Coding Guidelines - Perception of Soft Robots

Before starting coding, if you have not mentally present the study and study design, please (a) click through our main study on risks and benefits of soft robots & read the scenario texts on (i) socially assistive robots + information "soft", (ii) search and rescue robots + information "soft", (b) read the randomly created sub-word list generated by the Shiny-App which we provide to you. The random sub-word list was created after summarization of a part of the concepts that have been written by participants in our CAM-study, (c) read the coding guidelines below (the category system) several times, (d) the "to consider" section below and (e) the graphical representation of possible connections between the concepts at the end of the document (p. 7). It is strongly recommended to display the coding guidelines on a second screen during coding or to print them out in advance.

Link to main study:
(i) socially assistive: https://studien.psychologie.uni-freiburg.de/publix/564998da-e3fa-491b-8844-86464ce30ddc/df9ee6a1-1d9c-4013-a4f1-16e4785745b2/start
(ii) search and rescue: https://studien.psychologie.uni-freiburg.de/publix/12e3668a-de47-446f-80e4-47be40be2f78/df9ee6a1-1d9c-4013-a4f1-16e4785745b2/start

General remarks:

- After reviewing the coding guidelines, categorize the concepts from the provided word list into subordinated categories based on our coding guidelines, please consider the comment columns and the valence of the respective words.
    - Please use the "codes" of the categories for the coding process. This means you **enter the codes of the respective categories in the first column of the wordlist "Superordinated" of the word list**. The "codes" can be found in the first column of the coding guidelines.
    - Important: After our joint session, 4 new categories were added to the coding guidelines and two categories were modified: new categories: (i) technological limitations, (ii) technological possibilities, (iii) rest category, ambiv., (iv) rest category, neutr. Modified categories: (i) Perceived usefulness -> now: rest category, pos.: perceived usefulness, (ii) Perceived uselessness -> now: rest category, neg.: perceived uselessness.
- Important: The word list also includes our predefined concepts for the CAMs. Please do not summarize them into superordinate categories. The predefined terms are: "Vorteile", "Nachteile", "sozialer Assistenzroboter", "Rettungsroboter".[German]
- Please consider the valence and provided comments by the participants of the words for the coding process.
- Please carefully decide which category each concept belongs to.

How to proceed:

- To summarize the concepts, please do the following: In the third column of the wordlist (C) you will see concepts, which have been written by the participants in our study.
    - Based on our coding guidelines, you should consider which superordinate category the respective concept should be assigned to. You then enter the **code** of the superordinate term based on our coding guidelines in the first column ("Superordinate"). For example, the concept "Diskriminierung" could be superordinate into the category "MT" (Mistrust) from our coding guidelines.
- You should not summarize the predefined concepts (see general remarks above).

- If you are facing any problems or have any open questions regarding single concepts, please write them down in the respective comment column (2). Please do not talk to other rather during this process, because the independence of the raters is crucial.

Timeline for the three central tasks:

1. Done: finish the first coding process (each of us will code the same 100 concepts) till 20th of February
2. Done: we have a joint session to discuss possible ambiguities and difficulties in a hybrid meeting on the 27th of February
3. Please finish the second coding process till 18th of March

To consider

Please consider the following while coding:

- All the key points mentioned in the "Coding Rules" do not have to be fulfilled to assign the respective categories.
- If you have difficulties assigning the concepts from the wordlist to the given categories, sort them into the 4 "rest categories" (pos., neg., amiv., neutr. considering the valence) > you should use the comment column (2) to explain why you are assigning a term to a "rest category".
- In the end of the document, after the coding guidelines, you will find a graphic showing the complementary terms of the coding guidelines (such as safety vs. risk) as well the differentiation between the terms (coding rules).
- You can write comments if you have assigned concepts to categories but have questions/uncertainties.
- Please assign every word to one (but only) one of the provided codes in the category system.

If you have any questions or something is unclear, please send me an E-Mail:
louisa.estadieu@philosophie.uni-freiburg.de

Coding Guidelines for Assessing Risks and Benefits of (Soft) Robots

Code	Category	Definition	Coding Rules	Examples from data
SA	safety	The belief that the robot is ● operating safely ● performing tasks without posing risks and/or harm to humans ● performing tasks even safer than humans	Compared to "Trust" this concept primarily concerned with <u>operational safety</u> rather than overall reliability and integrity of the robot	"Accuracy of execution"; "Safer for humans"; "Precision and efficiency"; "Reduced risk of injury"; "Quick action"
R	risk	The belief that the robot is ● operating unsafely ● performing tasks that pose risks to humans ● performing tasks riskier than humans	Compared to "Mistrust" this concept emphasizes the assessment of risky behaviors and actions on the <u>operational level</u>, rather than focusing on the robot's overall reliability and integrity	"Lack of accuracy"; "Slight damage"; "Possible malfunctions"
T	trust	The belief that the robot ● performs with integrity and/or reliability ● is thus considered trustworthy	Compared to "Safety" this concept focuses on the overall <u>reliability and integrity</u> of the robot, rather than focusing solely on operational safety components	"Consistency of care"; "objective/neutral"; "Adaptation to patient/client"; "better knowledge capacity"; "does not evaluate"
MT	mistrust	The lack of belief that the robot ● performs tasks with integrity and/or reliability ○ leading to mistrust, skepticism and/or doubt	Compared to "Risk" this concept focuses on skepticism or doubts about the robot's <u>reliability and integrity</u>, rather than focusing on specific risky behaviors or actions on the operational level	"Discrimination against individuals"; "Favoritism"; "can be abused"; "Algorithm bias"

AP	anthropomorphism, pos.	Positive evaluation of the ● attribution of human characteristics and/or behaviors to the robot ● such as autonomy or emotions	Compared to "HRI, pos.," this concept focuses more on the evaluation of specific positive robot <u>characteristics</u> than on the interaction itself	"Autonomous decisions"; "Comforts the injured"; "Conversation partner for the lonely"
AN	anthropomorphis, neg.	Negative evaluation of the ● attribution or lack of human characteristics and/or behaviors to the robot ● such as autonomy or emotions	Compared to "HRI, neg.," this concept focuses more on the evaluation of specific negative robot <u>characteristics</u> than on the interaction itself	"Autonomous decisions"; "limited emotional intelligence"; "lack of empathy"
HRIP	Human-Robot-Interaction = HRI, pos.	Positive implications of the ● interaction between humans and robots	Compared to "Social impact, pos.," this concept emphasizes positive aspects of the direct <u>interaction of humans and robots</u>, not broader societal impacts	"Interaction with elderly/autistic individuals"; "Assistance with feelings of shame (objective/neutral)"
HRIN	Human-Robot-Interaction = HRI, neg.	Negative implications of the ● interaction between humans and robots	Compared to "Social impact, neg.," this concept emphasizes negative aspects of <u>interaction</u>, not broader societal impacts	"Unwanted emotional relationship"; "Impersonal"; "Exclusion"; "Dependency"
SIP	social impact, pos.	Positive impacts of ● robots ● human-robot-interaction on society	Compared to "HRI, pos.," this concept is not about the direct interaction between humans and robots but emphasizes positive <u>impacts</u> that robots and HRI might have on society, including long term impacts	"more job opportunities"; "promotes social interaction/integration"; ""ustice (objectivity)"; "reduces loneliness"

SIN	social impact, neg.	Negative impacts of • robots • human-robot-interaction society	Compared to "HRI, neg.," this concept emphasizes negative <u>impacts</u> that robots and HRI might have on society, including long term impacts	"fewer jobs"; "loss of human nature"; "loss of human interaction"; "loss of human skills"; "coarsening of human behavior"; "people being replaced"
TP	technological possibilities	Technological possibilities oft the robot	Compared to "Trust", "HRI, pos." & "Social impact, pos." this concept focuses solely on the technological possibilities of the robot	"permanent operational readiness"; "can overcome hurdles"
TL	technological limitation	Technological limitations of the robot	Compared to "Mistrust", "HRI, neg." & "Social impact, neg." this concept focuses solely on the technological limitations of the robot. It can be rated positively or negatively	"Limited know-how"; "limited autonomy"; "Defects"; "Limited learning ability"
LC	low cost	Low cost of • robot fabrication and deployment • on a personal and/or societal level	Compared to "High Cost," this concept focuses on economical aspects and cost-effectiveness in robot development and deployment	"inexpensive"; "cost-efficient"; "cheap labor"
HC	high cost	High cost of • robot fabrication and deployment • on a personal and/or societal level	Compared to "Low Cost," this concept highlights the higher financial investment required for robot development and deployment	"expensive"; "high costs"
RCPP	rest category, pos.: perceived usefulness	The belief that the robot is • useful and/or assistive and/or overall good • on a personal and/or societal level • the concept cannot be assigned, but is general positive	All <u>positive</u> terms that cannot be clearly assigned to one of the above-mentioned belong in this category. Can be concepts regarding usefulness, and others	"customized programming"; "for therapy"; "consistency of care"; "cognitive superiority"

RCPN	rest category, neg.: perceived uselessness	The belief that the robot is ● useless and/or overall bad ● on a personal and/or societal level 　　● the concept cannot be assigned, but is general negative	All negative terms that cannot be clearly assigned to one of the above-mentioned belong in this category. Can be concepts regarding perceived uselessness, and others	"no specific expertise"; "difficult to adapt"
RCA	rest category, ambiv.	● All terms that ● do not fall under above constructs, ● and/or are neither positive or negative but ambivalent	All terms that cannot be clearly assigned to one of the above-mentioned constructs because of their ambivalent valance belong in this category	"Unclear ethical behavior"; "Liability"
RCN	rest category, neutr.	● All terms that ● do not fall under above constructs, ● and/or are neigther positive/negative, ambivalent but neutral	All terms that cannot be clearly assigned to one of the above-mentioned constructs because of their neutral valence belong in this category	"Construction method"; "hybrid construction"

Graphical representation of possible connections between the concepts

In the following graphic the complementary terms of the coding guidelines (such as safety vs. risk) are shown, whereby only the codes are written. Arrows on the left or right indicate that these terms were mentioned in the coding rules of the respective other term, which highlights the need to differentiate between these terms.