Emotional robots and people: challenges for education

L. Devillers

devil@limsi.fr

29 September 2022

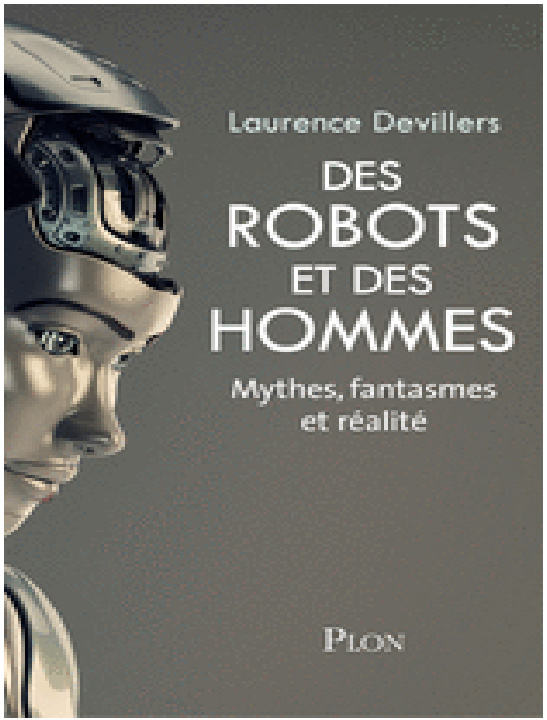
Laurence Devillers

Professor in AI at Sorbonne-University/CNRS (France) and a Human-Machine Affective Interaction Specialist, GPAI expert, convenor AHG6-AI-Enhanced nudging JTC21 CEN-CENELEC

Artificial Intelligence (AI)
Imitation

Artificial intelligence is a branch of computer science that aims to create « intelligent » machines.

The core problems of artificial intelligence include programming computers for certain traits such as: Knowledge, Reasoning, Problem solving, Perception, Learning, Planning, Ability to manipulate and move objects, etc.

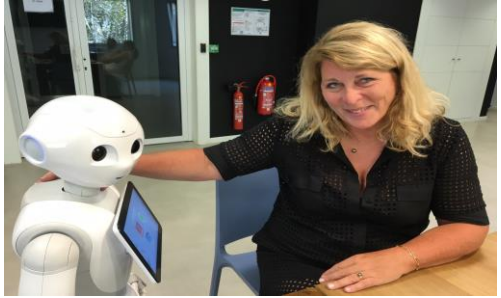


Artificial intelligence opens up important opportunities in the field of numerous applications such as education or health diagnosis and treatment support with the aim of, for example, better student or patient follow-up.

Social and Emotional Robots: Useful Artificial Intelligence without Consciousness

Myths, Fantasies and Reality

The challenge is to educate on the subject of AI in schools and more widely in society.



Team: « Affective and social dimensions of spoken interactions with (ro)bots: technological and ethical issues »

L. Devillers – Professor at Sorbonne University – LISN-CNRS (devil@limsi.fr)

Since 2001 team 5-10 researchers (15 PHD students, 5 in progress)

- Detection and analysis of affective and social components of audio signals exchanged between a human and his partner (human, robot or virtual agents)
- Dynamic speaker profil modeling (affective and interactionnel dimensions such as extraversion…) to adapt the strategy of a dialog system
- Combinaison of verbal and non verbal audio and audiovisual information
- Real-life emotions – mixtures of emotion (pos/neg) in H-H dialogs
- Induction techniques (WoZ) with machine (robots or agents)
- Ethical issues –nudging with affective computing

CHAIR AI HUMAAINE 2019-24: HUman-MAchine Affective INteraction & Ethics

The planned scientific work focuses on the detection of social emotions in human voice, and on the study of audio and spoken language manipulations (nudges), intended to induce changes in the behavior of the human interlocutor.

Main research lines: Four axis of research will be addressed within this chair

Axis 1	Socio-affective states detection from paralinguistic/linguistic cues using deep learning techniques and other statistical techniques
Axis 2	Nudging system using hybrid symbolic/deep reinforced learning and adaptive learning
Axis 3	Experiment in real use-cases: attention at school, nudging in education, assistive elderly
Axis 4	Dissemination of the results: methodology of audit experiments, evaluation metrics & Demystification: user empowerment

The human being and society in the digital challenge: ethical issues

Humans: Asymmetry of knowledge and power (data collection, consent, manipulation), Human-Machine co-adaptation

Societal consequences: change of relationship to the world, knowledge, emotions, interpersonal relations, virtual/real

Specificities of digital ethics: influence, fascination, questioning of human autonomy, social and political relationships, contribution of digital technology to sustainable development objectives

Challenges & opportunities: anticipating these developments and the conditions of our free will, understanding the digital divide

It is important to demystify and disseminate robotics and AI science -> AI READY

- Imagination of our contemporaries about robotics, emotions in machines and more generally AI are mainly founded on science-fiction narratives and myths.
- Expressions used by experts such as "robots are autonomous", "they make decisions", "they learn by themselves" are not understood as metaphors by those outside the technical research community.
- To mitigate ideas originating from science fiction that mainly underline gloomy consequences, it is important to demystify and disseminate robotics and AI science by engaging in public discussion and debate with all citizens.

AI: 3 periods

1rst : Expert systems
Perception, Learning, Abstraction, Reasoning

2nd : Statistical learning -> big data + machine learning

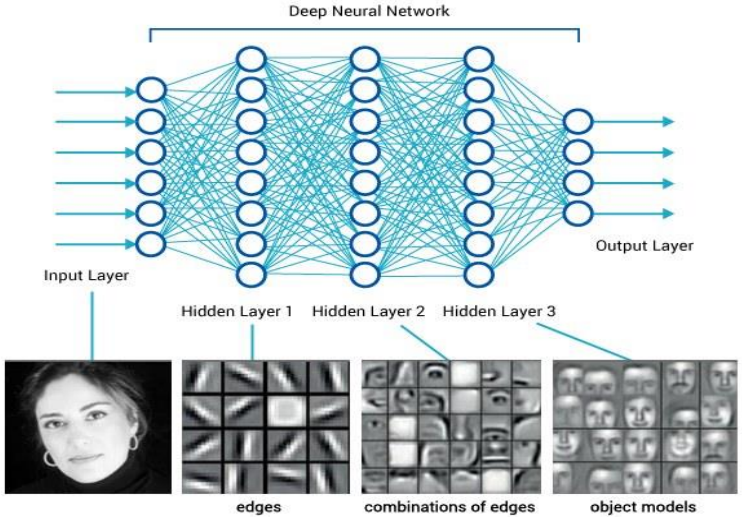
unsupervised/supervised learning/reinforcement learning
Perception, Learning, Abstraction, Reasoning

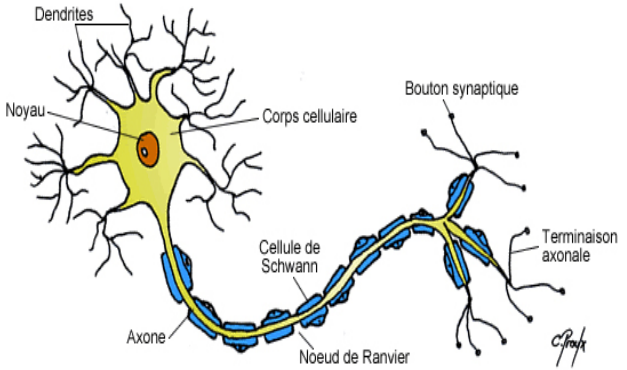
Learn without understanding

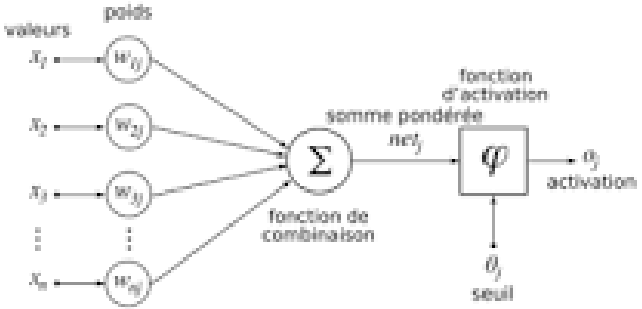
3rd : Contextual adaptation, common sense
Perception, Learning, Abstraction, Reasoning

Ethics of machines - > transparency, explicability, loyalty

Machine learning – artificial neuron (McCulloch & Pitts) 1943







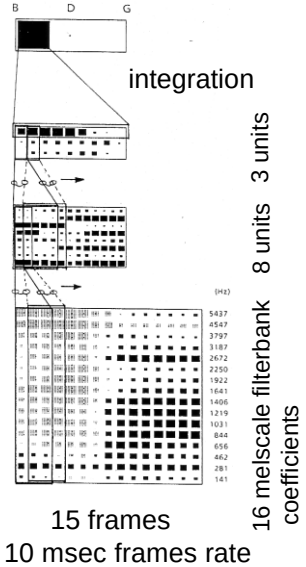
Time-Delay Neural Network (TDNN) - 1987

The TDNN – Waibel, Hinton, et al.
(the first “Convolutional” Net)

- Context at Each Layer
- Shift Invariant Training
- “Deep” Network
- Well Suited for Speech & Image Processing

L. Devillers PHD thesis

Hybrid TDNN/HMM - 1992



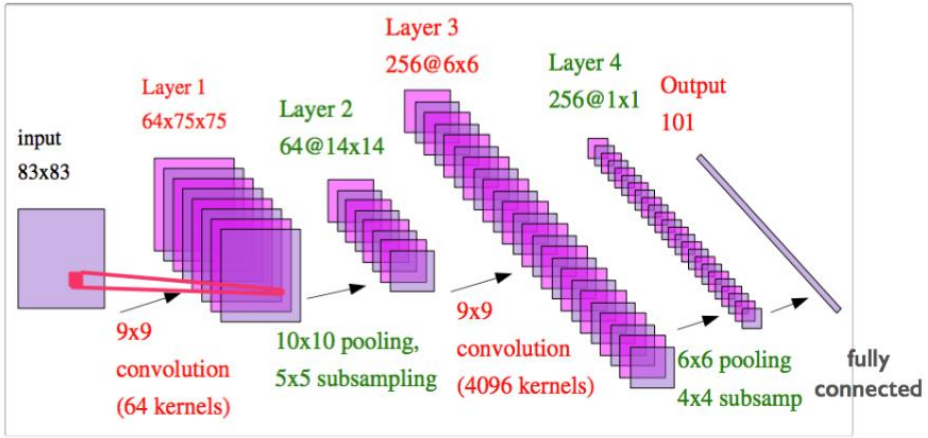
Output Layer

Hidden Layer 1

Hidden Layer 2

Input Layer

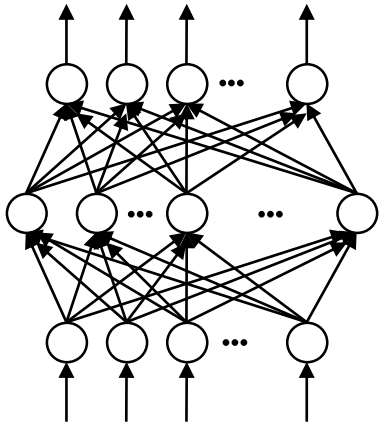
TDNN's/Convolutional Networks in Character Recognition (1998)



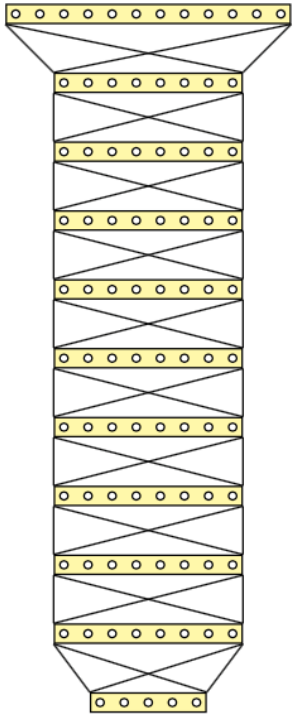
TDNN's are applied to Character Recognition and Image Processing as "Convolutional Neural Networks" [LeCun, 1998]

2010 Neural Nets: Bigger, Deeper, Faster

- NN's Revisited (~2010)
- More Units, More Layers (!),



→



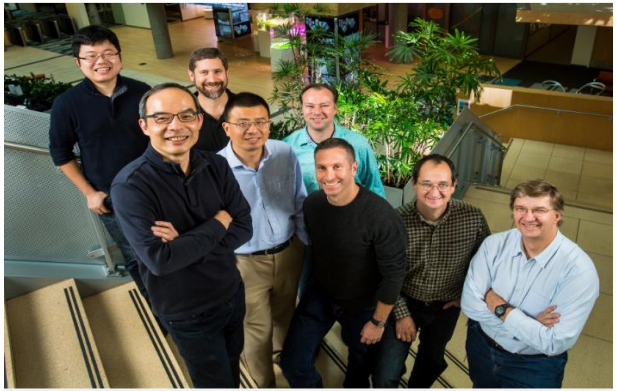
Since 2010:

- More Data
- More Computing
- More Neurons, More Layers
- Enormous Performance Gains!!! In Speech 30%!
- Surprise Gains (30%++)
- Took Community by Storm
- Branched to Every Field

Conversational Speech recognition

- Conversational Speech: High Error Rates Persisted for 25 years.
- In 2017, Human Parity Possible, Using Structures of TDNN/CNN
- Deep learning is the reigning monarch of AI.

Historic Achievement: Microsoft researchers reach human parity in conversational speech recognition



Microsoft researchers from the Speech & Dialogue research group include, from back left, Wayne Xiong, Geoffrey Zweig, Xuedong Huang, Dong Yu, Frank Seide, Mike Seltzer, Jasha Droppo and Andreas Stolcke. (Photo by Dan DeLong)

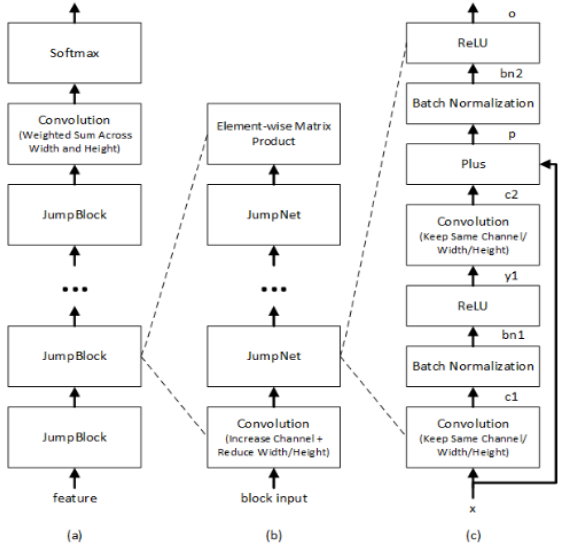


Fig. 1. LACE network architecture

Demystification
Narrow AI – Today

- What can it Do?
  - Can Approach Human/Super-Human functional Performance
  - Super-Human Access to Data
  - Tasks Requiring Judgement & Perception
  - Can Handle Ambiguity
- But:
  - Requires Massive Amounts of Labeled Data
  - World Reasonably Well-Defined
  - Cannot Easily Abstract to New Task, no common sense, no creativity (human sense)
- Anticipation or marketing ?
  - Some AI experts and also transhumanists write that conscience will emerge in future AI systems and that the Machines will feel emotions!!! The Machine has No body – A. Damasio

Fiction - AGI

The Voice technology revolution

- This new era of voice-enabled devices and their rapid adoption signals a major shift for the digital world and consumer interaction models with a potential to augment human welfare and well-being in many ways: improve the customer experience, personalize the communication, give medical advice, help elderly, create a virtual companion, etc.
- These social (ro/chat)bots can have negative consequences, unfair or manipulation if not regulated. We need to build trustworthiness within a sustainable market.

Assistant - Companion


Google Home


Amazon Echo


portal


HomePod
Siri


DuerOS
BAIDU


AI Cube
HUAWEI


SK telecom
NUGU
인공지능 음성인식 디바이스


kakao


Clova

From vocal assistants to social and affective robots

Azuma Hikari,

600M people use voice assistants at least once a week

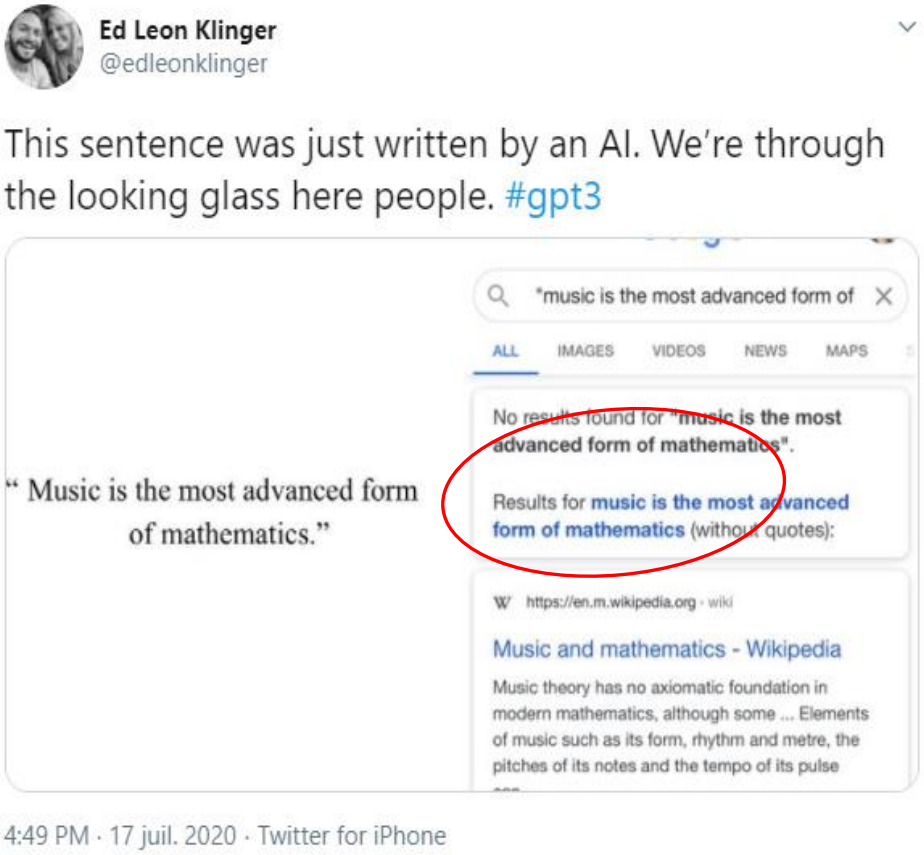
Conversational Agents

- A chatbot is a machine that interacts with users in natural language and is usually integrated into a digital platform
- Advancing conversational AI chatbots with Transformers (trained specifically on data from huge set of conversations) -> allow the chatbot to engage in free dialogue on a potentially infinite number of topics (LamDA Google).
- How to explore and improve the understanding of the functioning and limits of these transformers is a big issue!

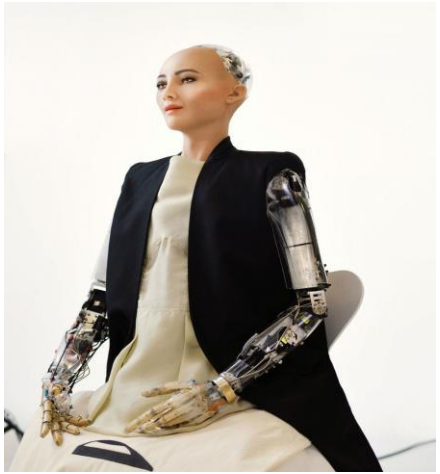
Goals of chatbots. Written by GPT-3

“The team at Google didn't make GPT-1 to solve any particular philosophical problem. They just wanted to see if they could make a computer reason like a human.”

https://pastebin.com/3AEtjv35



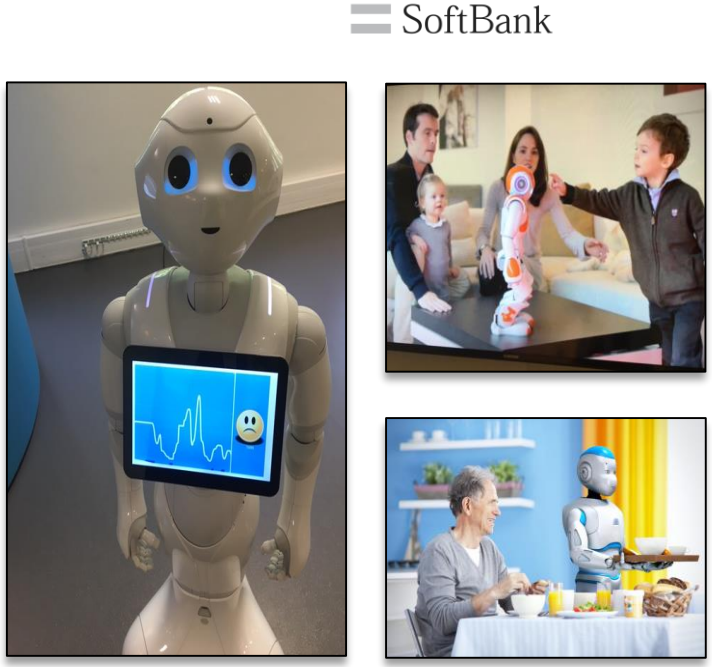
AI is imitation



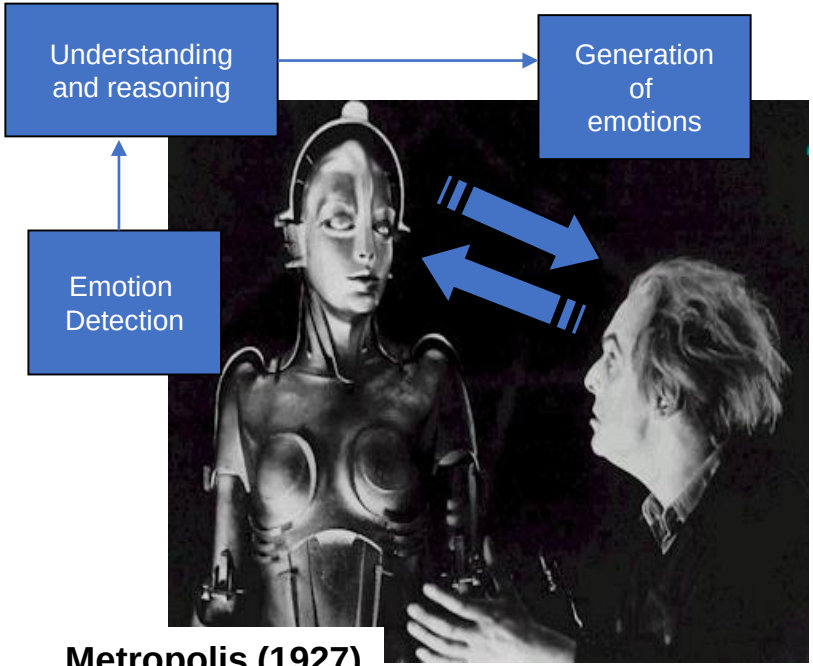
- Human language is an essential element in shaping, even determining, cultural characteristics, human perceptions, and even entire worldviews.
- **The language of conversational agents are devoid of lived experience, physically perception, feeling, or reasoning like a human. We creates a linguistic universe without bodily experience and understanding of meaning.**
- Thus, we converse with machines that can neither take responsibility for what they say nor be held accountable for it.
- Furthermore, tomorrow they could express themselves better than many of us... and seem more intelligent

Social robots at LISN

- We create social robots using conversational agents that can recognize, interpret, process and simulate human language and affect (even a kind of machine-humor)
- With the capacity of interpretation of the emotional state of humans, a robot can adapt his behavior and give an appropriate response to these emotions
- Naturally, it interacts differently with different individuals



Affective computing (R. Picard, 1997)



Metropolis (1927)
Fritz Lang

- Emotion is a key explanatory factor of behaviour
- We are not rational without being emotional (A. Damasio, "Descartes' error")

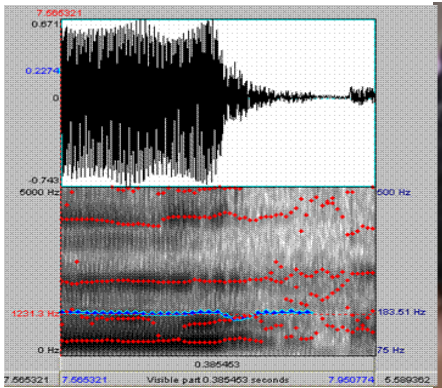
Emotions play a central role in social interaction

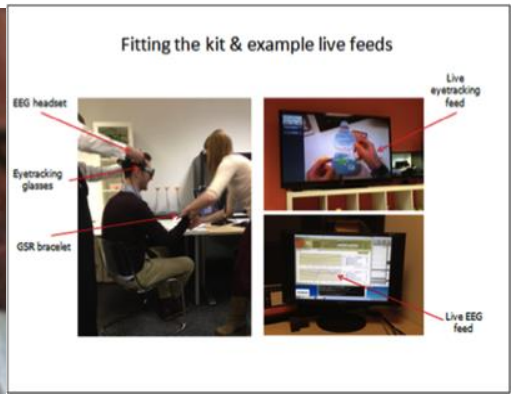
How we measure emotions?

Voice cues: Prosody : timber, energy, rythm, Affect bursts

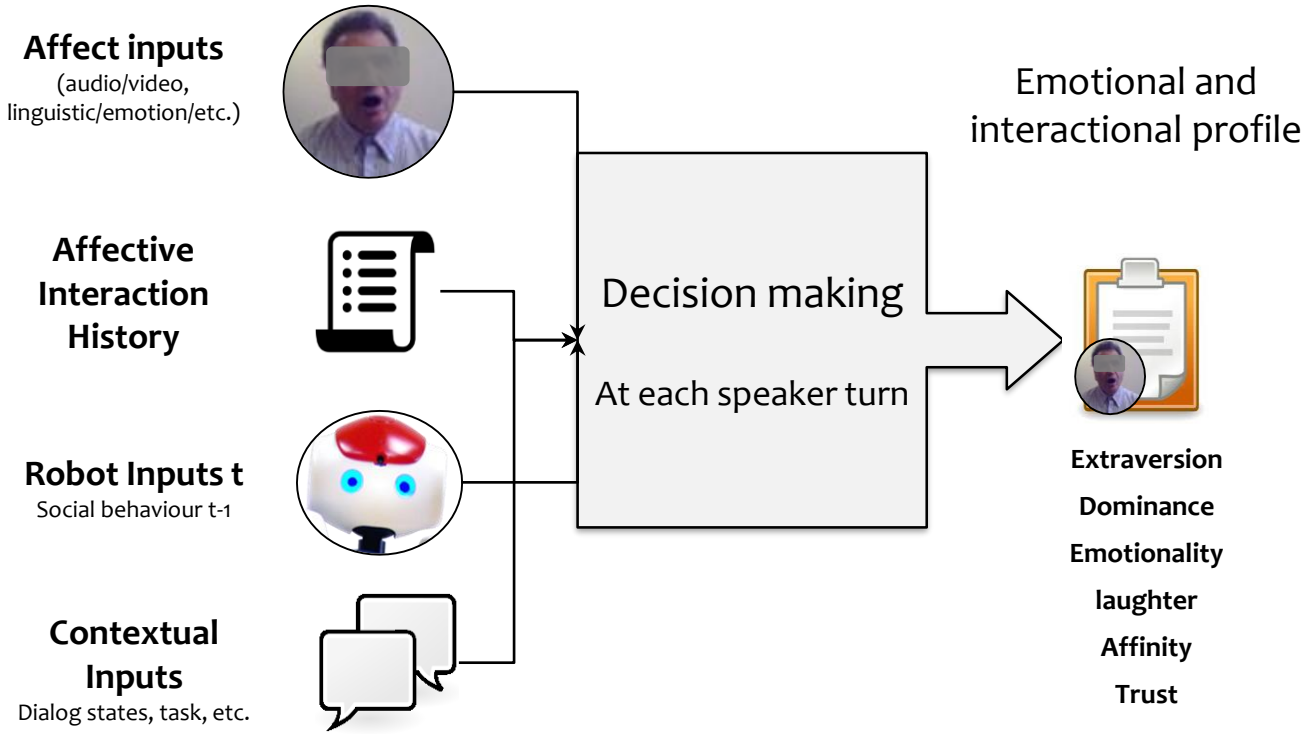
Visual cues: Facial expressions, Gaze tracking, Posture, Gestures, Movements, Behavior, Action

Physiological cues: Temperature, Respiration, Pupil dilation, Skin conductance ECG, EEG, Blood pressure, etc.





User model


Affect inputs
(audio/video, linguistic/emotion/etc.)
Affective Interaction History
Robot Inputs t
Social behaviour t-1
Contextual Inputs
Dialog states, task, etc.
Decision making
At each speaker turn
Emotional and interactional profile
Extraversion
Dominance
Emotionality
laughter
Affinity
Trust

Real tests
Broca Hospital (Paris, France)

Scenario: the robot negotiates for that the person drinks a glass of water. if the person does not want, the robots adapts its policies.

USC Institute for Creative Technologies

University of Southern California

SimSensei:
Virtual Human for Healthcare Support

& MultiSense:
Multimodal Perception and Learning

Albert (Skip) Rizzo, PI
Louis-Philippe Morency, PI

As part of DCAPS program:
Detection & Computational Analysis of Psychological Signals
(3rd Interim Progress Video)

The work depicted here was sponsored by the U.S Defense Advanced Research Projects Agency. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

Affective and social dimensions with chatbots /social robots

- Emerging interactive and adaptive systems using emotions modify how we will socialize with machines.
- More the dialog system is physically embodied, more it seems smart

Voice and appearance of (ro/chat)bots

Voice Assistants/ Social robots bring new dimensions (similar language, body movement or affective behavior) to the interaction and could become a mean of influencing or nudging individuals …

80% of the coders are men, 80% of the chatbots and social robots built have women names ALEXA, SAMANTHA, SOFIA, … or women face or body (Prometheus myth).

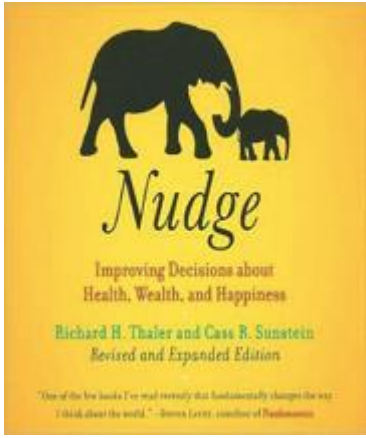
- Women characters/robots are assistants, nurses, sexual robots
- AI systems reinforce the stereotypes! It is also urgent to involve more women in AI and robotics.

2017: Nobel Prize in Economics - Richard Thaler

A nudge is a technique that consists in encouraging individuals to change their behavior without constraining them, using their cognitive biases: lacking self-control, acting emotionally, acting by conformity or laziness, etc.

Key insights for digital nudge enhanced by AI

- User-interface or socio-affective (ro/chat)bots design influences choices, even unintentionally
- Heuristics and biases influences online and off-line behavior
- Thorough design and testing can help achieve a designer's intended behavioral effects.



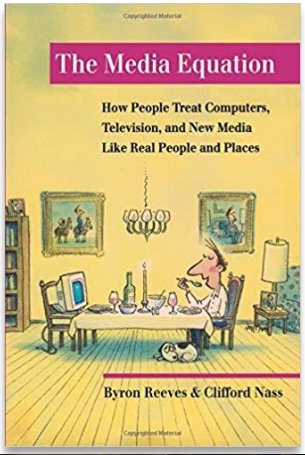
AI-enhanced nudging for Autistic Children

- Children are engaging in a creative play with the robot Moxie - drawing, story-telling, mindfulness practice, and more - that helps promote important life skills such as conversation, turn-taking, emotion regulation, and how to make friends.
- Moxie can learn to recognize a child's face and their particular learning needs for nudging.
- AI-enhanced nudging's persuasive power and intrusive capacity can also cause profound and long-lasting changes in the behaviour of users such as isolation, addiction, etc. especially with children and vulnerable people.

Can nudging be controlled in Extended Reality (XR)

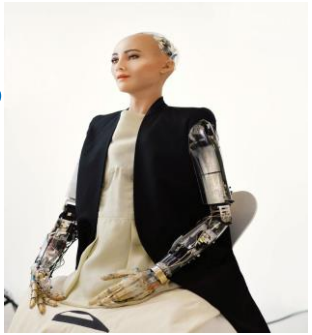
- Virtual reality "reunites" mother with dead daughter in South Korean doc.
- XR would open new ethical challenges such as emotional manipulation with confusion between reality and reality in XR (gray zone), relationship with digital objects…
- Other classical ethical principles and values also in XR: Privacy, Consent, Transparency, Responsibility, Dignity, Justice

What is the anthropomorphism?



- The attribution of human characteristics or behaviour to an object.
- The Media Equation (Reeves & Nass, 1996) explains that people tend to respond to media/computer/robot as they would either to another person by being polite, cooperative, attributing personality characteristics and even gender.
- So an object “which seems to be in the pain”, as the robot Atlas of Boston Dynamics, can inspire some empathy.

AI, Robotics and ethics



- Today, emerging interactive and adaptive systems using emotions modify how we will socialize with machines and maybe with humans
- These areas inspire critical questions centering on the ethics, the goals and the deployment of innovative products that can change our lives and society
- The issues surrounding AI machine behaviour cover a wide range of ethical topics: censorship of content, discrimination and unfairness, manipulation, marketing to children, changing our collective behaviours.

Private data – continuous monitoring
Hacking – cyber-security
Adaptive learning machine - control
Dehumanisation: isolation, dependance, addiction
Super-intelligent AI
More disparity in the society
Replacement of humans by machines in the work

Ethical issues of conversational agents

The CNPEN offers recommendations for governance, industry and citizens. A publication on the ethical issues of conversational agents was returned to the French Prime Minister on November 9 .

The Committee identified three main ethical issues regarding conversational agents: their status, the way in which they imitate language and emotions, and the public awareness of their effectiveness and limitations.

- List of recommendations, design principles, and research questions
- Examples: Vulnerable people, Work, the memory of the dead, Long-term effects

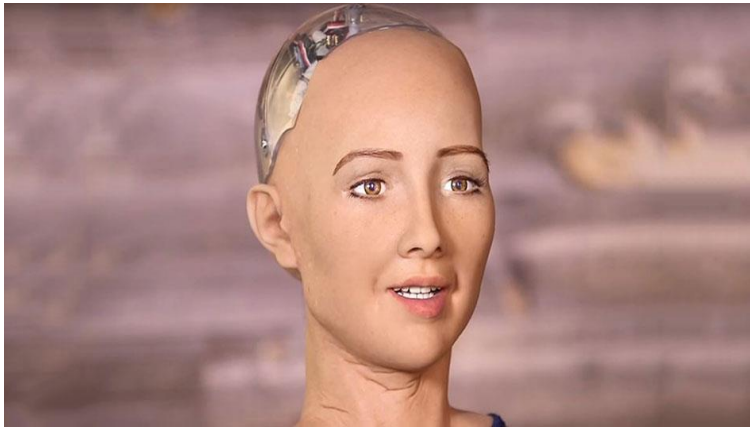
(L. Devillers, et al., CNPEN, 2021)

Ethical principles (EU experts)

Consider the level of trust in a chatbot/robot, its capabilities and limits and the capabilities and limits of the pair it forms with the user.

- Beneficence: promoting well-being, preserving dignity, and sustaining the planet
- Non-maleficence: privacy, security and “capability caution”
- Autonomy: the power to decide
- Justice: promoting prosperity and preserving solidarity
- Transparency and Explicability: enabling the other principles through intelligibility and accountability

Imitation of life



Issues

- What is a robot vs. a living being?
- Human identity vs. android
  Status of the robot in Human Society

Recommandations*

- Clarify the limits of imitation to avoid over attribution of capacities
- Keep a clear distinction between a living being and a machine (Law California 2019)

**Ethics in Robotics research, CERNA-Allistène (2014)*

Adaptive learning

TAY microsoft (2016)

It took less than 24 hours for Twitter to corrupt an innocent AI chatbot. « Tay » learnt from interaction with humans. It went from « humans are super cool » to full nazi in <24 hrs

Issues

- When the system continues to learn after deployment, its long-term behavior is difficult to control.
- Who is responsible if the machine malfunctions: the designer, the owner of the data, the owner of the system, its user.

Recommandations*

- The machine itself cannot be responsible.
- Users should be aware of the learning capacity of the machine. It can lead to new problems that affect the consent of both user and society.
- Machine should be controlled with benchmarks several times during the usage of the machine.
- Researchers should seek to contribute to societal debates and to the development of assessment benchmarks and protocols for broad dissemination of machine learning systems.
- For use in specialized professional sectors (medicine, law, transportation, energy, etc.), data collection and analysis requires collaboration between computer scientists and experts in those fields.

**Ethics in Machine learning, CERNA-Allistène (2018)*

Methodology for impact assessment, risks mitigation and metrics designation

Medical analogy

- like in the medical sector, AI providers can set up processes to ensure that serious incidents are identified and assessed. This enables developers and AI designers to be alerted to signals emerging across the world, which might not be obvious from local data alone

Metrics designation

- identify processes and key indicators, both horizontally (by industry and sectors) and vertically (by applications and technologies) supporting AI-enhanced Nudges requirements in order to develop conformity assessment methodologies and tools.

Conclusions

- The issues surrounding these Human (Ro)bot interaction cover a wide range of ethical topics: censorship of content, discrimination and unfairness, manipulation, changing our collective behaviours and culture.
- How will humans co-evolve, co-create, co-learn with these social (ro)bots or in the metaverse (work, education, health)? In particular, how will vulnerable people be protected from potential threats: addiction, manipulation, isolation?
- A lot of use cases can lead at ethical positive nudge such as areas of health, consumer goods and services, and green energy. But nudging poses an ethical threat and negative business implications.
- Social (ro)bots, AI-enhanced nudge raise ethical questions to which no standard today provides direct answers. We must define standards and ethical frameworks (soft-law) for conversational systems (chatbots, social robotics) and educate multi-disciplinary and multi-stakeholder.

Contact: devil@limsi.fr

- I am now convenor of a WG 4 on Foundational and societal aspects of AI at AFNOR-JTC21 CEN-CENELEC (12 countries).
- JTC 21 produces standardization deliverables in the field of AI and related use of data to address European market and societal needs