# Facilitators' guide for a workshop with the human-AI collaboration cards

peer

VUB SMIT STUDIES IN MEDIA, INNOVATION & TECHNOLOGY RESEARCH GROUP

Knowledge Centre Data & Society

"THE HYPER EXPERT COLLABORATIVE ASSISTANT"

## Table of Contents

# 1 Introduction

This guide is developed by imec-SMIT, Vrije Universiteit Brussel for task T2.2 'Co-creation with representations of AI systems' of the PEER project. The aim of task 2.2 is to develop representations of AI systems so that parts of the AI system become design materials in workshops with end-users. With these representations, users can **co-create their ideal human-AI collaboration form**.

This task evolved in the creation of **human-AI collaboration cards**; a card set to use within a workshop with end-users to define their ideal human-AI collaboration form. The aim of the **workshop** is to define **requirements for human-AI collaboration**, keeping in mind the envisioned AI systems of the use-cases in PEER.

In this guide, the facilitator can find more information about the human-AI collaboration cards, a detailed description of the workshop and its exercises, and lastly, some tips and tricks for facilitating the workshop.

# 2 Human-AI collaboration cards

The human-AI collaboration cards are made to discuss with the end-users what their ideal human-AI collaboration should look like. The human-AI collaboration cards consist of two different cards: the AI collaborator cards and the AI characteristic cards. By making use of the card set, end-users can build their ideal AI partner and define the characteristics, features, and elements of the human-AI collaboration. The complete card set can be found in the annex of this document.

## 2.1 AI collaborator cards

Within the cards, there are three **AI collaborator cards** that describe the type of human-AI collaboration: AI as 'an assistant, as 'a guide', and as 'an advisor'. These are based on the use-cases within the PEER project, however, when using the cards within another project, the facilitator and use-case owners are recommended to create their own AI collaborator card(s) based on the previous user research done within the project.

The AI collaborator cards that are included in the card set are:

- **'The assistant'**: the AI system supports its end-users in fulfilling their actions. It provides useful information at the right moment (on request for example). Within this human-AI collaboration, the level of autonomy of the end-users is rather high, and the authority of the AI system is rather low.
- **'The guide'**: the AI system gives clear instructions and directs the end-users to reach their goals. The end-users rely heavily on the system as they are following the instructions of the AI system. Therefore, the level of autonomy of the user is in this case rather low, and the authority of the AI system is more likely to be high.
- **'The advisor'**: the AI system proposes possible actions to its end-users, and the end-users can choose to accept these actions or reject them. The level of autonomy of the user is in this case high (as the users takes the final decision), and the level of authority of the AI system is rather high too, as the end-users expect the AI system to propose correct actions.

This is an interpretation of the preferred AI collaborator based on the PEER use-cases, of course, this might differ when actually talking with a group of end-users and amongst the end-users themselves. Some participants will perceive 'the advisor' as having less authority or providing less autonomy to the user, than other participants. Therefore, the AI collaborator cards must be seen as a conversation starter to discuss with the end-users how they envision their ideal human-AI collaboration. The facilitators elaborates on each of the cards in group to discuss what the characteristic is about and to discuss with the group of end-users (the participants) want/deem important this characteristic to be included in the end result. If needed, an empty AI collaborator card is included in the card set, so the facilitator and end-users can come up with their own type of AI collaboration and an appropriate definition.

## 2.2 AI characteristic cards

The **AI characteristic cards** describe several elements of a human-AI collaboration, such as trust enablers, attributes, capabilities, qualities, tone of voice, and interface types. By using these cards, the end-users can elaborate more on what their ideal human-AI collaboration would look like by expressing what characteristics they deem more important, or less. If a characteristic is missing according to the end-users, they are given the ability to add the missing characteristic on an empty card.

The elements of the AI characteristic cards are given a definition too, as these concepts (trust enablers, etc.) might have a different meaning among its users (i.e. the facilitator and participants of the workshop). It is your role as a facilitator to inform the end-users of the AI systems what is meant with each of the elements. By doing so, all participants are on the same line in the workshop, and it will smoothen the discussion in the workshop.

As a facilitator, it is crucial to be well prepared for the workshop, to have read all the cards beforehand and to understand clearly what is written on the AI collaborator and AI characteristic cards (i.e. definitions). In 'Table 1: Overview of the AI characteristics cards' an overview of the AI characteristic cards is given per element of the human-AI collaboration. The detailed definition of each of the AI characteristics cards can be read on the cards in the annex of this document. In the table, possible pros and cons of the AI characteristics on the technology design are given. These are just examples, for illustrative purposes, and they should capture what the possible effects (both benefits or potential risks) could be or might arise when a certain characteristic is embedded within the AI system, and how that might affect or not the functioning of the system. These pros and cons might come up in the discussion with end-users, as a facilitator it is good to know (or think of in advance about) some of these pros and cons as well to inform the end-users of possible effects when they talk about their preferred AI characteristics. These examples must give a clear idea to the end-users what the consequences of their choices are. For example, if they deem clarity more important than guidance, what could this mean for the final human-AI collaboration? When these characteristics are discussed, there is a good chance that end-users will automatically discuss what level of clarity is important for them and this type of information is crucial for the design of the technology.

| Element of human-AI collaboration | Definition of element | Related AI characteristics | Some examples of pros and cons of AI characteristics |
|---|---|---|---|
| Trust enablers | Enablers are AI characteristics that can increase the feeling of trust in AI systems. | 1. Transparency<br>2. Empathy<br>3. Reliability<br>4. Explainability<br>5. Ethically<br>6. Privacy<br>7. Accuracy<br>8. Controllability | • If the AI system is highly transparent, it might frighten or overburden the end-user with the amount of information.<br>• On the other hand, showing how certain decisions are made, and based on which inputs (transparency), will |

| | | | |
|---|---|---|---|
| | | | make the system more understandable (explainability)and thus, more trustworthy, leading to better acceptance and adoption<br>• If the AI system is good in explaining why it made a certain decision, it might also confuse end-users with little technical knowledge.<br>• However, if the transparency and explainability are done in a language appropriate for each user group (e.g., simple language, no legalese) is expected to be an advantage for the users and for the acceptance of the AI system. |
| System attributes | Attributes allow AI systems to interact, work together, and adjust efficiently in changing environments. They enhance the usability and significance of AI systems for end-users. | 1. Sensing<br>2. Predictability<br>3. Directivity<br>4. Directability<br>5. Adaptability<br>6. Awareness sharing<br>7. Customisability<br>8. Traceability | • If the AI system is highly adaptable, the end-user might become frustrated as it might not be clear based on what the AI system is adapting (again).<br>• On the other hand, a highly adaptable system might be what the users need (e.g., a system adapting quickly to their needs, or to changing inputs and environments)<br>• A system that is highly customisable might quickly become confusing, because of the many parameters involved.<br>• For a system to be good at sensing, adapting, etc, it needs to collect a lot of data about its user and its environment.<br>• However, a system that is traceable will provide good overview and explanation on how decisions were being made and based on what |
| Capabilities | Capabilities refer to the tasks or functions that an AI system can perform | 1. Recognition<br>2. Prediction<br>3. Reasoning<br>4. Generation<br>5. Recommendation | • If a function of the AI system is to make predictions and recommendations, it might give wrong predictions as it bases its predictions on |

| | | | |
|---|---|---|---|
| | exceptionally well, particularly in scenarios where the AI assists humans in a collaborative setting. | | historical data about past evens and actions (bias).<br>• However, if a system provides recommendations based on users' input and settings (preferences) will ensure that the outputs are aligned with the users' needs and requirements. |
| Collaboration qualities | Qualities are characteristics that influence the level of comfort and distinction in the human-AI collaboration. | 1. Clarity<br>2. Familiarity<br>3. Ergonomics<br>4. Responsiveness<br>5. Consistency<br>6. Guidance | • More guidance is not always better, as it might also have the opposite effect – because of the overload on information given, the end-user is lost and does not know what to do anymore.<br>• However, if this is being done with clarity (in a clear manner), it will be beneficial for the users since it will diminish the possibilities of misunderstandings and overload. |
| Tone of voice | Tone of voice refers to the way the AI-system is interacting with the user and communicating its personality. | 1. More formal<br>2. More serious<br>3. More respectful<br>4. More straightforward<br>5. More casual<br>6. More funny<br>7. More irreverent<br>8. More enthusiastic<br>9. More simple<br>10. More complex<br>11. Adjustable | • A formal tone of voice might give a severe and top-down impression, whereas a more casual tone of voice might be considered as unreliable.<br>• However, being able to choose their own tone of voice will make sure that users feel more familiar, more at ease, and more confident with the AI collaborator. |
| Interface types | Interface types refer to the medium through which the human and the AI system can collaborate. A system can use a combination of different interface types. | 1. Visual commands<br>2. Visual information<br>3. Auditory commands<br>4. Auditory information<br>5. Haptic commands<br>6. Haptic information | • The interface types are very context dependent. An auditory interface can for example only be of value in a quiet surrounding.<br>• However, being able to choose the interface that fits the specific users' needs will ensure their easiness of use and appropriateness for the contexts of use. |

Table 1: Overview of AI characteristics cards

## 2.3 Workshop template

A template is used within the workshop (see 3.2) to structure the exercise of 'AI character cards'. In this exercise, each participant is given the opportunity to decide which AI characteristics they prefer for their ideal human-AI collaboration. On the template, each participant can write which of the AI characters are highly important, fairly important, not at all important for their ideal human-AI collaboration. By doing so, their individual preferences are captured too.

The participants can categorise the AI characteristics in the corresponding row, going from highly important to not important at all for them to consider in the final AI technology design and human-AI collaboration. Each AI characteristic card is given a number, so the participants only need to write down the number of the corresponding AI characteristic in the template.

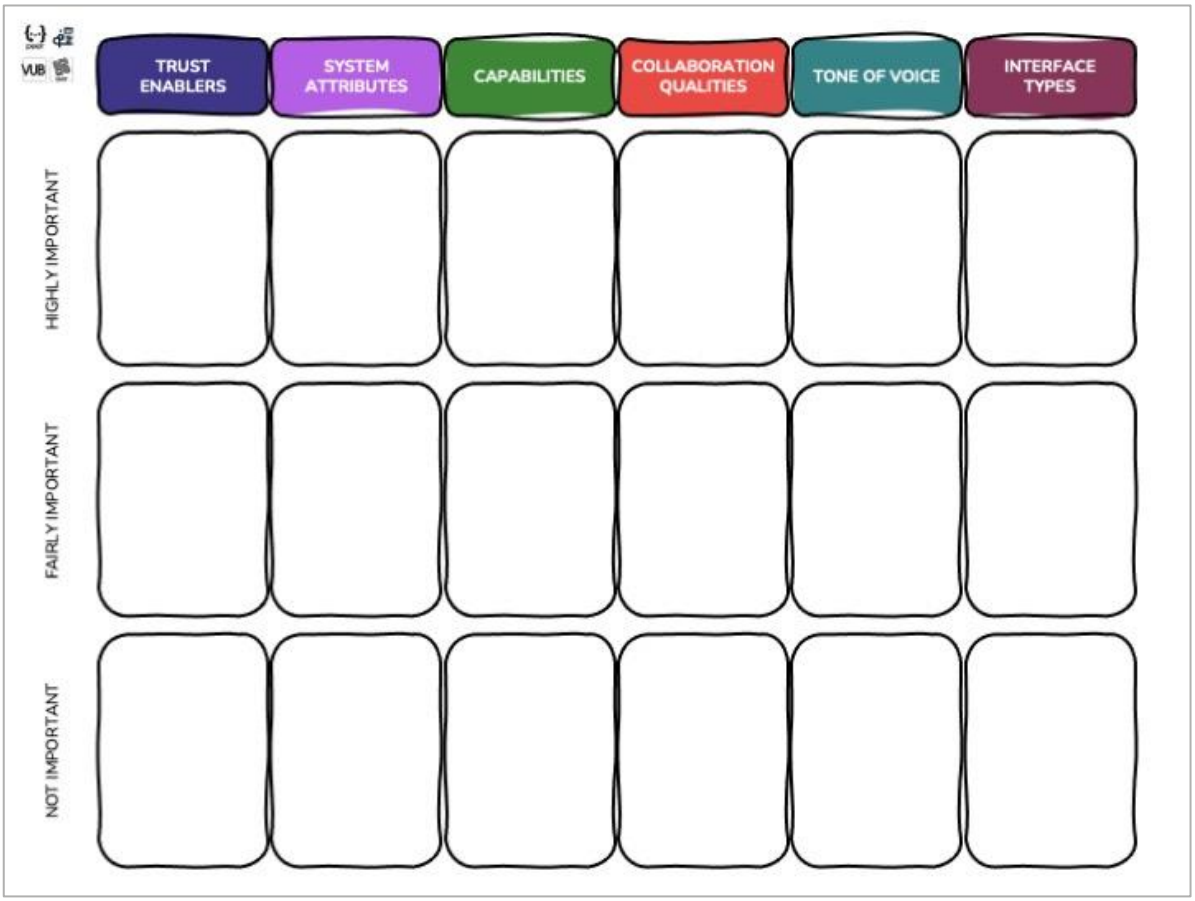

Figure 1: The workshop template

# 3 Workshop with human-AI collaboration cards

## 3.1 Before the workshop

Before the workshop starts, some crucial aspects need to be organised. First, the **participants** of the workshop need to be **recruited**. Ideally, minimum 10 participants will attend the workshop; this number allows for a diverse group in terms of age, gender, education, socio-economic status and similar. In case of a work context, include not just the employees who will directly use the system, but also employees that are/might be indirectly included/impacted by the system; it could be helpful to also include a member of a trade union, as this person also has good knowledge about complaints, wants and needs of other employees.

Make sure to schedule at least **two hours** for the workshop. With the introduction, outro and a break in between, a two-hour workshop will be over quickly. However, more time for a workshop is often more difficult to schedule given the availability of the participants and/or the company's production processes schedule.

Second, a **room** in which the workshop will take place must be reserved; preferably in the facilities of the company or the organisation itself. It is necessary to have a room that is large enough for at least 11 persons (10 participants and one facilitator) to sit comfortably, away from too much noise and a room in which there are tables present, so the participants can write and do the different exercises within the workshop. Do also take accessibility of the location into account, to make sure all participants can be present. Do not forget to order **catering**, so the participants can have a drink and maybe eat a snack during the workshop. A **fee** to participate in the workshop can be foreseen as well if preferred.

The **facilitator** can either be a member of the organisation of the use-case or an external facilitator (i.e. freelancer). If the facilitator is a member of the organisation, this person cannot be a direct colleague (or manager) of the participants to ensure participants can speak freely and without any issues of power. A good facilitator can connect a group of participants, create enthusiasm and willingness to work together and reach results.

In order to **prepare** for the workshop, the facilitator must read this facilitator's guide, the human-AI collaboration cards, and of course, the background information on the use-case. If needed, the cards must be **translated** to the language of the participants.

## 3.2 Workshop guide

In the table below the different exercises within the workshop are described. For each of the exercises, the duration and the goal of the exercise, the steps on how to execute the exercise, the material needed, and some tips and tricks for the facilitator are mentioned. This guide will help the facilitator

to execute the workshop, but of course, the facilitator is free to adapt the workshop guide and change the exercises if deemed necessary.

| **Duration** | **Exercise** |
|---|---|
| 10 min | Intro |
| Goal | Participants are informed on what the workshop will be about, are comfortable, and ready for the workshop. |
| Steps | 1. The facilitator welcomes all participants and explains what the aim of the workshop is; to discuss what their ideal human-AI collaboration would look like based on the use-case. If needed, the facilitator describes what the use-case is about, and what the state of the technology design is so far.<br>2. The facilitator gives more information about what will be done with the results of the workshop; the results will be taken into consideration in the further development of the prototype of the use-case – of course it is up to the technical partners and the use-case owners to see what is feasible.<br>3. The facilitator details some guidelines to ensure the workshop will go smoothly, such as the participants can speak freely, listen to each other, give feedback and ideas without being interrupted, and if needed, participants can write down their input if they prefer not to speak for the group. |
| Material | / |
| Tips & tricks | • Be very open and transparent about what can and cannot be discussed in the workshop. What aspects of the technology design are still open for input and open to change, and what aspects are fixed and already decided upon. This will set the boundaries of the workshop. |

| **Duration** | **Exercise [in group]** |
|---|---|
| 20 min | AI collaborators |
| Goal | Participants must agree on an AI collaboration type for the use-case. |
| Steps | 1. The facilitator explains the 3 AI collaborator cards of the card set (see 2.1) and asks the participants what AI collaborator card they would prefer in the use-case. If they do not agree with one of the AI collaborator cards, they can also think of another AI collaboration type they would prefer.<br>2. The participants write their preference individually on a post-it. So, each participant writes one of the AI collaborator cards (or a new collaborator card) down on a post-it.<br>3. The facilitator gathers all the post-its and looks what AI collaborator card is the most preferred by the participants. The participants are given the opportunity to explain why they wrote their preference down.<br>4. The group can discuss what the envisioned AI system must be for them and the kind of role it must play within this human-AI collaboration.<br>5. The goal is to choose one type of AI collaborator at the end of this exercise. If a new type is agreed upon, the empty template is filled in by the facilitator. |
| Material | • Post-its<br>• Writing material<br>• AI collaborator cards (+ empty template of AI collaborator card) |

| | |
|---|---|
| Tips & tricks | • This exercise allows the group to choose the AI characteristics more easily in the next exercise as they can reflect on whether a particular AI characteristic contributes to the type of AI collaboration they have chosen in this exercise.<br>• Instead of writing their preference down on a post it, the participants can also discuss in group what AI collaborator they want the use-case to be. Guidance from the facilitator is needed in this discussion to ensure everyone can speak freely and listens to each other.<br>• Wrap-up the discussion after 20 minutes so there is still enough time for all the other exercises in the workshop. If needed, the facilitator must shut down the discussion and go to the next exercise.<br>• It could help if the facilitator already has a preferred AI collaborator in mind, based upon the previous user research done within the use-case. If needed, the facilitator can suggest this type of AI collaboration to shorten the discussion amongst the participants. |

| **Duration** | **Exercise [individually]** |
|---|---|
| 45 min | AI characteristics |
| Goal | Participants define what AI characteristics the human-AI collaboration must have. |
| Steps | 1. The facilitator explains what the AI characteristic cards (see 2.2) are about and what information can be found on the cards. Every participant receives a template (see 2.3) and some writing material.<br>2. The facilitator gives more information about the 6 categories of human-AI collaboration, and in which the AI characteristics are grouped. Participants can ask questions about these 6 elements.<br>3. Afterwards, the facilitator reads each AI characteristic card, and if needed gives more information about the card and what it is about. The facilitator asks if an AI characteristic was not mentioned and is missing in the card set. If so, an empty AI characteristic card can be filled in.<br>4. In the meantime, the participants write down the number of the AI characteristic card (in the upper left corner) on the template and categorise it according to the level of importance for them. Is the AI characteristic highly important, fairly important, or not important at all for their ideal human-AI collaboration? It is important that the participants do this exercise individually, so all opinions are gathered and considered in the next exercises of the workshop.<br>5. The facilitator collects all the templates. |
| Material | • AI characteristic cards (for the facilitator, for all participants)<br>• A template for each participant<br>• Writing material |
| Tips & tricks | • The facilitator can also do this exercise in group and discuss the different AI characteristics with all participants at once. This could be done when, for example, the goal is to find consensus within the group. However, a group discussion will make it more difficult to record the individual differences and preferences.<br>• There are 6 categories of AI characteristics, make sure there is enough time to discuss each category and categorise all AI characteristic cards on the template. |

| | |
|---|---|
| | • The facilitator can also hand out a card set to each participant, so they can read the AI characteristic cards themselves. This might fasten the exercise a bit as they can always look what is written on the AI characteristic card. |

| **Duration** | **Exercise [in group]** |
|---|---|
| 35 min | Ideal human-AI collaboration |
| Goal | The participants consolidate the AI characteristics and determine what their ideal human-AI collaboration would look like. |
| Steps | 1. The facilitator informs the group about big differences in the categorisation of the AI characteristics (i.e. when an AI characteristic is categorised as not important at all by one participant and as highly important by another – and vice versa). The participants can discuss these differences together and try to come to a consensus on the categorisation of the AI characteristics.<br>2. By doing so, the participants discuss what role their AI collaborator should take on, and which characteristics it should entail to make their collaboration pleasant and productive: how does it behave, how does it communicate and what can it (not) do?<br>3. The facilitator consolidates the discussion and writes down on which of the characteristics all participants agreed or not, or on which a consensus was not found. This gives a clear idea to the use-case owners and technical partners of PEER which of the characteristics are more or less important to consider in the development of the AI systems. |
| Material | • Filled in templates of the participants<br>• An empty template<br>• Writing material |
| Tips & tricks | • This exercise can be done in a group discussion in which the facilitator guides the discussion towards the end-goal of this exercise; an overview of characteristics all participants agreed on or not, or on which a consensus was not found. Other possibilities are to use an empty template on which the facilitator consolidates the discussion (the template guides the discussion) or a card sorting exercise. In the card sorting exercise, the cards are placed on a table into the six categories. The group can categorise the cards into the same categories on the template: highly important, fairly important and not important at all. This exercise is of course done in group as a consolidation exercise to find consensus amongst the participants. |

| **Duration** | **Exercise** |
|---|---|
| 10 min | Outro |
| Goal | Participants are thanked for their presence and input and are informed about the next steps. |
| Steps | 1. The facilitator thanks the participants for their active contribution and collaboration in the workshop and describes what will happen next. The output of this workshop will serve as input for the use-case owners and the technical partners within the PEER project to further develop the MVP of the consequential decision-making system. As mentioned before, the facilitator stresses again that not all preferences mentioned in the workshop will reach it to the final AI system as other choices need to be made such as cost, effort to ensure this preference, etc. |

| | |
|---|---|
| | 2. If there are still some questions about the cards, the workshop or the use-case, the facilitator takes the time to respond. |
| Material | / |
| Tips & tricks | / |

## 3.3 After the workshop

A report of this workshop is made by the facilitator that contains the individually filled in templates by the participants, and the consolidation done in the exercise 'ideal human-AI collaboration'. Remarks and things that stood out in the workshop are noted down in the report as well as these might contain important information about the ideal human-AI collaboration too. The report can be as concise as wanted but note that both the individual preferences of the human-AI collaboration and the preference of the group must be included so the use-case owners and the technical partners have a clear idea of which of the characteristics are more or less important to consider in the development of the AI system.