UiO : Department of informatics
University of Oslo

Ethical challenges and opportunities with (care) robots

Universal Design as an ethical charter for inclusive robot design and interaction

Diana Saplacan, Researcher
Robotics and Intelligent Systems
Department of Informatics, University of Oslo
Oslo, Norway

- *diana.saplacan at ifi.uio.no*
- *Linkedin: /dianasaplacan*

Imagine how it would be if different types of users of Social and Assistive Robots (SARs) would be able to use the same robot

Photo taken from Bing

Persona

Eve, 85 years old, lives at home on her own

Image source: https://upload.wikimedia.org/wikipedia/commons/thumb/3/31/Noto_Emoji_Oreo_1f475.svg/768px-Noto_Emoji_Oreo_1f475.svg.png

UiO : Institutt for informatikk

Possible scenarios with social and assistive robots used within home- and healthcare

1. Informed consent

2. Other people around

Photo credit: Diana Saplacan

Photo credit: Vegard Søyseth

robot look vs. robot capabilities vs. security vs. privacy vs. safety vs. context of use

The problem of the robots that may look *too* cute – may also lead to you trusting them more

*”But then, again, to bring it into robotics world and to physical robots, **I think that’s equally important to look at the diversity of robots that are made, and also how they have been made to look, you know. And you know, who they are accomodating to. *cause I think that’s interesting to see as well, when you look at the different robot designs** (…) In what cultures are they made, you know, how? Where are they made to look like they look? And of course, maybe that is one of the problems with the more humanoid robots in who’s image are the humanoid robots made, and they are not necessarily very accomodating in two different cultures and ethnicities. So, that is, definirely, also I think an interesting question, when it comes to design, and I think we see some interesting robot design as well.”*

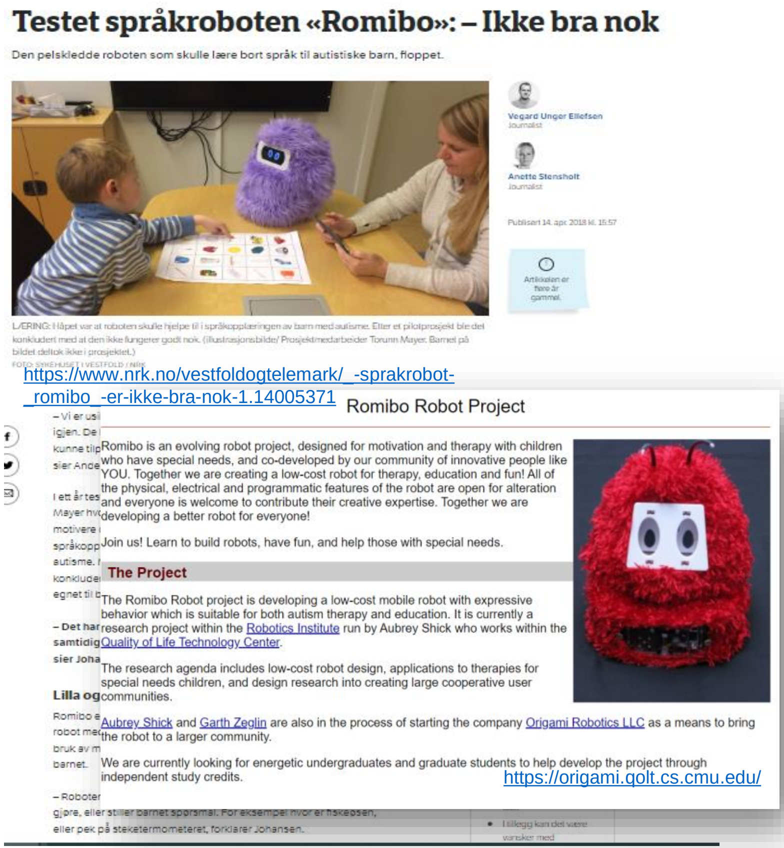
”I always show that Japanese Lovot, which is so cute, I mean, How they have, you know, tried to explore design, tried to make the robots more, you know, cuter (in japanese Kawaii). I think we should bear in mind what happens with using this concept on our work – it is because of the robots and how we see them, for example if we make them too cute.

*”**it’s a little menacing that you would probably allow it to do anything you know.** So, that is also one side I think about the design, but at the same time, it’s probably, you know, like it’s with the PARO as well, it’s more likeable for more people.”*

*”Yeah, but this can also lead to deception, because it might be very menanacing, a **very menancing robot which looks cute, and you trust it, you know, it increases your trust towards it. It’s more, not that it will hurt you or something, but it might collect data that you don’t want to be collected.”***

Robot look vs. robot capabilities vs. privacy vs. transparency

Examples from Saplacan et al. (to be submitted, 2022). Robots and Boundary Objects: A Study from the Nordic Context presenting experts’ views on Social and Assistive Robots in home- and healthcare. Book chapter in Cambridge Handbook on Law, Policy, and Regulations for HRI. Cambridge University Press.



Testet språkroboten «Romibo»: – Ikke bra nok

Den pelskledde roboten som skulle lære bort språk til autistiske barn, floppet.

Vegard Unger Ellefsen
Journalist

Anette Stensholt
Journalist

Publisert 14. apr. 2018 kl. 15:57

Artikkelen er flere år gammel.

LÆRING: Håpet var at roboten skulle hjelpe til i språkopplæringen av barn med autisme. Etter et pilotprosjekt ble det konkludert med at den ikke fungerer godt nok. (illustrasjonsbilde/ Prosjektmedarbeider Torunn Mayer. Barnet på bildet deltok ikke i prosjektet.)

FOTO: SYKEHUSET I VESTFOLD / NRK

– Roboten gjøre, eller stiller barnet spørsmål. For eksempel hvor er fiskeposen, eller pek på steketermometeret, forklarer Johansen.

https://www.nrk.no/vestfoldogtelemark/_-sprakrobot-_romibo_-er-ikke-bra-nok-1.14005371

Romibo Robot Project

Romibo is an evolving robot project, designed for motivation and therapy with children who have special needs, and co-developed by our community of innovative people like YOU. Together we are creating a low-cost robot for therapy, education and fun! All of the physical, electrical and programmatic features of the robot are open for alteration and everyone is welcome to contribute their creative expertise. Together we are developing a better robot for everyone!

Join us! Learn to build robots, have fun, and help those with special needs.

The Project

The Romibo Robot project is developing a low-cost mobile robot with expressive behavior which is suitable for both autism therapy and education. It is currently a research project within the Robotics Institute run by Aubrey Shick who works within the Quality of Life Technology Center.

The research agenda includes low-cost robot design, applications to therapies for special needs children, and design research into creating large cooperative user communities.

Aubrey Shick and Garth Zeglin are also in the process of starting the company Origami Robotics LLC as a means to bring the robot to a larger community.

We are currently looking for energetic undergraduates and graduate students to help develop the project through independent study credits.

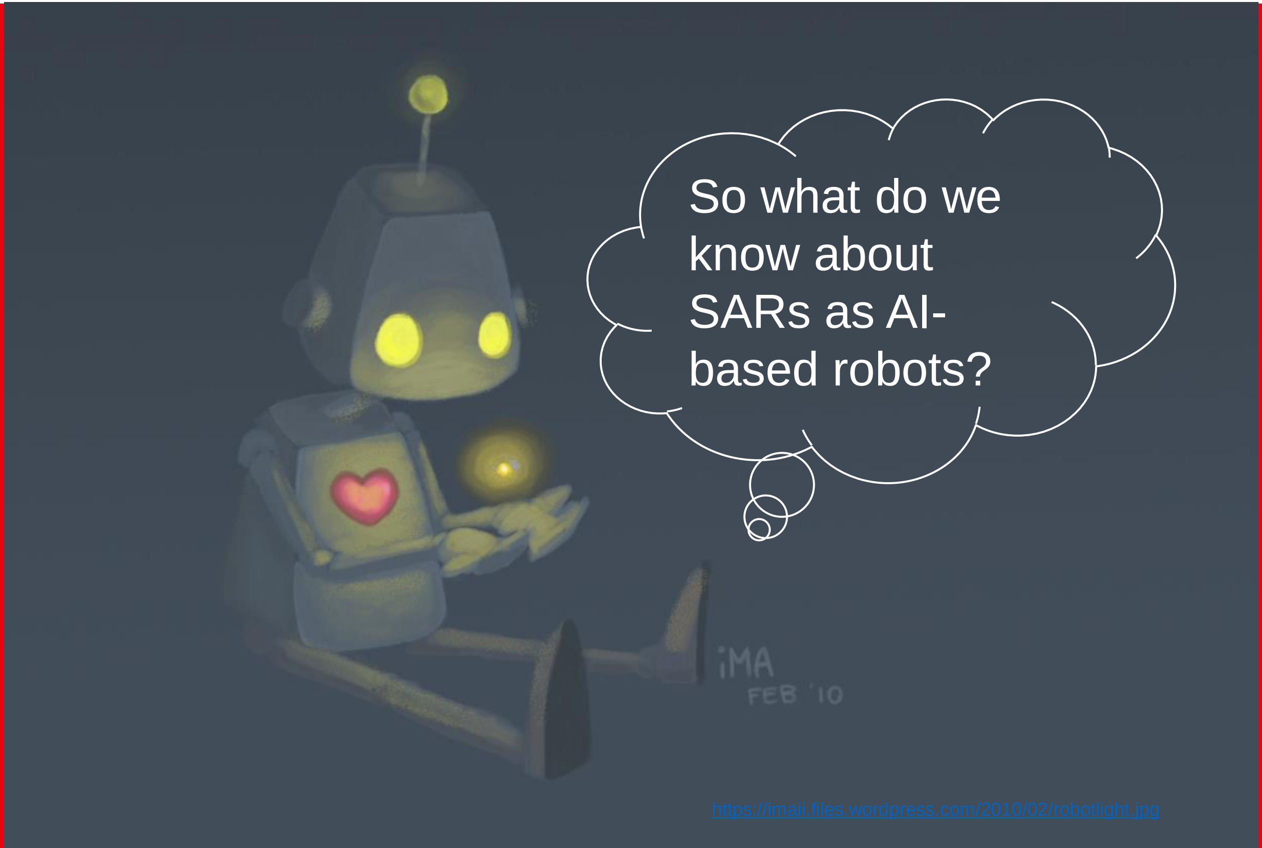
https://origami.qolt.cs.cmu.edu/

Examples from Saplacan et al. (submitted, 2022). Robots and Boundary Objects: A Study from the Nordic Context presenting experts' views on Social and Assistive Robots in home- and healthcare. Book chapter in Cambridge Handbook on Law, Policy, and Regulations for HRI. Cambridge University Press.

Toys as connected devices used in therapy sessions

- Concrete example: a connected device (e.g., a robot, **Romibo**) is **marketed and sold by the manufacturer as a toy** – but it is, **in practice used as a medical device** (e.g., with people with Autism)
- The robot was recognized as a medical device in U.S., and as a toy in Europe - according to an informant
- The manufacturer/producer does not conform with all the health/data standards regarding privacy, confidentiality, safety – only with the Toy Directive
- **The take-away points:**
  - The design issue: the robot does not have a mouth – the children with Autism do not know where the sound comes from
  - The manufacturer has currently the power to decide how the connected device should be sold. This has implications on the user and the user data.
  - If a robot is assessed as a medical device, a medical device should not be used all the time (e.g., mental health issue)

robot look vs. robot capabilities vs. accountability vs. how the robot is used in practice



So what do we know about SARs as AI-based robots?

https://imaii.files.wordpress.com/2010/02/robotlight.jpg

Identified Ethical Challenges with Social and Assistive Care Robots

1) **Lack of legal framework and harmonized standards regulating AI and (care) robots**
2) **Decreased human contact**
3) **The elderly feeling objectified and that they lose control (over their data)**
4) **The elderly perceiving that their privacy is lost**
5) **The elderly feeling deception and infantilization**
6) **The elderly's concern on who is responsible**

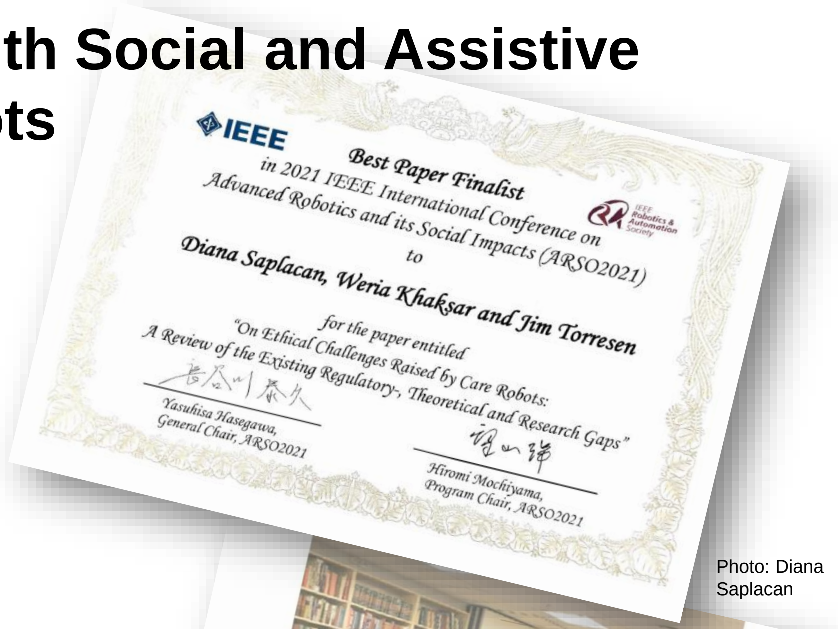


Photo: Diana Saplacan

Photo: MECS

Saplacan, D.; Khaksar, W.; Torresen, J. (2021), [On Ethical Challenges Raised by Care Robots: A Review of the Existing Regulatory-, Theoretical-, and Research Gaps](https://ieeexplore.ieee.org/document/9542844), In *Proceedings of the 2021 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO) Virtual Conference, July 8-10, 2021*. IEEE Robotics and Automation Society. ISBN 978-1-6654-4952-6. Paper 38. s 219 - 226, https://ieeexplore.ieee.org/document/9542844

Identified Opportunities with Social and Assistive Care Robots

- Robots can help with routine, manual and repetitive tasks
- Robots may transport medicine
- Robots may balance the feeling of loneliness and anxiety
- Robots may allow the care-receiver to interact remotely with care-takers
- Robots may save time in the rehabilitation process
- Robots may provide accurate and real-time health data
- Robots may prevent falls
- Robots may provide reminders, or help with household chores

- Conclusion: the benefits may **outweight** some of the ethical concerns or challenges

Photo: Diana Saplacan

Photo: MECS

Saplacan, D.; Khaksar, W.; Torresen, J. (2021), On Ethical Challenges Raised by Care Robots: A Review of the Existing Regulatory-, Theoretical-, and Research Gaps, In *Proceedings of the 2021 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO) Virtual Conference, July 8-10, 2021*. IEEE Robotics and Automation Society. ISBN 978-1-6654-4952-6. Paper 38. s 219 - 226, https://ieeexplore.ieee.org/document/9542844

Robots within home- and healthcare

- Robots as welfare technologies and how social and assistive robots used in home- and healthcare challenge the notion of welfare

- Privacy, security, and safety issues of robots to be used with vulnerable users

- The challenge of introducing social and assistive robots within home- and healthcare services

D. Saplacan and J. Torresen, "Healthcare Professionals´ Attitudes towards the Organization of Care Services and the Adoption of Welfare Robots in Norway," *2022 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO)*, 2022, pp. 1-8, doi: 10.1109/ARSO54254.2022.9802979.

Diana Saplacan and Jim Tørresen. 2022. Robots as Welfare Technologies to Reduce Falls Amongst Older Adults: An Explorative Study from Norway. In Human Aspects of IT for the Aged Population. Technology in Everyday Living: 8th International Conference, ITAP 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26 – July 1, 2022, Proceedings, Part II. Springer-Verlag, Berlin, Heidelberg, 88–106. https://doi.org/10.1007/978-3-031-05654-3_6

What we also know: AI and Human Rights

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Pr…

PRINCIPLED ARTIFICIAL INTELLIGENCE:

Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI

Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Christopher Nagy, Madhulika Srikumar

BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

C. Persons with disabilities

Even though more than 650 million people worldwide have a disability of one form or another (two thirds of whom live in developing countries), most have long been neglected and marginalized by the State and society. It is only in recent years that persons with disabilities have brought about a paradigm shift in attitudes towards them. This has seen a move away from regarding them as "objects" of charity and medical interventions towards their empowerment as "subjects" of human rights, including but not limited to the right to health.

The right to health of persons with disabilities cannot be achieved in isolation. It is closely linked to non-discrimination and other principles of individual autonomy, participation and social inclusion, respect for difference, accessibility, as well as equality of opportunity and respect for the evolving capacities of children.[17]

Persons with disabilities face various challenges to the enjoyment of their right to health. For example, persons with physical disabilities often have difficulties accessing health care, especially in rural areas, slums and suburban settings; persons with psychosocial disabilities may not have access to affordable treatment through the public health system; women with disabilities may not receive gender-sensitive health services. Medical practitioners sometimes treat persons with disabilities as objects of treatment rather than rights-holders and do not always seek their free and informed consent when it comes to treatments. Such a situation is not only degrading, it is a violation of human rights under the Convention

[17] These and other principles are reflected in art. 3 of the Convention on the Rights of Persons with Disabilities, which was adopted by the United Nations General Assembly in its resolution 61/106 of 13 December 2006.

D. How does the principle of non-discrimination apply to the right to health?

Discrimination means any distinction, exclusion or restriction made on the basis of various grounds which has the effect or purpose of impairing or nullifying the recognition, enjoyment or exercise of human rights and fundamental freedoms. It is linked to the marginalization of specific population groups and is generally at the root of fundamental structural inequalities in society. This, in turn, may make these groups more vulnerable to poverty and ill health. Not surprisingly, traditionally discriminated and marginalized groups often bear a disproportionate share of health problems. …s have shown that, in some societies, ethnic minority …ous peoples enjoy fewer health services, receive less …nd are less likely to have adequate housing and safe … their children have a higher mortality rate and suffer …tion than the general population.

…mination is compounded when an individual suffers … discrimination, such as discrimination on the basis … national origin or age. For example, in many places … receive fewer health and reproductive services and … more vulnerable to physical and sexual violence than …on.

…and equality are fundamental human rights principles …ents of the right to health. The International Covenant … and Cultural Rights (art. 2 (2)) and the Convention … Child (art. 2 (1)) identify the following non-exhaustive …nation: race, colour, sex, language, religion, political … national or social origin, property, disability, birth …cording to the Committee on Economic, Social and …er status" may include health status (e.g., HIV/AIDS) … States have an obligation to prohibit and eliminate … grounds and ensure equality to all in relation to access …e underlying determinants of health. The International …limination of All Forms of Racial Discrimination (art. 5) …ates must prohibit and eliminate racial discrimination …ight of everyone to public health and medical care.

…and equality further imply that States must recognize …e differences and specific needs of groups that …icular health challenges, such as higher mortality … to specific diseases. The obligation to ensure non-…es specific health standards to be applied to particular

7

Office of the United Nations High Commissioner for Human Rights

World Health Organization

The Right to Health

Fact Sheet No. 31

…eference to human rights, …ework for their overall …on of human rights may …t to bear on individuals … of standards like "privacy" …r complex situations in which separate principles …on with one another.

Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar. "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI." Berkman Klein Center for Internet & Society, 2020. Available at: http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420 | WHO, "The Right to Health," Office of the United Nations High Commissioner for Human Rights, Fact Sheet No. 31. [Online]. Available: https://www.ohchr.org/documents/publications/factsheet31.pdf

What we also know: Themes amongst AI principles

- Privacy
- Accountability
- Safety and security
- Transparency and explainability
- Fairness and non-discrimination
- Human control of technology
- Professional responsibility
- Promotion of human values

3.8. Promotion of Human Values

With the potential of AI to act as a force multiplier for any system in which it is employed, the Promotion of Human Values is a key element of ethical and rights-respecting AI.[304] The principles under this theme recognize that the ends to which AI is devoted, and the means by which it is implemented, should correspond with and be strongly influenced by social norms. As AI's use becomes more prevalent and the power of the technology increases, particularly if we begin to approach artificial general intelligence, the imposition of human priorities and judgment on AI is especially crucial. The Promotion of Human Values category consists of three principles:

PRINCIPLES UNDER THIS THEME

- 44% Human Values and Human Flourishing
- 31% Access to Technology
- 64% Leveraged to Benefit Society

3.5. Fairness and Non-discriminatio

Algorithmic bias – the systemic under- or over-prediction of probabilities for a specific population – creeps into AI systems in a myriad of ways. A system might be trained on unrepresentative, flawed, or biased data.[217] Alternatively, the predicted outcome may be an imperfect proxy for the true outcome of interest[218] or the outcome of interest may be influenced by earlier decisions that are themselves biased. As AI systems increasingly inform or dictate decisions, particularly in sensitive contexts where bias long predates their introduction such as lending, healthcare, and criminal justice, ensuring fairness and non-discrimination is imperative. Consequently, the Fairness and Non-discrimination theme is the most highly represented theme in our dataset, with every document referencing at least one of its six principles: "non-discrimination and the prevention of bias," "representative and high-quality data," "fairness," "equality," "inclusiveness in impact," and "inclusiveness in design."[219]

Within this theme, many documents point to biased data – and the biased algorithms it generates – as the source of discrimination and unfairness in AI, but a few also recognize the role of human systems and institutions in perpetuating or preventing discriminatory or otherwise harmful impacts. Examples of language that focuses on the technical side of bias include the Ground Rules for AI conference paper ("[c]ompanies should strive to avoid bias in A.I. by drawing on diverse data sets")[220] and the Chinese White Paper on AI Standardization ("we should also

PRINCIPLES UNDER THIS THEME

- 89% Non-discrimination and the Prevention of Bias
- 36% Representative and High Q
- 56% Fairness
- 25% Equality
- 42% Inclusiveness in Impact
- 47% Inclusiveness in Design

Percentage reflects the number of docum the dataset that include each principle

Inclusiveness in Impact

"Inclusiveness in impact" as a principle calls for a just distribution of AI's benefits, particularly to populations that have historically been excluded. There was remarkable consensus in the language that documents employed to reflect this principle, including concepts like "shared benefits" and "empowerment":

Document	Language of principle
Asilomar AI Principles	Shared Benefit: AI technologies should benefit and empower as many people as possible.[242]
Microsoft's AI principles	Inclusiveness – AI systems should empower everyone and engage people. If we are to ensure that AI technologies benefit and empower everyone, they must incorporate and address a broad range of human needs and experiences. Inclusive design practices will help system developers understand and address potential barriers in a product or environment that could unintentionally exclude people. This means that AI systems should be designed to understand the context, needs and expectations of the people who use them.[243]
Partnership on AI Tenets	We will seek to ensure that AI technologies benefit and empower as many people as possible[244]
Smart Dubai AI principles	We will share the benefits of AI throughout society: AI should improve society, and society should be consulted in a representative fashion to inform the development of AI[245]
T20 report on the future of work and education	Benefits should be shared: AI should benefit as many people as possible. Access to AI technologies should be open to all countries. The wealth created by AI should benefit workers and society as a whole as well as the innovators.[246]
UNI Global Union's AI principles	Share the Benefits of AI Systems: AI technologies should benefit and empower as many people as possible. The economic prosperity created by AI should be distributed broadly and equally, to benefit all of humanity.[247]

The European High Level Expert Group guidelines add some detail around what "benefits" might be shared: "AI systems can contribute to wellbeing by seeking achievement of a fair, inclusive and peaceful society, by helping to increase citizen's mental autonomy, with equal distribution of economic, social and political opportunity."[248] There is a clear connection to the principles we have catalogued under the Promotion of Human Values theme, especially the principle of "leveraged to benefit society."

4. International Human Rights

In recent years, the human rights community has become more engaged with digital rights, and with the impacts of AI technology in particular. Even outside of human rights specialists, there has been an increasing appreciation for the relevance of international human rights law and standards to the governance of artificial intelligence.[327] To an area of technology governance that is slippery and fast-moving, human rights law offers an appealingly well-established core set of concepts, against which emerging technologies can be judged. To the broad guarantees of human rights law, principles documents offer a tailored vision of the specific – and in some cases potentially novel – concerns that AI raises.

Accordingly, when coding the principles documents in our dataset, we also made observations on each document's references to human rights, whether as a general concept or specific human-rights related documents such as the Universal Declaration of Human Rights, International Covenant on Civil and Political Rights, the United Nations Guiding Principles on Business & Human Rights and the United Nations Sustainable Development Goals. Twenty-three of the documents in our dataset (64%) made a reference of this kind. We also noted when documents stated explicitly that they had employed a human rights framework, and five of the thirty-six documents (14%) did so.

Given the increasing visibility of AI in the human rights community and the apparent increasing interest in human rights among those invested in AI governance, we had expected that the data might reveal a trend toward increasing emphasis on human rights in AI principles documents. However, our dataset was small enough, and the

Nature of actor	Number of documents	Number with any reference to human rights	
Civil society	5	4	80%
Government	13	6	46%
Intergovernmental organization	3	2	67%
Multistakeholder initiative	7	4	57%
Private sector	8	7	88%
Total	36	23	64%

There are multiple possible explanations for this. It may be that the agencies or individuals in government who have been tasked with drafting and contributing to principles documents were not selected for their expertise with human rights law, or it may be that national laws, such as the GDPR, are perceived as more relevant.

The documents also exhibit significant variation in the degree to which they are permeated by human rights law, with some using it as the framework of the whole document (denoted by a star in the data visualization), and others merely mentioning it in passing (denoted by a diamond). Using a human rights framework means that the document uses human rights as a basis for further ethical principle for the development and use of AI systems. Only five documents use a human rights framework. Three are civil society documents and two are government documents from the EU: Access Now report, AI for Europe, European High Level Expert Group guidelines, Public Voice Coalition AI guidelines, and Toronto Declaration.

Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar. "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI." Berkman Klein Center for Internet & Society, 2020. Available at: http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420

What we also know: Ethical guidelines: Ethics of AI (UNESCO)

unesco

Distribution: limited

SHS/IGM-AIETHICS/2021/JUN/3 Rev.2
25 June 2021
Original: English and French

DRAFT TEXT OF THE RECOMMENDATION ON THE ETHICS OF ARTIFICIAL INTELLIGENCE

Pursuant to the 40 C/Resolution 37, and in accordance with the UNESCO Constitution and the Rules of Procedure concerning recommendations to Member States and international conventions covered by the terms of Article IV, paragraph 4, of the Constitution, the Director-General of UNESCO convened an Ad Hoc Expert Group (AHEG) for the preparation of a draft text of a Recommendation on the Ethics of Artificial Intelligence, and submitted the draft text of the Recommendation to the special committee meeting of technical and legal experts, designated by Member States.

The special committee meeting was held in two phases, the first from 26 to 30 April 2021 and the second from 21 to 25 June 2021. Intersessional consultations were also organized in the period from 1 to 18 June 2021 (12 days). The special committee meeting revised the draft Recommendation and approved the present text for the submission to the General Conference at its 41st session for adoption.

VALUES:

- **Respect, protection and promotion of human rights and fundamental freedoms and human dignity**
- **Environment and ecosystem flourishing**
- **Ensuring diversity and inclusiveness**
- **Living in peaceful, just and interconnected societies**

Ensuring diversity and inclusiveness

19. Respect, protection and promotion of diversity and inclusiveness should be ensured throughout the life cycle of AI systems, consistent with international law, including human rights law. This may be done by promoting active participation of all individuals or groups regardless of race, colour, descent, gender, age, language, religion, political opinion, national origin, ethnic origin, social origin, economic or social condition of birth, or disability and any other grounds.

20. The scope of lifestyle choices, beliefs, opinions, expressions or personal experiences, including the optional use of AI systems and the co-design of these architectures should not be restricted during any phase of the life cycle of AI systems.

21. Furthermore, efforts, including international cooperation, should be made to overcome, and never take advantage of, the lack of necessary technological infrastructure, education and skills, as well as legal frameworks, particularly in LMICs, LDCs, LLDCs and SIDS, affecting communities.

PRINCIPLES

- **Proportionality and Do No Harm**
- **Safety and security**
- **Fairness and non-discrimination**
- **Sustainability**
- **Right to Privacy, and Data Protection**
- **Human oversight and determination**
- **Transparency and explainability**
- **Responsibility and accountability**
- **Awareness and literacy**
- **Multi-stakeholder and adaptive governance and collaboration**

UNESCO Digital Library, "Draft text of the Recommendation on the Ethics of Artificial Intelligence," 2021. https://unesdoc.unesco.org/ark:/48223/pf0000377897 (accessed Aug. 10, 2021).

What we also know:
Directives:
Universal Design – European Accessibility Act (EU 2019/882)

- European Accessibility Act: A directive on accessibility requirements
- Aims to a well-functioning of the European ”internal market for accessible products and services, by removing barriers created by divergent rules in Member States”

Businesses will benefit from:

- common rules on accessibility in the EU leading to costs reduction
- easier cross-border trading
- more market opportunities for their accessible products and services

Persons with disabilities and elderly people will benefit from:

- more accessible products and services in the market
- accessible products and services at more competitive prices
- fewer barriers when accessing transport, education and the open labour market
- more jobs available where accessibility expertise is needed

7.6.2019 | EN | Official Journal of the European Union | L 151/70

DIRECTIVE (EU) 2019/882 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 17 April 2019

on the accessibility requirements for products and services

(Text with EEA relevance)

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Accessibility standardisation

Common European accessibility standards help remove barriers for people with disabilities and others (e.g. the elderly). When applied across Member States, these standards also improve the functioning of the internal market, by removing barriers to free movement of goods and services.

Actions

The Commission has instructed European standards organisations, which include CEN, CENELEC and ETSI, to develop and implement accessibility standards. These include standards for

- ICT accessibility resulting in European Standard EN 301 549
- accessibility to the built environment, leading to European Standard EN 17210, which is currently in the process of consultation
- accessibility following “Design for all” standards, resulting in European Standard EN 17161
- accessibility of websites and mobile applications, updating European Standard EN 301 549

Key EU legislative instruments (the directive on web accessibility, the European accessibility act, the public procurement directives) refer to the possible use of accessibility standards. The Commission encourages the participation of all relevant stakeholders in these processes, including persons with disabilities.

https://ec.europa.eu/social/main.jsp?catId=1202
https://ec.europa.eu/social/main.jsp?catId=1485&langId=en

What we also know: Ethical Values in AI based technologies

- **On Values:**
  - Value is typically associated with what is “good” or “desirable.”
  - Rather than being descriptive, values are normative and express what is “good.”
  - EU High-Level Expert Group on AI + IEEE formulated ethical principles and moral values that should be adhered to design and deployment of AI

1. Respect for autonomy
2. Non-maleficence
3. Fairness
4. Transparency
5. Explainability
6. Accountability.

- Value-Sensitive Design (VSD) (see Friedman, 1996)

Ethics in AI:

Some of these papers talk whether we can embed ethics and values in AI and how.For instance, Dignum et al. (2018) propose the idea of:

- **Ethics *by* design**: designing technologies with ethics in mind.
- **Ethics *for* design:** following standards, legal aspects, rules etc.
- **Ethics *in* design**: how we think about design of algorithms, AI, technologies: what are our moral theories, ethical theories. Often Values Sensitive Design is discussed.

- AI systems: technical artifacts, human agents, and institutions, artificial agents and certain technical norms that regulate interactions between artificial agents and other elements of the system.

Friedman, Batya. 1996. “Value-Sensitive Design.” ACM Interactions, 1996. https://old.vsdesign.org/publications/pdf/friedman96valuesensitivedesign.pdf. │Friedman, Batya, and David Hendry. 2019. Value Sensitive Design: Shaping Technology with Moral Imagination. Cambridge, Massachusetts: The MIT Press. │Dignum, Virginia. 2018. “Ethics in Artificial Intelligence: Introduction to the Special Issue.” *Ethics and Information Technology* 20 (1): 1–3. https://doi.org/10.1007/s10676-018-9450-z. │ Dignum, Virginia, Matteo Baldoni, Cristina Baroglio, Maurizio Caon, Raja Chatila, Louise Dennis, Gonzalo Génova, et al. 2018. “Ethics by Design: Necessity or Curse?” In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 60–66. AIES ’18. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3278721.3278745.

This means that we should embed values in the design of robots

- Respect, protection, and promotion of human rights
- Ensuring diversity and inclusiveness

- Ensure accessibility of SARs for diversity of users
- Operationalize the values to be embedded in the design of SARs robots

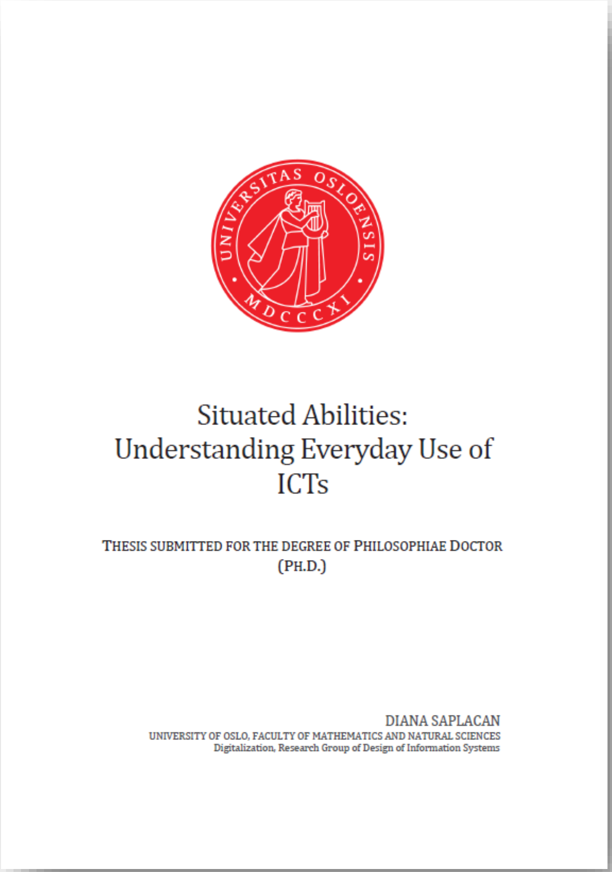
But HOW to do this?

How to implement values in the design and interaction of robots?
How to let values guide our design of SARs?

What we also know:
About abilities/disabilities and robots

- On situated abilities as a salutogenic (as opposed to pathogenic approach) focusing on individuals' low or high abilities on an ability spectrum
- Promotes the idea of seeing our selves in relation to others

"Kittay (2011) argues that human beings form dependencies relationships with others at various stages in their lives, where they may depend more or less on others. She argues that **being less abled is an inherent characteristic of humans** (Kittay 2011). She states: "From this perspective, we reason that **our societies should be structured to accommodate inevitable dependency within a dignified, flourishing life – both for the cared for and for the carer**. Finally, if we see ourselves as always **selves-in-relation**, we understand that our sense of **well-being is tied to the adequate care and well-being of another. Caregiving work is the realization of this conception of self, both when we give care generously and when we receive it graciously**" (Kittay 2011, p. 54, in Saplacan, 2020)



Situated Abilities:
Understanding Everyday Use of ICTs

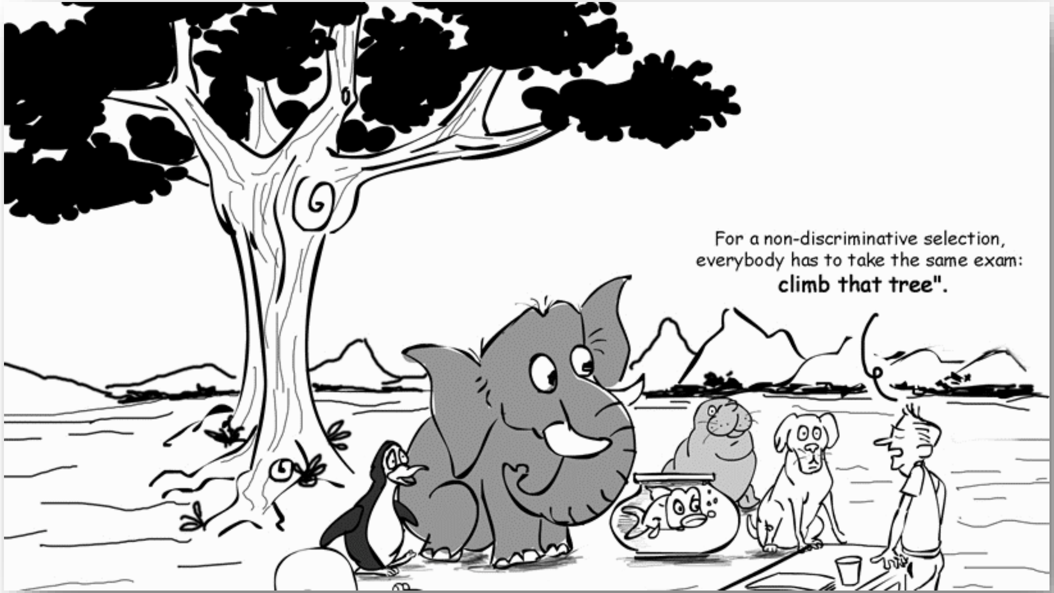
THESIS SUBMITTED FOR THE DEGREE OF PHILOSOPHIAE DOCTOR (PH.D.)

DIANA SAPLACAN
UNIVERSITY OF OSLO, FACULTY OF MATHEMATICS AND NATURAL SCIENCES
Digitalization, Research Group of Design of Information Systems

Saplacan, D. (2020). *Situated Abilities: Understanding Everyday Use of ICTs*. PhD Thesis. University of Oslo.

Proposing a Universal Design framework for Design of SARs
Universal Design: Definition and History

- **“The design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design.”** (R. Mace, 1941-1998)



The 7 principles of Universal Design (UD)

#	Principle	Example
1	Equitable use	Use of a ramp for getting into a bus: it provides equal ability to step onto a bus for both people in a wheelchair and without a wheelchair, such as a woman with a stroller
2	Flexibility in use	The use of a table with an adjustable height is good for both abled people, people with back problems, people sitting in wheelchairs, or children
3	Simple and intuitive use	An iconic example is the iPhone design with its buttons in the same place in different versions.
4	Perceptible information	Consistency in using symbols for volume or radio buttons, send icons, or save icons on buttons.
5	Tolerance for error	The undo button provides reliable feedback. Another example is the oven lock button for children's safety.
6	Low physical effort	The height of ATMs provides easy access and low physical effort for people of different heights, including children and people sitting in a wheelchair
7	Size and space for approach and use	The gates of a metro-station or security control at the airport should be large enough to accommodate individuals of different sizes, or people sitting in a wheelchair

On Universal Design of robots – considering their verbal and non-verbal communication

- **how they *appear* (LOOK)**
  - **a robot should not have any physical design that is not useful**, e.g., the robot should not have any arms and fingers if the robot will not be used to pick up things; e.g., a robot to be used in therapy session with autistic children should always be equipped with a mouth
- **how they *move* (MOVEMENT)**
  - **a robot should move according to the social and cultural norms of the context of use**, e.g. nodding head as a yes, in the Indian culture, respecting the distance between itself and the human the robot interact with depending on the culture – see proxemics (Hall, 1966);
- **how they *interact* (BEHAVE)** with us humans, including people with different (situated) abilities
  - **the robot should be able to behave apropiately together with the specific users it interact with**, e.g. see the Romibo example; a robot used with deaf people should be able to use sign language; a robot should be able to interact according to the digital litearcy level and ways of interaction, the verbal and non-vebral language used by the users (e.g, specific to the elderly, children, adults with high or low digital literacy etc.). The robot's ways of interaction should be adaptable to the users.

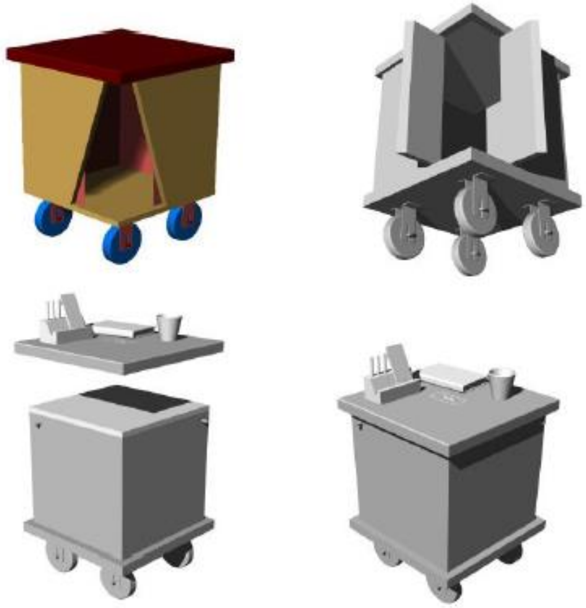
Universal Design (UD) and guidelines

#	UD Principle	Guidelines
1	Equitable Use: The design is useful and marketable to people with diverse abilities.	1a. Provide the same means of use for all users: identical whenever possible; equivalent when not. 1b. Avoid segmentation or stigmatization of any users. 1c. Provisions for privacy, security, and safety should be equally available to all users. 1d. Make the design appealing to all users.
2	Flexibility in use: The design accommodates a wide range of individual preferences and abilities.	2a. Provide choice in methods of use. 2b. Accommodate right- or left-handed access and use. 2c. Facilitate the user's accuracy and precision. 2d. Provide adaptability to the user's pace.
3	Simple and Intuitive Use Use of the design is easy to understand, regardless of the user's experience, knowledge, language skills, or current concentration level.	3a. Eliminate unnecessary complexity. 3b. Be consistent with user expectations and intuition. 3c. Accommodate a wide range of literacy and language skills. 3d. Arrange information consistent with its importance. 3e. Provide effective prompting and feedback during and after task completion.
4	Perceptible Information The design communicates necessary information effectively to the user, regardless of ambient conditions or the user's sensory abilities.	4a. Use different modes (pictorial, verbal, tactile) for redundant presentation of essential information. 4b. Provide adequate contrast between essential information and its surroundings. 4c. Maximize "legibility" of essential information. 4d. Differentiate elements in ways that can be described (i.e., make it easy to give instructions or directions). 4e. Provide compatibility with a variety of techniques or devices used by people with sensory limitations.
5	Tolerance for Error The design minimizes hazards and the adverse consequences of accidental or unintended actions.	5a. Arrange elements to minimize hazards and errors: most used elements, most accessible; hazardous elements eliminated, isolated, or shielded. 5b. Provide warnings of hazards and errors. 5c. Provide fail safe features. 5d. Discourage unconscious action in tasks that require vigilance.
6	Low Physical Effort The design can be used efficiently and comfortably and with a minimum of fatigue.	6a. Allow user to maintain a neutral body position. 6b. Use reasonable operating forces. 6c. Minimize repetitive actions. 6d. Minimize sustained physical effort
7	Size and Space for Approach and Use Appropriate size and space is provided for approach, reach, manipulation, and use regardless of user's body size, posture, or mobility.	7a. Provide a clear line of sight to important elements for any seated or standing user. 7b. Make reach to all components comfortable for any seated or standing user. 7c. Accommodate variations in hand and grip size. 7d. Provide adequate space for the use of assistive devices or personal assistance.

Universal Design principles – applied to ”care” robots (inclusive robot design)

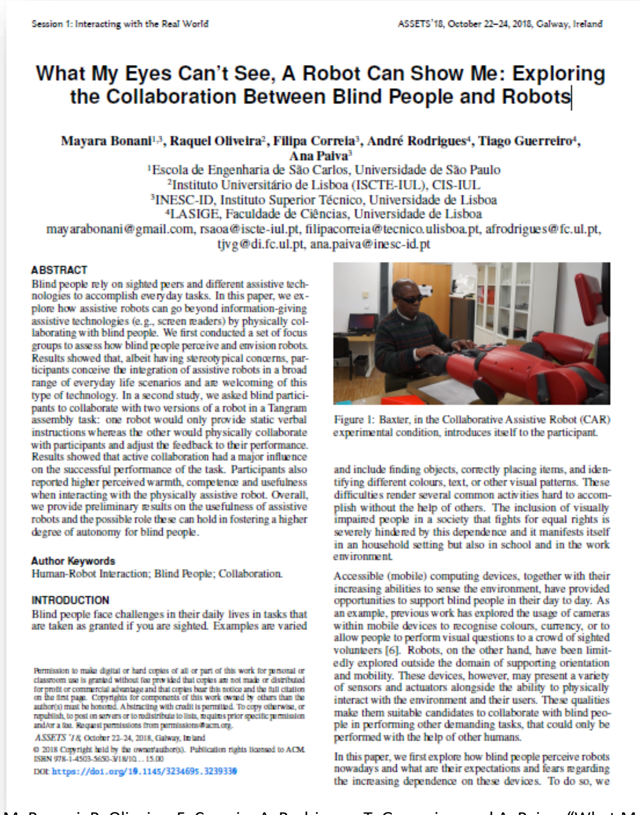
#	UD Principle	Example on UD principle applied to the physical layer	Example on UD principle applied to the virtual interaction layer
1	Equitable use	The physical design of the robot should be appealing to different types of users. For instance, the robot could be equally used by elderly patients without the feeling of infantilization, but also by children patients. The physical design and form should be appropriate for a diversity of the users. For instance, the size of the arm manipulator, and hands and grip of the robot should be appropriate to be used by both adults and children.	The robot can adjust its interaction to the user. It can interact through speech for those preferring a such interaction, or through displaying a text through a screen for those that are hearing impaired, or through color feedback, for those who need simple interaction.
2	Flexibility in use	The robot has an adjustable height: it can go up if the human user is standing, or it can go down, if the human user is sitting.	The robot interaction types should be multimodal and customizable depending on the type of user that is interacting with the robot.
3	Simple and intuitive use	Design of different components of the robot should be simple and intuitive to use. For instance, a robotic arm should be designed looking alike a human arm. The stop button should always be visible and placed in a specific place of the same type of robot. The stop button should always be red and have a stop icon, and/or the word stop on it.	The robot should use clear language that is understood by the user.
4	Perceptible information	The design of the symbols used by in the navigation display should follow the international guidelines. The symbol for play, stop, go back, move up and down, volume up and down etc. should be used accordingly.	The language used should be clear language and adjusted according to the mother tongue or the used language of the human user. For instance, if the user has Norwegian as his/her mother tongue, then the robot should be able to interact in Norwegian. Another situation is that the robot avoids giving technical errors to non-technical users. For instance, the robot displaying or indicating error 451 does not say much to a user. Instead, the robot should display or say in clear language what is eventually wrong.
5	Tollerrance for error	The robot shall be equiped with wheels that can navigate different types of floors, including slippery floors, but also floors that have carpets, or being able to go over the doorstep without getting stuck.	The interaction of the robot should be designed with tollerance for error in mind, without the robot “loosing” its patience, or becoming rude if the user takes more time to execute a task. If the robot is designed to indicate the human user to eat breakfast or to move around, but the human user refuses to do so, the robot should try to understand the reason why the human user does not execute the tasks, rather than punishing the user.
6	Low physical effort	The robot’s physical design should allow different users to adopt a neutral body position, and a minimum effort. Incorporating an adjustable height to the robot is an illustrative example for this principle.	The robot should be able to adjust its interaction speech depending if the user is an elderly patient, a child, a medical staff, or a technical staff. The robot should not make the human user to him- or herself adjust to the robot language in order to make him – or herself understood, but the other way around (see example of current chatbots that make the human user adjust his- or her language to the chatbot).
7	Size and space for approach and use	The physical design of the robot should be appropriate to its functionalities and aim. For instance, if a robot shall be designed for its use within a home, then it should not take too much space. Its height should not be greater than the humans height, however it should not be too small, such as that the human user may stumble into it while walking in the home. For instance, if the size of the robot is too small, a user sitting in a wheelchair, or a user with back problems will have to bend to reach the robot if the robot gets stuck. These situations should be avoided.	The size of the display, arms and grips, if any, should be appropriate to the size of the robot. However, the display of the robot should be enough big so an elderly person or someone with sight impairments can easily see the text, icons or symbols displayed.

Example: T-ABLE robot – applying Universal Design principles on the physial characteristic of a robot



D. Saplacan, J. Herstad, and T. Schulz, “T-ABLE - The Robotic Wood Table: Exploring situated abilities with familiar things,” *Int. J. Adv. Intell. Syst.*, vol. 13, no. 3 & 4, Dec. 2020, [Online]. Available: https://www.iariajournals.org/intelligent_systems/index.html
J. Herstad, T. W. Schulz, and D. Saplacan, “T-able: An Investigation of Habituating Moving Tables at Home,” *Univers. Des. 2021 Spec. Mainstream Solut.*, pp. 238–251, 2021, doi: 10.3233/SHTI210400.

Example: Robots for- low vision or blind people


Session 1: Interacting with the Real World ASSETS '18, October 22–24, 2018, Galway, Ireland

What My Eyes Can't See, A Robot Can Show Me: Exploring the Collaboration Between Blind People and Robots

Mayara Bonani[1,3], Raquel Oliveira[2], Filipa Correia[3], André Rodrigues[4], Tiago Guerreiro[4], Ana Paiva[3]
[1]Escola de Engenharia de São Carlos, Universidade de São Paulo
[2]Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL
[3]INESC-ID, Instituto Superior Técnico, Universidade de Lisboa
[4]LASIGE, Faculdade de Ciências, Universidade de Lisboa
mayarabonani@gmail.com, rsaoa@iscte-iul.pt, filipacorreia@tecnico.ulisboa.pt, afrodrigues@fc.ul.pt, tjvg@di.fc.ul.pt, ana.paiva@inesc-id.pt

ABSTRACT

Blind people rely on sighted peers and different assistive technologies to accomplish everyday tasks. In this paper, we explore how assistive robots can go beyond information-giving assistive technologies (e.g., screen readers) by physically collaborating with blind people. We first conducted a set of focus groups to assess how blind people perceive and envision robots. Results showed that, albeit having stereotypical concerns, participants conceive the integration of assistive robots in a broad range of everyday life scenarios and are welcoming of this type of technology. In a second study, we asked blind participants to collaborate with two versions of a robot in a Tangram assembly task: one robot would only provide static verbal instructions whereas the other would physically collaborate with participants and adjust the feedback to their performance. Results showed that active collaboration had a major influence on the successful performance of the task. Participants also reported higher perceived warmth, competence and usefulness when interacting with the physically assistive robot. Overall, we provide preliminary results on the usefulness of assistive robots and the possible role these can hold in fostering a higher degree of autonomy for blind people.

Author Keywords

Human-Robot Interaction; Blind People; Collaboration.

INTRODUCTION

Blind people face challenges in their daily lives in tasks that are taken as granted if you are sighted. Examples are varied and include finding objects, correctly placing items, and identifying different colours, text, or other visual patterns. These difficulties render several common activities hard to accomplish without the help of others. The inclusion of visually impaired people in a society that fights for equal rights is severely hindered by this dependence and it manifests itself in an household setting but also in school and in the work environment.

Accessible (mobile) computing devices, together with their increasing abilities to sense the environment, have provided opportunities to support blind people in their day to day. As an example, previous work has explored the usage of cameras within mobile devices to recognise colours, currency, or to allow people to perform visual questions to a crowd of sighted volunteers [6]. Robots, on the other hand, have been limitedly explored outside the domain of supporting orientation and mobility. These devices, however, may present a variety of sensors and actuators alongside the ability to physically interact with the environment and their users. These qualities make them suitable candidates to collaborate with blind people in performing other demanding tasks, that could only be performed with the help of other humans.

In this paper, we first explore how blind people perceive robots nowadays and what are their expectations and fears regarding the increasing dependence on these devices. To do so, we

Figure 1: Baxter, in the Collaborative Assistive Robot (CAR) experimental condition, introduces itself to the participant.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ASSETS '18, October 22–24, 2018, Galway, Ireland
© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-5650-3/18/10 ... 15.00
DOI: https://doi.org/10.1145/3234695.3239330


Figure 1: Baxter, in the Collaborative Assistive Robot (CAR) experimental condition, introduces itself to the participant.

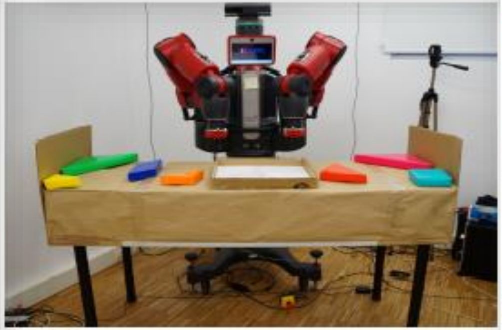


Figure 2: Initial setup of the assembling task.

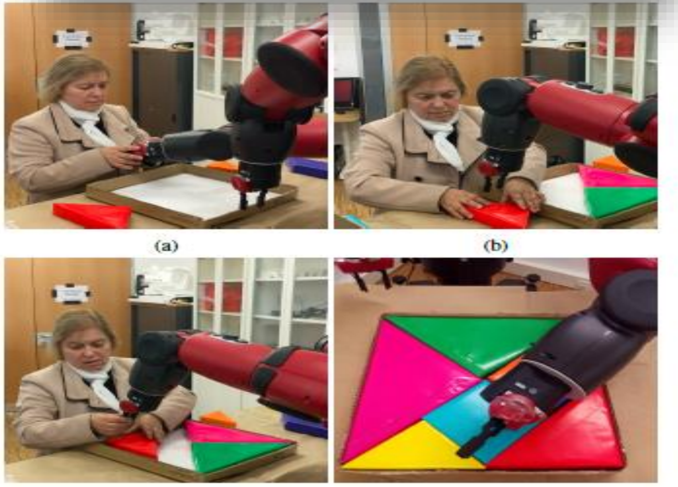

(a) (b) (c) (d)

Figure 3: (a) (1) Baxter extends its arm; (b) (2) Baxter indicates the position of the next piece to assemble; (c) (3) Baxter indicates the final position for the current piece; (d) (4) Baxter provides orientation instruction with its arm.

M. Bonani, R. Oliveira, F. Correia, A. Rodrigues, T. Guerreiro, and A. Paiva, "What My Eyes Can't See, A Robot Can Show Me: Exploring the Collaboration Between Blind People and Robots," in *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, New York, NY, USA, Oct. 2018, pp. 15–27. doi: 10.1145/3234695.3239330.

Example: TIAGo for deaf people using sign language

2022 31st IEEE International Conference on Robot and
Human Interactive Communication (RO-MAN)
August 29 - September 2, 2022. Naples, Italy.

Nothing About Us Without Us: a participatory design for an Inclusive Signing Tiago Robot

[1] Emanuele Antonioni [2] Cristiana Sanalitro [3] Olga Capirci [3] Alessio di Renzo [4] Maria Beatrice D'Aversa
[5] Domenico Bloisi [1] Lun Wang [1] Ermanno Bartoli [1] Lorenzo Diaco [6] Valentina Presutti
[1] Daniele Nardi

[1] Dept. of Computer, Control, and Management Engineering Sapienza University of Rome, Rome (Italy)
[2] International Telematic University UNINETTUNO Rome, (Italy)
[3] Institute for Cognitive Sciences and Technologies - National Research Council (ISTC-CNR) Rome, (Italy)
[4] School of LIS - SILIS Group Rome, (Italy)
[5] Dept. of Mathematics, Computer Science, and Economics, University of Basilicata, Potenza, (Italy)
[6] LILEC, University of Bologna, Bologna (Italy)

***Abstract*—The success of the interaction between the robotics community and the users of these services is an aspect of considerable importance in the drafting of the development plan of any technology. This aspect becomes even more relevant when dealing with sensitive services and issues such as those related to interaction with specific subgroups of any population. Over the years, there have been few successes in integrating and proposing technologies related to deafness and sign language. Instead, in this paper, we propose an account of successful interaction between a signatory robot and the Italian deaf community, which occurred during the Smart City Robotics Challenge (SciRoc) 2021 competition [1]. Thanks to the use of a participatory design and the involvement of experts belonging to the deaf community from the early stages of the project, it was possible to create a technology that has achieved significant results in terms of acceptance by the community itself and could lead to significant results in the technology development as well.**

***Index Terms*—Robotics, Sign Language, Inclusivity**

I. Introduction

The World Federation of the Deaf (WFD), in the Statement on Sign Language Work (2014), considers the exclusion of the Deaf Community and their national organizations from sign language work a violation of the linguistic human rights of deaf people. Decisions regarding sign languages should always remain within the linguistic community, in this case, deaf people: Nothing About Us Without Us. In 2019 the WFD drew up a Position Paper on Accessibility about Sign Language Interpreting and translation and technological developments. In this document, the importance of recognizing some unique characteristics of sign languages is strongly claimed, which becomes possible only through consideration of the context and the cultural norms on which they are based: "It is of paramount importance that Deaf communities are part of the design in the machine learning developments for any automated sign language translations". There are many innovations and technologies currently available in the field of robotics. Over the last few years, several studies have tried to explore the possibilities of these innovations in automated sign language translation to increase the inclusion of deaf people who use this language. However, from the results of the studies, promising developments have not always emerged, especially in terms of acceptability. This study has been held as part of the organization of the SciRoc 2021 competition held in Bologna. The organization of the competition included tasks related to the daily city context for social robots, so it was natural to take into account the concept of inclusivity and accessibility of technologies. The concept of inclusivity is fundamental to modern society, and an essential step towards it is the possibility for the robot to understand and perform sign language (SL). We also believe that to obtain an appreciable result accepted by the deaf community, they must be involved in all phases of the project. Our study started from these considerations, and all the phases of the work carried out are based on participatory design or co-design, not "for" deaf people but "with" deaf people. Two deaf and one hearing socio-linguistic professionals, sign language experts, from the Language and Communication across Modalities (LaCAM) Lab of the Institute for Cognitive Sciences and Technologies - National Research Council (ISTC-CNR) created a dataset of signs based on linguistic and technological reflections. The team of the RoCoCo lab of the Sapienza University of Rome, as an expert in social robotics, provided all the technical specifications of the Tiago robot available for the SciRoc competition. In fact, Tiago is equipped with a single robotic arm whose movement is rather limited for the torsion of the wrist, while for the hand, there are three motors that do not

[1] https://sciroc.org/

978-1-7281-8859-1/22/$31.00 ©2022 IEEE

+++ studies with people with autism and robots etc.

Universal Design of Robots (UD-ROBOTS)

Bufdir

Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

Universal Design of Robots (UD-Robots)

Robots, being physical and digital, must adhere to different sets of guidelines if they are to be universally designed. Which guidelines are necessary for the universal design of a robot, and how can we evaluate a robot to see if it is universally designed? This project aims at examining existing guidelines to see how they apply to robots, discussing with potential users of robots in different user cases, devising a method for evaluating robots, and using this method to evaluate several robots.

UNIVERSITAS OSLOENSIS MDCCCXI

VID

Questions

1. What is accessibility of robots for you?

2. Do you see any benefits and/or challenges of applying universal design to robots?

Thank you.

UiO : **Department of informatics**

Feel free to reach me on:

diana.saplacan@ifi.uio.no
LinkedIn: /dianasaplacan

Thank you for your attention