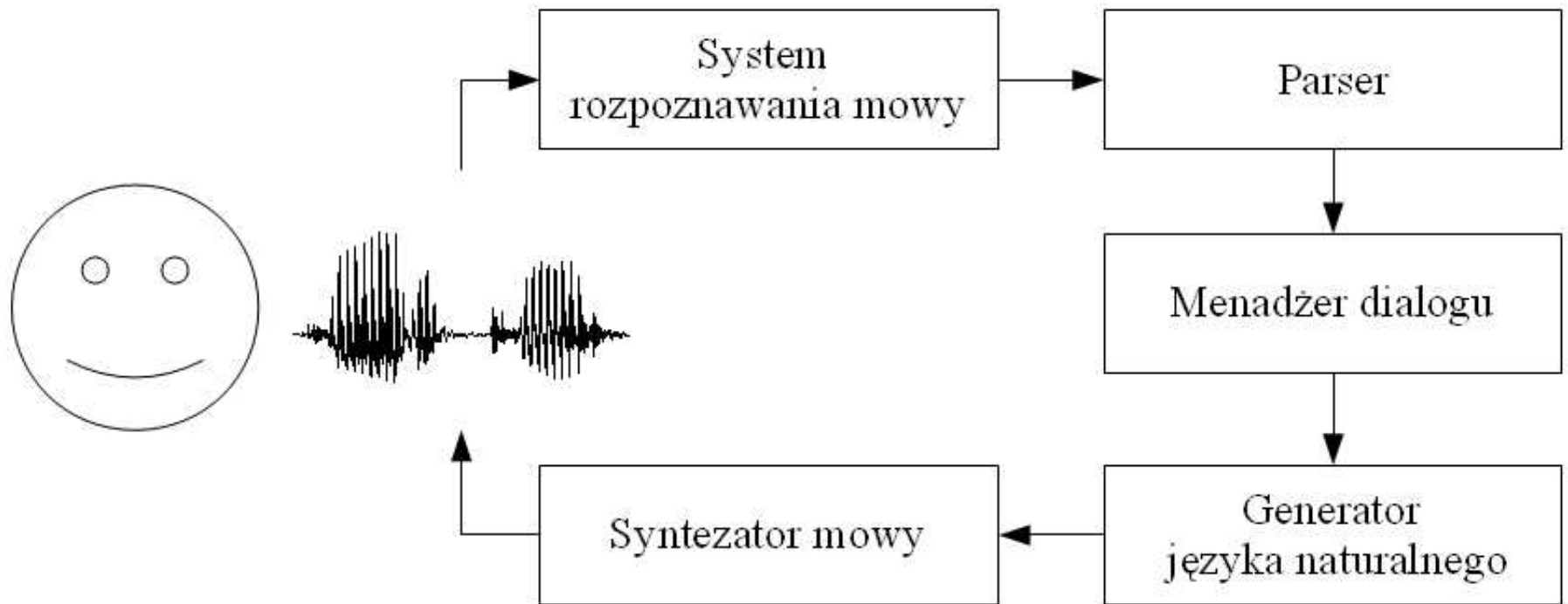
Perhaps you're really talking about yourself?

>Yes, I am.

You seem quite sure.

Ogólny schemat systemu dialogowego

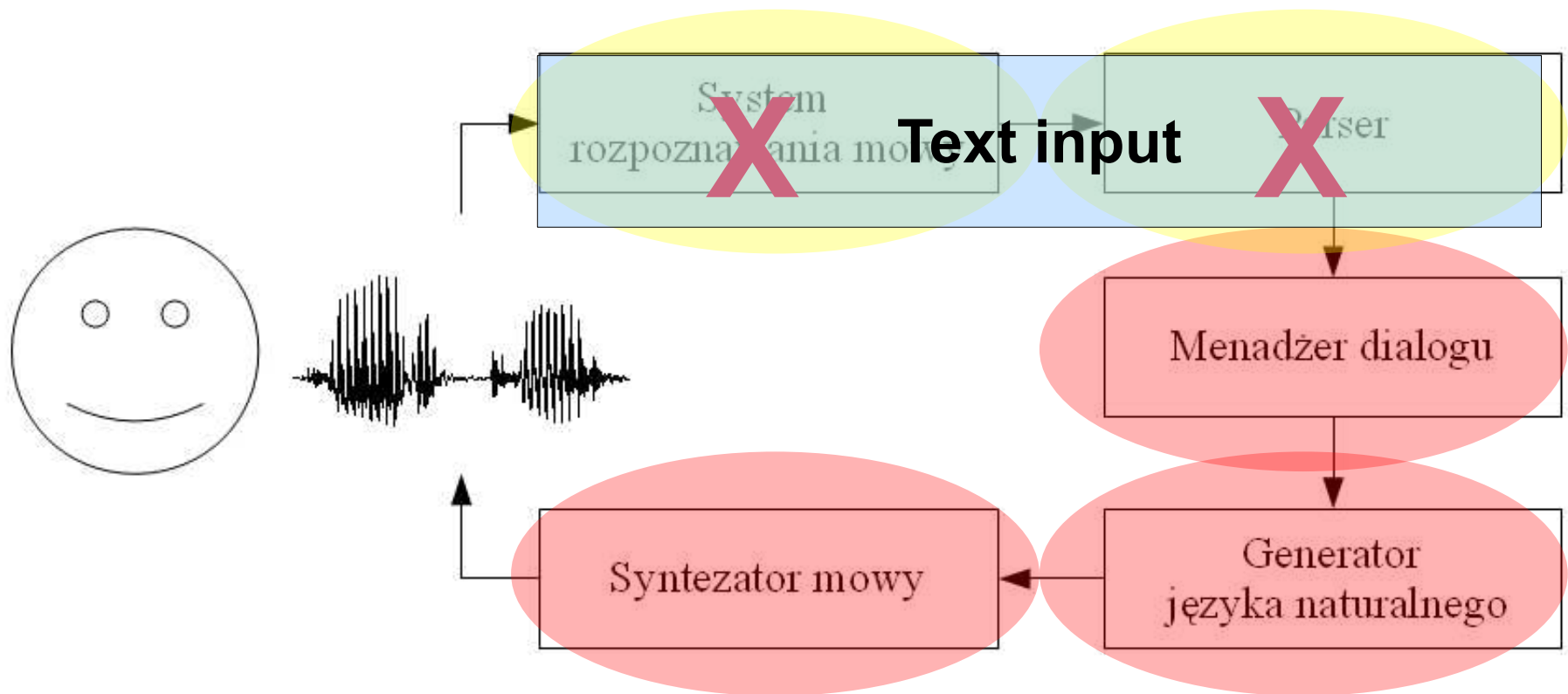
Spoken dialogue system



Przykład

- The Big Bang Theory - Howard's phone
 - <https://www.youtube.com/watch?v=Y0hl1-06gOo>
- The Big Bang Theory - Siri Voice Recognition
 - human-computer relationships
 - human-computer misunderstandings
 - <https://www.youtube.com/watch?v=8kloHgfZPW0>
 - <https://www.youtube.com/watch?v=9Gvltvuujh8>

Schemat Text-In-Voice-Out



Schemat TIVO jest często stosowany w systemach poleceń i kontroli oraz w systemach protezowych (np. system Stephen Hawkinga)

Multimodalne systemy dialogowe

- tekst
- mowa
- grafika
 - mimika twarzy
 - gesty
- dotyk
- ruch

Radio Rex

The first machine that recognized speech was probably a commercial toy named “Radio Rex” which was sold in the 1920’s. Rex was a celluloid dog that moved (by means of a spring) when the spring was released by 500 Hz acoustic energy. Since 500 Hz is roughly the first formant of the vowel [eh] in “Rex”, the dog seemed to come when he was called.



David, E.E. & Selfridge, O.G. 1962. Eyes and Ears for Computers. Bell Telephone Laboratories, Murray Hill, N.J. Proceedings of the Institute of Radio Engineers. May, 1962. Volume: 50 Issue:5 On page(s): 1093 – 1101

“Put-that-there”

- Richard A. Bolt. 1980. “Put-that-there”: Voice and gesture at the graphics interface, *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, strony 262-270
- <https://www.youtube.com/watch?v=RyBEUyEtxQo>

Systemy dialogowe end-to-end (E2E)

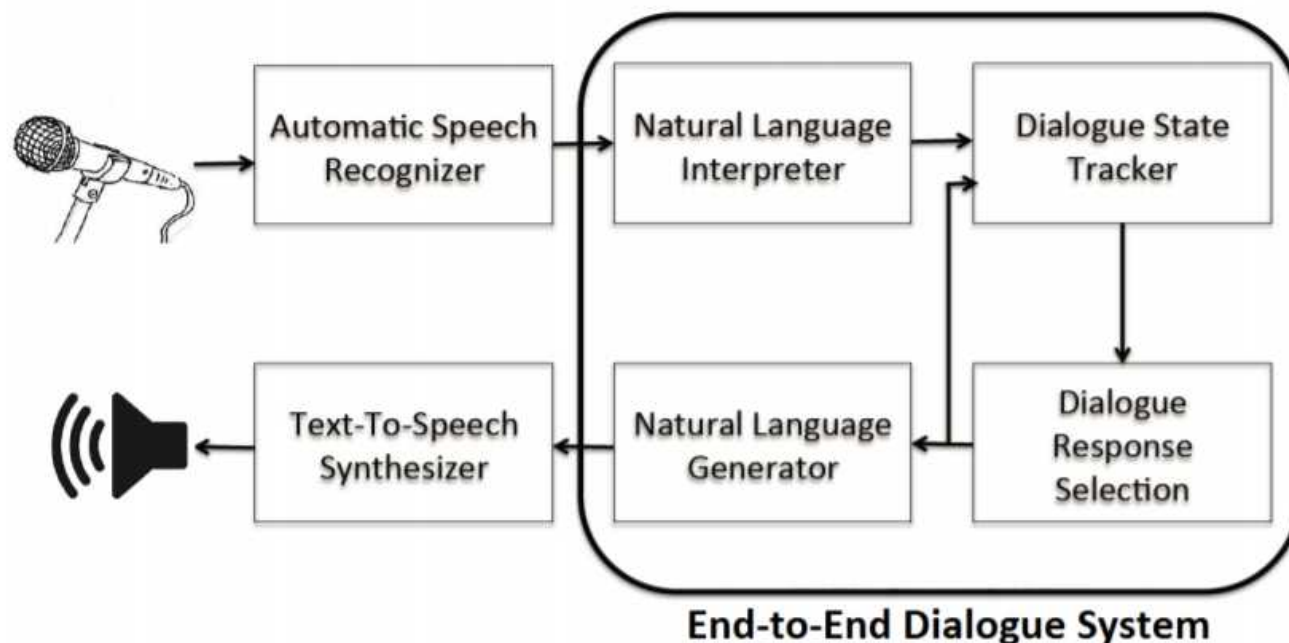


Figure 1: An end-to-end dialogue system replaces the traditional components of a dialogue system with a single statistical model.

R. Lowe et al. 2017. Training end-to-end dialogue systems with the Ubuntu Dialogue Corpus, *Dialogue & Discourse* 8(1) (2017) 31–65

Afektywne systemy dialogowe

- System nie tylko rozumie, co do niego mówimy, ale też potrafi ocenić nasz stan emocjonalny i się do niego przystosować, a nawet na niego wpływać
 - Jeśli jesteśmy zdenerwowani w danym momencie, to system może próbować nas uspokoić
 - np. architektura oparta na drzewach decyzyjnych

Afektywne systemy dialogowe

- ocena stanu emocjonalnego na podstawie
 - prozodii wypowiedzi
 - długości wypowiedzi
 - języka/słownictwa
 - cech wizualnych (uśmiech, nadmierna gestykulacja)
 - tętna

Typy systemów dialogowych

- informacyjne
- transakcyjne (umożliwiają subskrypcję do czasopism lub transakcje bankowe)
- diagnostyczne (udzielają pomocy technicznej)
- edukacyjne (tutoriale do nauki np. języków)
- rozrywkowe: entertaining & chatting
- mobilni wirtualni personalni asystenci – Virtual Personal Assistants (VPA), np. Google Now, Siri, Cortana, Bixby

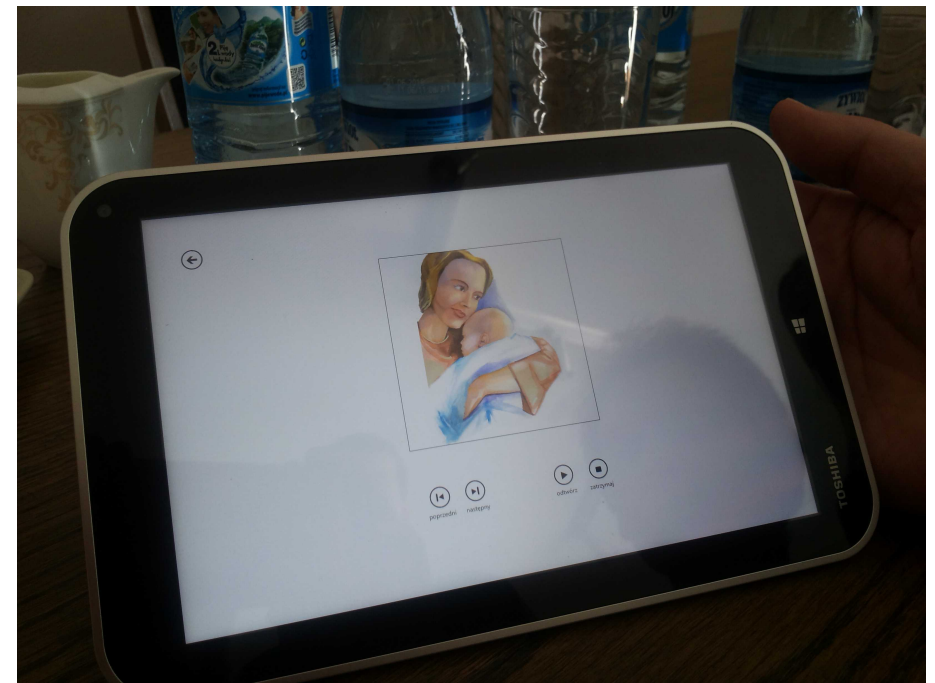
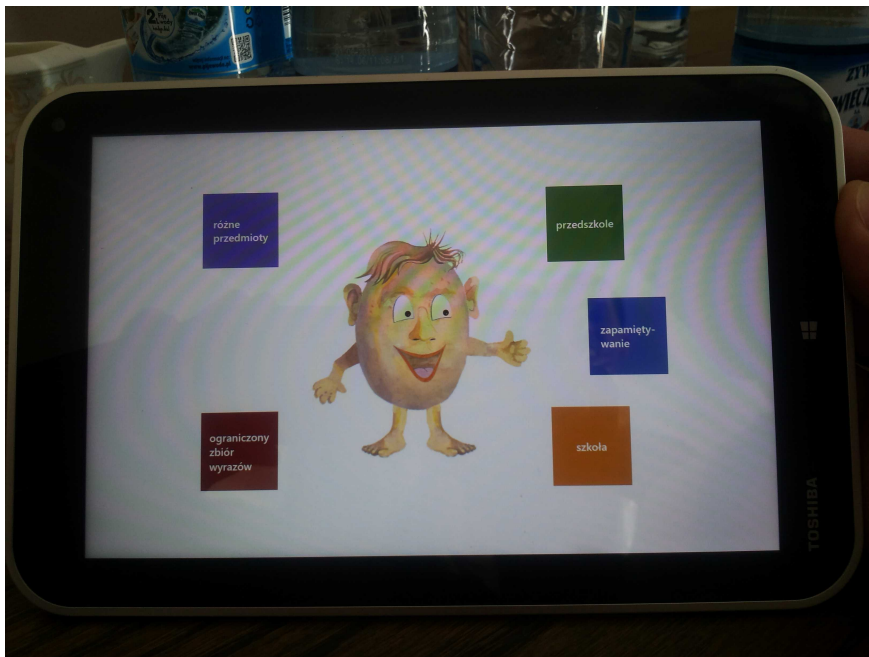
Przykładowe systemy dialogowe

Intelligent Animated Agents for Interactive Language Training

- assesses children's understanding of selected fields of study
- designed for profoundly deaf children
- *"Baldi" the Virtual Tutor Helps Hearing-Impaired Children to Learn Speech*
- <http://www.nsf.gov/od/lpa/news/press/01/pr0119.htm>
- <https://www.youtube.com/watch?v=Z9aqgT-AFx0>
- https://www.youtube.com/watch?v=mDfKdRh5h_g

Testy percepcyjne dla dzieci z implantami ślimakowymi

- output: audio
- input: dotknięcie konkretnego obrazka



Dialogue system for elderly people



DS: *When and where the picture was taken?*

U: *At the LREC conference.*

DS: *Who is in the picture?*

U: *Dafydd Gibbon. Do you know Dafydd Gibbon?*

DS took initiative:
Yes, he is my brother in law. (the knowledge database is updated)

U: *When is the next LREC conference?*

DS activates the real time Internet search engine...

The Semaine Project

- The Semaine project is an EU-FP7 1st call STREP project and aims to build a SAL, a Sensitive Artificial Listener, a multimodal dialogue system which can:
 - interact with humans with a virtual character
 - sustain an interaction with a user for some time
 - react appropriately to the user's non-verbal behaviour
- <https://semaine-db.eu/>
- https://www.youtube.com/watch?v=6KZc6e_EuCg

SimSensei

- This effort seeks to enable a new generation of clinical decision support tools and interactive virtual agent-based healthcare dissemination/delivery systems that are able to recognize and identify psychological distress from multimodal signals.
- <http://ict.usc.edu/prototypes/simsensei/>
- <https://www.youtube.com/watch?v=ejczMs6b1Q4>

NAO

- face detection and recognition
- text-to-speech (8 languages available)
- speech recognition (word-spotting function)
- whole body motion
- sound localisation (4 microphones)
- <https://www.youtube.com/watch?v=nNbj2G3GmAo>

Alpha 2

- The World's First Humanoid Robot for the Family. Intelligent, Interactive and Expandable!
- <https://www.youtube.com/watch?v=E1AtfHm4hf8>

Android-Robots from Japan

- <https://www.youtube.com/watch?v=3IFuv1AVouM>
- <https://www.youtube.com/watch?v=hIHrvQ7D5OU>
- Actroid (2:30)
<https://www.youtube.com/watch?v=V7TKBFJMbi0>

My Jibo – Family Robot

- <http://www.myjibo.com/>

NTT Disruption to release the jibo robot for healthcare and education.

The company has strengthened its portfolio in social cognitive agents, aiming to improve people's lives through tech.

- <https://www.youtube.com/watch?v=H0h20jRA5M0>

ASIMO

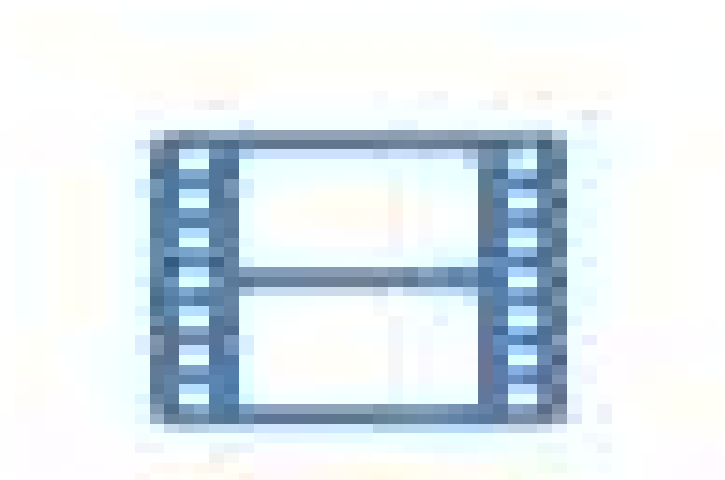
- <https://www.youtube.com/watch?v=skXYr8BzjpM>

Pytanie: Gdzie jeszcze można spotkać systemy dialogowe?

Pytanie: Gdzie jeszcze można spotkać systemy dialogowe? -

- infolinie
 - BOK
 - komunikacja miejska
 - kina
- muzea/galerie
 - U: *Co jest na tym obrazie?*
- samochód: <https://www.youtube.com/watch?v=t29GFBk-2-E>
 - DS: *Wyczuwam, że jesteś nietrzeźwy. Nigdzie nie pojedziesz!* <https://www.youtube.com/watch?v=QTwynnBMX5M>
rozmieszczenie mikrofonów
- inteligenty dom

Inteligentny dom



Nie tylko mowa

- W systemach dialogowych nie tylko można wykorzystać tekst lub mowę, ale też znaki migowe
- THETOS: Translator i słownik Polskiego Języka Migowego
 - tłumaczy teksty w języku polskim na polski język migowy
- <http://thetos.polsl.pl/>

THETOS demo

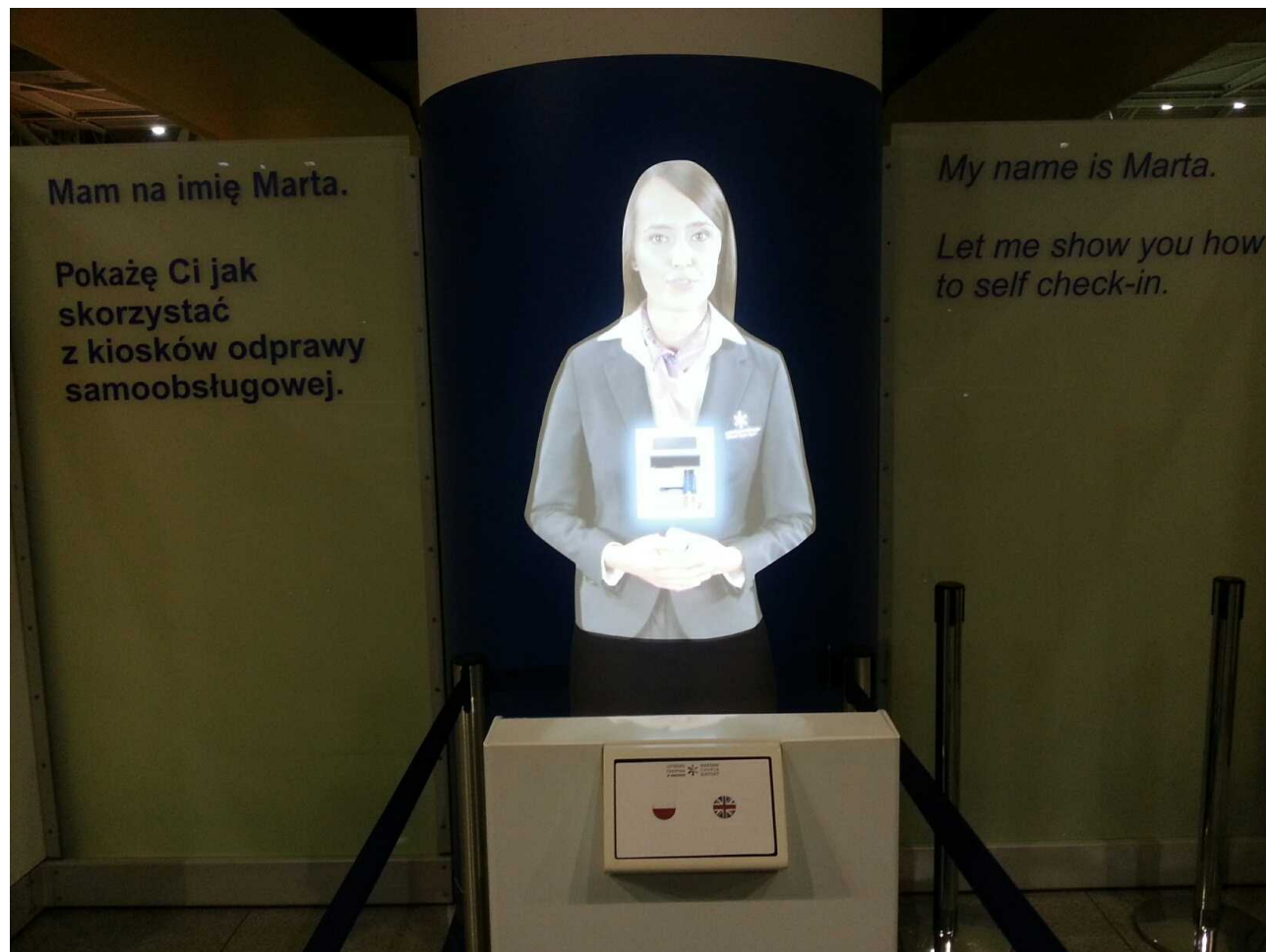


Android Ms. Aiko Chihira & sign language

- <https://www.youtube.com/watch?v=ttvQrMu1bEk>
- <https://www.youtube.com/watch?v=4BRycib5IWI> (2:37)

Tańsze rozwiązanie...? Hologram

- Marta, wirtualna asystentka na Lotnisku Chopina w Warszawie
- http://warszawa.wyborcza.pl/warszawa/1,34889,15210832,Awatar_pomoze_w_odprawie_na_Okeciu__Pierwszy_w_Polsce.html



Czas na rodzicielstwo...

- Babysitting robots
- <https://www.youtube.com/watch?v=7brpD9kBCMo>
- http://www.spacedaily.com/reports/In_Japan_robot_babysitter_always_ready_to_play_999.html

If parents want to leave their children with the kid-sized machine, they arrange for the youngsters to wear special badges that bear codes which the robot can read.

The robot can then identify the children by their names and ages and chat with them, even though its vocabulary remains limited.

- NEC's PaPeRo - tells jokes, gives quizzes and uses radio-frequency identification chips to track kids.
- <http://www.wired.com/2008/12/babysittingrobo/>

Sharkey does, however, take instruction from psychologist Harry Harlow's famous and controversial tests on the importance of maternal care for monkeys, and ostensibly people: when nursed by inanimate objects, they grew up to be withdrawn and socially dysfunctional.

"The question is, if robots could take care of your children, would you let them?"

PaPeRo

<https://www.youtube.com/watch?v=gZSXiV3AbJA>

NEC's (Japan) range of "Partner-type Personal Robots", or "PaPeRos", are designed for maximum interaction with humans. The Papero Childcare robot has been specially adapted with two (stereoscopic) camera "eyes" to recognize surroundings and faces, eight microphone "ears" to detect and recognize speech – even from multiple sources – and mobile-phone connectivity that allows PaPeRo to stay in touch with the absent parents. It can also speak with a 3,000-word vocabulary and react to touch via nine sensors. Children in the care of PaPeRo wear an ultrasonic transmitter and wireless microphone, allowing the robot to track their positions at all times.

From:

<http://www.guinnessworldrecords.com/world-records/first-robot-babysitter/>

Japanese baby robots

- Yatoro

- <http://www.telegraph.co.uk/news/newsvideo/weirdnewsvideo/7505847/Robot-baby-aims-to-increase-birthrate-in-Japan.html>
(unavailable)
- <http://www.telegraph.co.uk/expat/expatnews/7420381/Japanese-baby-bot-with-runny-nose-teaches-parenting-skills.html>
- <http://infocult.typepad.com/infocult/2010/04/the-baby-from-the-uncanny-valley.html>
- https://www.youtube.com/watch?v=LdkXzURJLLo&feature=emb_logo

Another Japanese robot treads the line between sweetness and horror. Baby Yataro is intended to make Japanese adults more willing to have children.

The mechanics are interesting: limbs are metallic skeleton, and the sweet face is a balloon upon which expressions are literally projected. Balloon boy, at last!

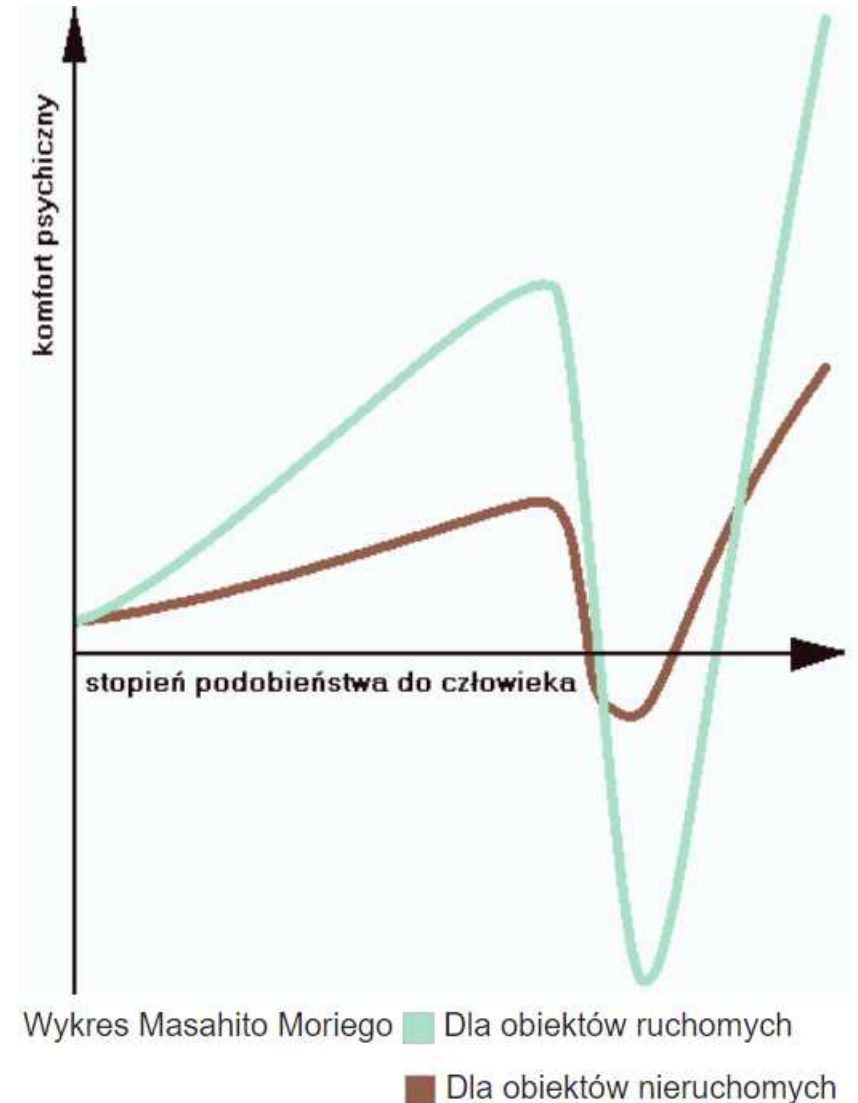
- Noby

- <http://www.telegraph.co.uk/news/newsvideo/weirdnewsvideo/7831807/Japanese-scientists-invent-baby-robot.html>
- <https://www.dailymotion.com/video/xdp1rl>

Dolina niesamowitości

- **Dolina niesamowitości** – termin stosowany w hipotezie naukowej, zgodnie z którą robot, rysunek lub animacja komputerowa wyglądający bądź funkcjonujący podobnie (lecz nie identycznie) jak człowiek, wywołuje u obserwatorów nieprzyjemne odczucia, a nawet odrazę.
- English: uncanny valley

Wikipedia



PARO Therapeutic Robot

- Paro is a therapeutic robot baby harp seal, intended to be very cute and to have a calming effect on and elicit emotional responses in patients of hospitals and nursing homes, similar to animal-assisted therapy except using robots. (Wikipedia)
- Cute Baby Seal Robot - PARO Therapeutic Robot #DigInfo

<https://www.youtube.com/watch?v=oJq5PQZHU-I>

Conversations with robots

- Talking robots, CommU and Sota

<https://www.youtube.com/watch?v=lb0WY9u5h4c>

NAOs - Two robots, one brain

- Researchers taught a robot to defy orders. Now they've built him a friend who silently communicates with him.
- <https://www.facebook.com/quartznews/videos/1360893547277677/?pnref=story>

Talking heads

- <https://www.youtube.com/watch?v=x8B7OxjB05k>
- <https://www.youtube.com/watch?v=TMxKcbQcnK4>
- <http://www.hansonrobotics.com/hansons-living-robots/>
- <https://www.youtube.com/watch?v=IhVu2hxm07E>
- https://www.youtube.com/watch?time_continue=330&v=Bg_tJvCA8zw
(2:20)

Sophia

Hanson Robotics' human-like robot

- <https://www.facebook.com/cnbcinternational/videos/749203651847444/?fref=nf>
- <http://www.cnn.com/2016/03/16/could-you-fall-in-love-with-this-robot.html>

Sophia's lifelike skin is made from patented silicon and she can emulate more than 62 facial expressions. Cameras inside her "eyes," combined with computer algorithms, enable her to "see," follow faces and appear to make eye contact and recognize individuals. A combination of Alphabet's Google Chrome voice recognition technology and other tools enable Sophia to process speech, chat and get smarter over time. Hanson is working with IBM and Intel to explore integrating some of their technologies.

- Kobieta-robot pierwszy raz w historii otrzymała obywatelstwo. Zostało przyznane przez Arabię Saudyjską
- <https://nauka.rocks/sophie/>
- <https://www.facebook.com/nauka.jest.niesamowita/videos/1720724478000810/>
- <https://www.hansonrobotics.com/sophia-2020/>
- AI Artificial Intelligence – Sztuczna inteligencja

Śpiewający robot

- 2009:
 - <https://www.youtube.com/watch?v=mfxkhzGqZIs>
 - Vocaloid: tekst + melodia => śpiew (Yamaha)
- 2010:
 - https://www.youtube.com/watch?v=_migLQ802Go
- 2018, HRP-4C (Miim):
 - <https://www.youtube.com/watch?v=QCuh1pPMvM4>

LovePlus – dating sim

- <https://en.wikipedia.org/wiki/LovePlus>
- Dating sims (or dating simulations) are a video game subgenre of simulation games, usually Japanese, with romantic elements. They are also sometimes put under the category of neoromance. The most common objective of dating sims is to date, usually choosing from among several characters, and to achieve a romantic relationship.

From *Wikipedia*, https://en.wikipedia.org/wiki/Dating_sim

- How to play?
 - <https://www.youtube.com/watch?v=y-rBkRWolZc>
- First marriage between a man and a virtual character
 - <https://www.youtube.com/watch?v=hsikPswAYUM>

Akira Yukino



Akira Yukino

- Post by akirayukino96, Aug 31, 2014 at 3:48am

*Does anybody else know that IF in the later future they create another LovePlus game, will they possibly add Akira Yukino? I only ask because Akira is soo **adorable** and I don't see any harm in her taking part of the game series...*

Samantha

- Dual-purpose robot: sex mode and family mode
- While it is true Samantha looks far more realistic than some other models, it is pretty clear this is still a robot we are dealing with. She has built-in artificial intelligence, which means people can have normal conversations with her at any given time. This is especially useful when running her in “family mode”, as holding a conversation is pretty important in this regard. She is, according to the creators, capable of telling jokes and even providing motivational quotes. It’s a very interesting concept, especially considering how this robot is rather affordable. (£3,500)
- <https://themerkle.com/sex-robot-samantha-offers-a-family-mode-to-make-it-less-weird/>

Mark 1 (?)

- <https://www.rickyma.hk/#https://www.youtube.com/watch?v=fn89fmsz1g>
- <https://www.rickyma.hk/mark-1-news/>
- speech synthesis
- facial expression
- voice & face recognition
- object tracking
- body movement – head, hands, legs

Mark 1 like Scarlett Johansson

Robi

- <https://www.youtube.com/watch?v=RPjlJKRa7Uw>
- <https://www.youtube.com/watch?v=nm1hprKd1Jc>

Kirobo

- "kibō" (希望) "hope" in Japanese + "robo"
 - <https://en.wikipedia.org/wiki/Kirobo>
- Toyota's Robot Astronaut Heading for International Space Station
 - <https://www.youtube.com/watch?v=VcuFk-QK5CM>
- Kibo Robot Project !!!NEW!!! 2013
 - https://www.youtube.com/watch?v=XS_B-yLceSA
- Kirobo the first robot to have a conversation in space
 - <https://www.youtube.com/watch?v=xqShesZ3v-g>

Warto wiedzieć

Co to jest?

- efekt Elizy
- bot i botyka
- Jabberwacky i Cleverbot
- VPA (Virtual Personal Assistant)
- Autonomous robot / Robot autonomiczny
 - William Grey Walter
 - In one experiment he placed a light on the "nose" of a tortoise and watched as the robot observed itself in a mirror. "It began flickering," he wrote. "Twittering, and jigging like a clumsy Narcissus." Walter argued that if it were seen in an animal it "might be accepted as evidence of some degree of self-awareness." (*Wikipedia*)

Test it!

- <http://www.cleverbot.com/>
- <http://www.jabberwacky.com/>

Psylon: *How are you?*

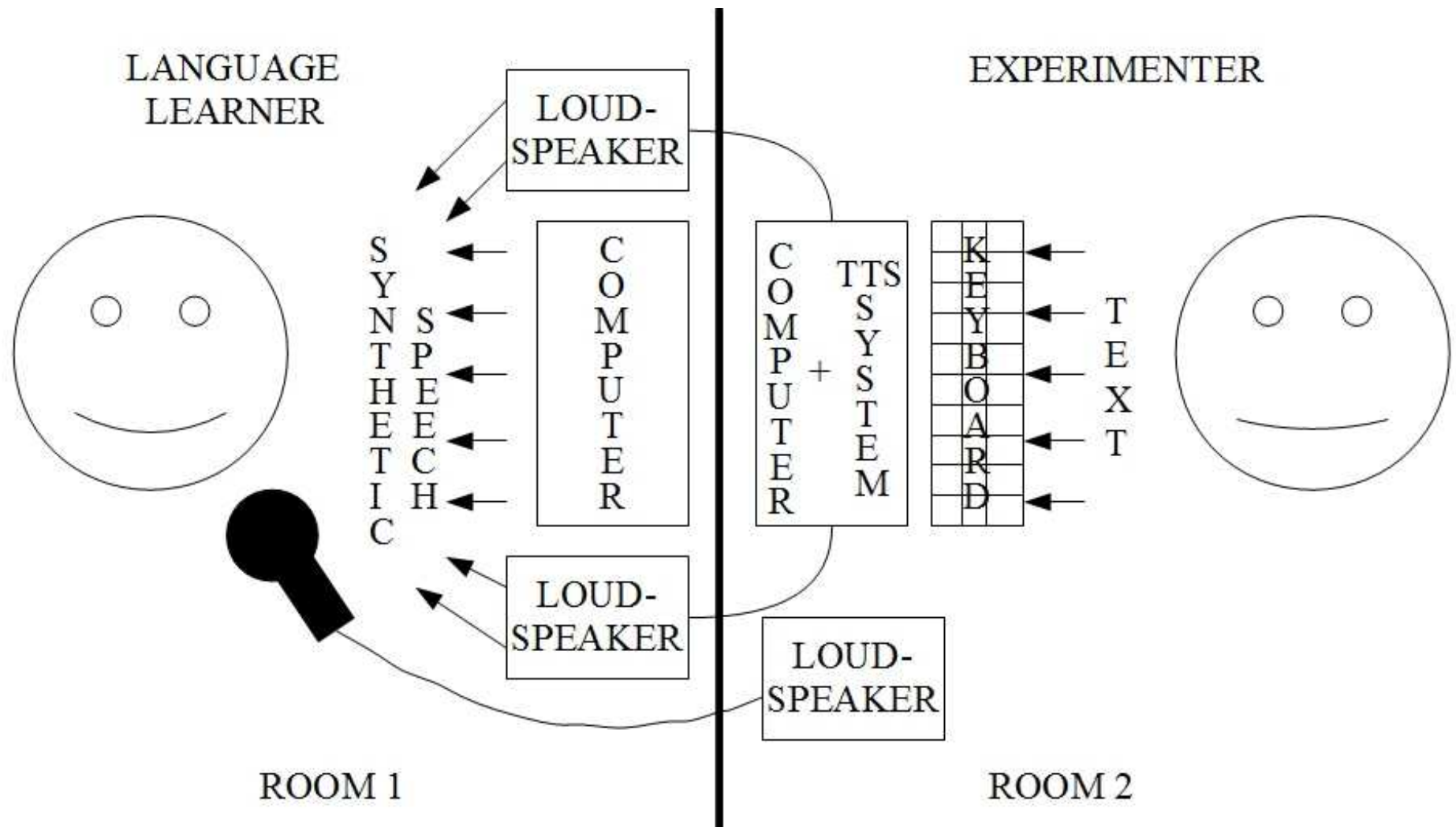
I: I am a little bit tired.

Psylon: *Then why won't you sleep?*

I: I have to work.

Psylon: *Then start working.*

Eksperyment Czarnoksiężnika z Oz



Co to jest test Turinga?

Can a Machine Think? (1950)

- *If a computer could think, how could we tell?*
- imitation game

Co to jest test Turinga?

- Można wskazać jeszcze jedną wątpliwość dotyczącą testu Turinga: W niektórych przypadkach, aby zaliczyć test, maszyna musiałaby symulować brak posiadanej wiedzy czy umiejętności. "Zdradzenie" się z taką wiedzą czy umiejętnościami powodowałoby, że nie zaliczy testu. Przykładem może być prośba o podanie wyniku działania matematycznego typu: 1234235 razy 2,23 (wynik: 2 752 344,05).

Wikipedia

Nagroda Loebnera

- <https://www.ocf.berkeley.edu/~arihuang/academic/research/loebner.html>
- <https://www.ocf.berkeley.edu/~arihuang/academic/research/intro.html>
- Mitsuku wins for the 5th time in 2019
- Eugene Goostman

Mitsuku: <http://www.mitsuku.com/>



<https://www.pandorabots.com/mitsuku/>

⚡ POWERED BY PANDORABOTS



Shop Kuki Bot Battle Gear!

Meet Kuki!

Mitsuku, or Kuki to her close friends, is a [record-breaking](#), five-time winner of the Loebner Prize Turing Test and the world's best conversational chatbot (according to folks like [Google AI Research](#)). As featured in the [New York Times](#), [Wall Street Journal](#), [BBC](#), [Guardian](#), [Wired](#), and [Radiolab](#).



Mitsuku a.k.a. Kuki

World's best conversational AI

[Chat Now](#)

Cleverbot
[https://
www.clever
bot.com/
amused](https://www.cleverbot.com/amused)

Cleverbot comes very close to passing the Turing Test
Techniche 2011, IIT Guwahati, India, 3rd September 2011

A high-powered version of Cleverbot took part alongside humans in a formal Turing Test at the Techniche 2011 festival. The results from 1,334 votes were astonishing...

Cleverbot was judged to be 59.3% human.

The humans in the event achieved just 63.3%.

"It's higher than even I was expecting, or even hoping for. The figures exceeded 50%, and you could say that's a pass. But 59% is not quite 63%, so there is still a difference between human and machine." *Rollo Carpenter*

During the event people voted how human-like responses seemed, from 0 to 10. Thirty of the audience volunteered, and chatted on 3 screens in 10 rounds of 4 minutes each. Half the conversations were human-human.

The Turing Test was proposed by early computer scientist Alan Turing. 'Thinking' and 'intelligence' are hard to define, so he suggested a test, with people and machines communicating via text. If a person could not tell human from machine more than half the time, the machine should be called intelligent.

Turing made a prediction that in 5-minute conversations machines would pass the test 30% of the time by the year 2000. That mark has now surely been passed. And with 30 conversations, 100 separate voting individuals and a total of 1,334 votes cast, the Turing Test at IIT Guwahati holds mathematical significance. Even so, different tests with different people and circumstances can and certainly will see different results.

"It was mesmerising to watch a bot chatting just like a human, and people finding it so hard to distinguish." *One of the humans at the event*

Our AI has been learning online for 15 years, but has recently seen an exponential growth curve of data and visitors. It is normal for there to be 100,000 conversations a day, each with 20 or more interactions. There will have been a billion things said to Cleverbot by the end of this year alone.

<http://www.projectmurphy.net/>

What if a Bethoven was a rock star?

17:51



You asked: " What if a Bethoven was a rock star? " so I will attempt to picture that.

17:51

You can say:

"More" - For more ideas,

🙄 - Didn't like it, change the photo,

😬 face - Didn't like it, change the face,

😊 - I like it, show another photo,

😊 face - I like it, show another face,

😡 - To report this image as offensive,

"Attributions" - To see the attributions of the photo sources



17:51

Synteza mowy

- System przetwarzania tekstu pisanego na mowę
- Text-to-Speech (TTS)
- TTS powinien być w stanie przeczytać każdy tekst, ale w praktyce nie jest to takie proste do zrealizowania

Synteza mowy

- Parametryczna synteza mowy
 - synteza formantowa
 - synteza artykulacyjna (VTdemo, vocal tract demo)
- Konkatencyjna synteza mowy
 - synteza difonowa
 - synteza trifonowa
 - unit selection
- Synteza mowy oparta o ukryte modele Markova (Hidden Markov Models HMM) albo sieci neuronowe (Neural Networks, NNs)
 - ang. HMM-based speech synthesis
 - Statistical Parametric Synthesis
- E2E TTS (End-to-End Text-to-Speech)

Automatyczne Rozpoznawanie Mowy (ARM)

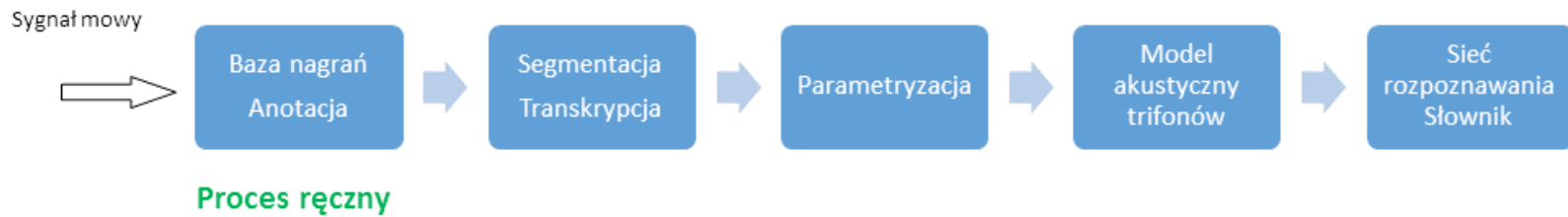
- Automatic Speech Recognition (ASR)
- Konwersja sygnału mowy na tekst

Poznański System Rozpoznawania Mowy Polskiej ARM

- <https://speechlabs.pl/oferta/arm/>
- Projekt „Zaawansowany system automatycznego rozpoznawania i przetwarzania mowy polskiej na tekst, dedykowany dla służb odpowiedzialnych za bezpieczeństwo państwa”.
- Dwa główne procesy:
 - proces uczenia się – proces ten rozpoczyna się od zebrania odpowiedniej bazy nagrań, która następnie zostaje poddana procesowi anotacji (segmentacja i transkrypcja mowy). Parametryzacja tak przygotowanej bazy pozwala na zbudowanie modelu akustycznego trfonów, który staje się podstawą Dekodera w procesie rozpoznawania mowy. Obok modelu akustycznego budowany jest Słownik (lista słów, obecnie ok. 450.000) z teksów pisanych języka polskiego (transkrypcja nagrań, artykuły z gazet, teksty prawnicze takie jak wyroki, akty prawne, umowy)
 - proces rozpoznawania: jądrem procesu rozpoznawania mowy jest Dekoder, który zamienia parametry akustyczne wyekstrahowane z sygnału mowy na tekst (wyrazy znajdujące się w Słowniku systemu ARM), a następnie model językowy koryguje błędy Dekodera (na podstawie n-gramów) i tak wygenerowany i skorygowany tekst jest prezentowany użytkownikowi

Schemat budowy systemu rozpoznawania mowy polskiej ARM

Proces uczenia



Proces rozpoznawania



Demenko, G., Cecko, R. Szymański, M., Owsianny, M., Francuzik, P., Lange, M. (2012). Polish speech dictation system as an application of voice interfaces. W: Dziech, A., Czyżewski, A. (Red.) Proceedings of 5th International Conference on Multimedia Communications, Services and Security, Kraków 2012 (pp. 68–76). Springer for Research and Development.

Przeczytaj: Inteligentne Systemy Dialogowe

- http://en.wikipedia.org/wiki/Dialog_system

Dziękuję za uwagę!

*A teraz pokonwersujcie z
systemami dialogowymi dostępnymi online.*