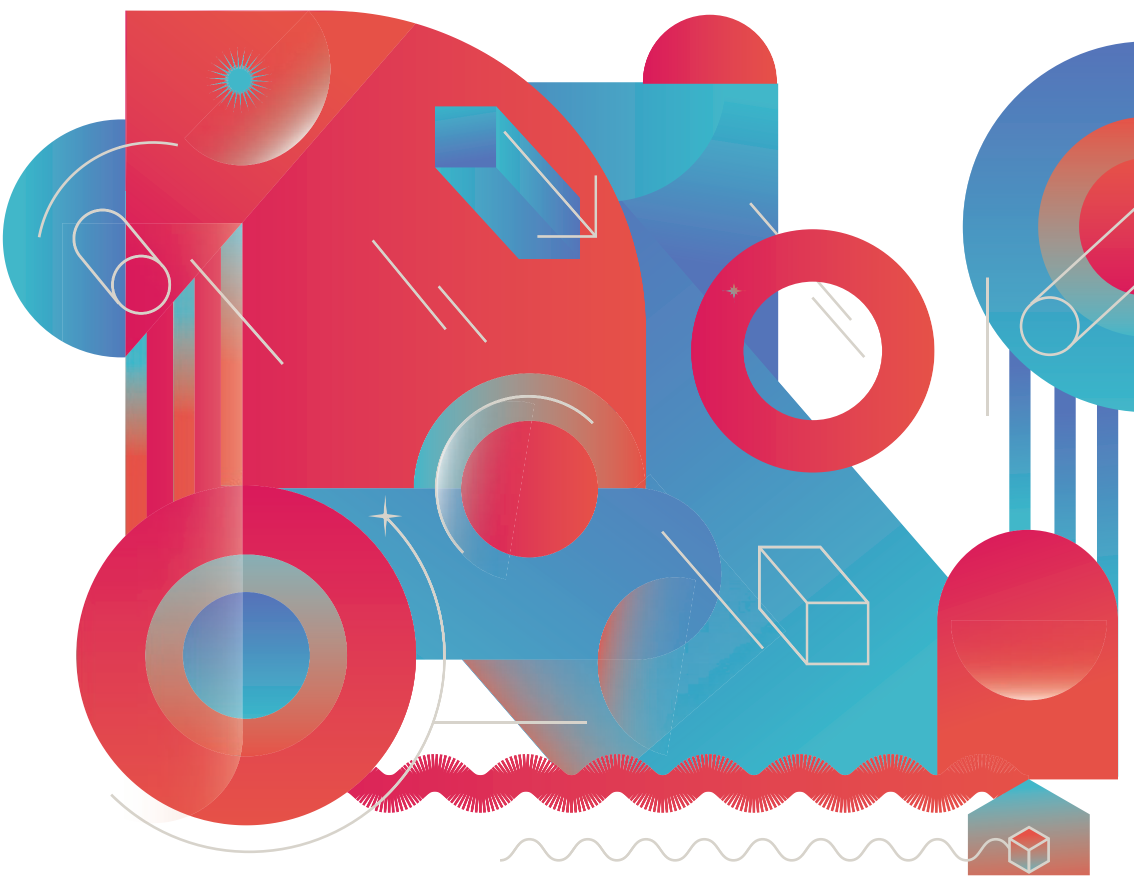
EMBEDDING INTELLIGENCE

Designerly reflections on AI-infused products

edited by Davide Spallazzo, Martina Sciannamè



D.I. FrancoAngeli
DESIGN INTERNATIONAL

The Design International series is born in 2017 as a cultural place for the sharing of ideas and experiences coming from the different fields of design research, becoming a place in which to discovering the wealth and variety of design, where different hypotheses and different answers have been presented, drawing up a fresh map of research in international design, with a specific focus on Italian design.
Different areas have been investigated through the books edited in these years, and other will be explored in the new proposals.
The Scientific Board, composed by experts in fashion, interior, graphic, communication, product and industrial, service and social innovation design, interaction and emotional design, guarantee the high level of the accepted books. After the first selection by the Scientific Board, the proposals are submitted to a double review by other international experts.

OPEN ACCESS
la soluzione FrancoAngeli

Il presente volume è pubblicato in open access, ossia il file dell'intero lavoro è liberamente scaricabile dalla piattaforma **FrancoAngeli Open Access** (http://bit.ly/francoangeli-oa).

FrancoAngeli Open Access è la piattaforma per pubblicare articoli e monografie, rispettando gli standard etici e qualitativi e la messa a disposizione dei contenuti ad accesso aperto. Oltre a garantire il deposito nei maggiori archivi e repository internazionali OA, la sua integrazione con tutto il ricco catalogo di riviste e collane FrancoAngeli massimizza la visibilità, favorisce facilità di ricerca per l'utente e possibilità di impatto per l'autore.

Per saperne di più:
http://www.francoangeli.it/come_pubblicare/pubblicare_19.asp

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati
possono consultare il nostro sito Internet: *www.francoangeli.it* e iscriversi nella home page
al servizio "Informatemi" per ricevere via e-mail le segnalazioni delle novità.

EMBEDDING INTELLIGENCE

Designerly reflections on AI-infused products

edited by Davide Spallazzo, Martina Sciannamè

D.I. FrancoAngeli DESIGN INTERNATIONAL OPEN ACCESS

Cover image by Sara Sciannamè

ISBN e-book Open Access: 9788835141914

Date of first publication: September 2022

Contents

Forewords

Davide Spallazzo, Martina Sciannamè
Department of Design, Politecnico di Milano

Artificial Intelligence – AI – is not new. Though, today it is living a moment of great hype. We now realize that AI permeates industrial products, services, and interfaces for public use, whether at home, at work, or in the public sphere.

The examples of AI-infused systems are uncountable. They range from smart thermostats – which adapt to their users' habits and optimize energy consumption while maintaining a high level of environmental comfort – to the most pervasive yet ephemeral systems that augment tools we use daily, such as Netflix or Amazon's recommendations.

These systems are characterized by a continuous learning process based on a statistical analysis of massive volumes of data. Accordingly, they learn and adapt their behavior over time.

Smart assistants, now permeating our domestic landscape, can be considered the first evident manifestation of AI-infused systems in our life. They represent a frontier edging closer to the notion generally associated with artificial intelligence: sentient robots capable of replicating human behavior.

Furthermore, these applications illustrate the dual interpretations of AI that have long enlivened the scientific debate: on the one hand, McCarthy's position emphasizes the development of a super-brain capable of simulating human behavior; on the other hand, Engelbart's position emphasizes the augmentation of human potential through AI (Winograd, 2006).

Engelbart's concept of augmentation views AI as a tool capable of augmenting rather than replacing human intellect and potential (Engelbart, 1962). This perspective is much more akin to that historically expressed in the field of Human-Computer Interaction (HCI) (Grudin, 2006), which placed a greater emphasis on the user than on the machine.

Winograd characterizes these two opposing positions as two distinct approaches to the subject of artificial intelligence. A rationalistic approach predicated on the conviction that fundamental aspects of thought can be captured in a formal symbolic representation. A design-oriented approach centered on people's interactions with their surrounding environments rather than modeling the mechanisms within intelligent systems (Winograd, 2006).

These concepts bring the realm of AI closer to Interaction Design, a discipline that takes a holistic and comprehensive view of interaction. It is defined as the mutual effect of individuals, artifacts, and the situations in which they exist; it involves discourse, connection, and social interaction (Kolko, 2011).

Given that Interaction Design is defined as the art of promoting human-to-human interaction through goods and services (Saffer, 2009), it is easy to see how AI looks to belong squarely within the purview of designers.

Moving from these premises, the edited book represents one of the first systematic studies on the design implications connected to the empowerment of industrial products with AI capabilities. It aims at providing a multifaceted view of AI-infused systems from a design standpoint. Far from being comprehensive and exhaustive of the subject matter, it is structured into five essays that problematize the embedding of intelligence within products.

The essays represent the point of view of the researchers involved in the Meet-AI research project, aimed at better understanding the nature of AI-infused systems from a design standpoint and their implications in terms of user experience.

Chapter 1 analyses smart assistants from a pure design/interaction design standpoint. Going through systematic reading of these devices' main characteristics, the chapter underlines a still immature embodiment of their potential and the need to enter a more mature stage of development, at least from a design standpoint.

Chapter 2 shifts the attention to the user experience – UX – enabled by AI-infused systems and its evaluation. It conducts a critical analysis of current UX evaluation methods against the peculiarities of

AI-infused systems and suggests the need for a specific method for such systems. It further advances the necessity to reconsider traditional UX dimensions.

Chapter 3 moves from these results, reporting the preliminary study conducted in the Meet-AI research project to elicit novel UX dimensions to better frame and assess the experience enabled by AI-infused systems.

Chapter 4 focuses on one of the most common characteristics of such systems: conversational interfaces. It contextualizes the potential disciplinary role of design in defining the peculiarity of these interfaces.

Chapter 5 concludes the book by opening the debate on the concept of meaningfulness. It problematizes the role of AI-infused systems in the lives of their owners not just as functional tools but also as relational mediators which shape the long-term aims, objectives, and behaviors.

References

Engelbart, D. (1962). *Augmenting human intellect: a conceptual framework*. PN.

Grudin, J. (2006). "Turing Maturing: The Separation of Artificial Intelligence and Human-computer Interaction". *Interactions*, 13(5), 54-57.

Kolko, J. (2011). *Thoughts on Interaction Design*, Second Edition (2 edition). Morgan Kaufmann.

Saffer, D. (2009). *Designing for Interaction: Creating Innovative Applications and Devices* (2nd edition). New Riders Pub.

Winograd, T. (2006). "Shifting viewpoints: Artificial intelligence and human–computer interaction". *Artificial Intelligence*, 170(18), 1256-1258.

1. AI-infused products so far. An analysis from a design standpoint

Mauro Ceconello
Department of Design, Politecnico di Milano

Since ubiquitous computing was first defined in 1991, computers have been pervading our everyday life. Nowadays, many AI-infused smart assistants are present in the domestic domain, and even robots-like devices, able to simulate human behavior, entered the market.

The contribution aims at analyzing and debating these devices in the design domain, focusing on their characters in terms of appearance, behavior, and interaction abilities.

The formal aspect couldn't be detached from function and the relationship with the context; these devices are designed for home use; therefore, shape, materials and finishing need to be considered. Some devices are mere facilitators for routine activities to be executed with vocal inputs or mediated through apps without exploring all the capabilities they could perform learning through continuous conversations with their owners.

From the analysis of products available on the market, the chapter emphasizes an incomplete exploration of the topic from a design point of view: a poor translation of functions into tangible shapes and a lack of interaction design basics, such as input and output modalities, feedback systems, and processes discoverability. The discussion frames the results in a broader review highlighting alternative paths for the design of future home virtual assistants.

1.1 AI-infused devices

Artificial Intelligence (AI) is the substance designers will be called on to engage within the subsequent years, according to Paola Antonelli's statement in an interview during the AI-Artificial

Imperfection roundtable held in New York in March 2018 (Antonelli, 2018). This comment by the curator of MoMA's design division may appear provocative, but it establishes a confluence between the areas of AI and design.

For years, AI has been quietly infiltrating systems and devices that we use daily. It has spread in several facets of people's lives, especially in four social domains: Healthcare, Education and Employment, Governance and Social Development, and Media and Entertainment (Burr, Taddeo and Floridi, 2020), and it is destined to become even more pervasive, similarly to electricity (Kelly, 2016). The wide availability of data-enhanced computational capabilities and improved algorithms has led to the diffusion of technologies integrating AI systems and the need for publicly acknowledging it. As a result, some artefacts have been created explicitly for this purpose. Multi-purpose home assistants synthesized the core characteristics of AI in the form of smart speakers that can learn from their owners' continual interactions. Furthermore, cheap domestic robots have reached the market in recent years, paving the way for a frontier closer to the idea generally associated with AI: sentient robots capable of simulating human behavior.

There are currently a few theoretical insights and experiments depicting the convergence of AI and Design. As a result, this contribution aims to further argument about the importance of design in the context defined above, paving the way for future research.

The concept is to frame the transitional period we are experiencing, looking for a path to follow in design research and practice, as the world is gradually moving towards the Ubiquitous Computing that Weiser envisioned in 1991 (Weiser, 1991). Computation is spreading throughout physical space and across multiple devices to build environments that help people in their everyday activities (Kaptelinin and Nardi, 2006). The focus is on virtual assistants, which summarize the first wave of AI's materialization in the domestic setting.

As far as 2020, studies have shown that already 320 million of these devices have been installed, and they are forecast to reach 640 million installed devices by 2024 (Canalys, 2020).

The study here reported analyzes ten devices that integrate AI agents (Table 1 lists the objects of analysis), the only ones on the market with the features indicated below. They are (i) versatile house assistants without a defined goal (e.g., senior help), (ii) especially developed as first-party hardware, (iii) already marketed or coming soon, and (iv) capable of controlling other smart home products. In the case of a product family, the research considers the first one released in its most recent version or its evolution. Some products are still under development, while others have not passed the prototype stage and have never reached the market. However, it is essential to include them in the analysis to provide a comprehensive picture of research on personal home assistants and give a complete description of the devices' functionalities and capabilities.

The study of virtual household assistants begins with a synthetic description of the devices. It continues by examining their interaction aesthetics (Petersen, Hallnäs and Jacob, 2008), which serves as proof of AI embodiment. In terms of preliminary considerations, it includes physical appearance, use, and interactivity; the purpose is not to evaluate the user experience of the analyzed items, but to illustrate state of the art in the field of domestic assistants to stimulate future thought.

As a result, neither UX evaluation procedures (Hassenzahl, 2001) nor usability testing (Brooke, 2013) has been used to evaluate the listed features. Official videos and documentation from the manufacturers were used to assess the products. Given the research's goal and the fact that some of the products are not still on the market, this method was deemed appropriate, ensuring homogeneity of the investigation where first-hand experiences would not be possible.

Table 2 lists the items that were analyzed (in order of the first model's release date) and spans the basic parameters that were used in the analysis: (i) physical appearance, (ii) input and output modalities, (iii) feedback mechanisms, and (iv) function discoverability (Saffer, 2010), with a focus on how proactive such artifacts are.

These factors were chosen to examine and define the selected items fundamentally from the aspect of product/interaction design. A choice motivated by the central hypothesis about domestic virtual assistants; acknowledging the great market success of some of them, this paper aims at showing how immature they are in some respects and encourage designers to take on a leading role in identifying a language and meaning besides a form.

1.2 Domestic smart assistants: a variegate picture

The panorama of domestic devices integrating AI capabilities varies in terms of appearance, dimensions, and functionalities.

In the following, we briefly describe each device object of analysis in its last release, whether it is available on the market, or its production has been interrupted or never passed the prototyping phase. The objective is to describe the devices to support the qualitative analysis that follows in the next section.

Amazon Echo	**Google Nest**	**Mji Tapia**	**Asus Zenbo**	**Jibo**
https://amzn.to/3ztwKZ5	https://bit.ly/3PTdlrP	https://mjirobotics.co.jp/en/	https://zenbo.asus.com/	https://jibo.com/
Apple HomePod	**Emotech Olly**	**InGen Dynamics Aido**	**Home Connect Mykie**	**Samsung Galaxy Home**
https://apple.co/3cZSDr	https://www.emotech.ai/	https://www.getaido.com/	https://bit.ly/3zutzkd	https://bit.ly/3zQLK4Q

Table 1.1 – List of the analyzed domestic assistants (with links to commercial webpages).

Amazon Echo

The Echo family is undoubtedly the most renowned in the field of domestic smart assistants and anticipated by far the other products. The traditional Echo speaker that entered the market in 2014 it's today substituted by less generic devices focusing on the quality of sound or the presence of a screen. For the sake of the study, we analyze the Amazon Echo Studio that derives directly from the first commercialized.

It is essentially a cylinder covered in grey fabric. A longitudinal cut in the lower part of the body favors the clearness of the sound. The top of the device presents a LED ring for light feedback and four physical buttons for adjusting the volume, turning the microphone off and prompting Alexa.

Google Home/Nest

Google Nest audio is the last release of the 2016 product (Google Home). It appears like a small parallelepiped with major measures in its front view, and a thinner lateral view. Available in light and dark grey, it is covered with fabric. The physical interface consists of a toggle to turn the microphone off on the back of the device.

Mji Tapia

Tapia is a presented as "home communication robot watching over your family", a personal companion to take care of elderly people or to help children in their learning activities. It appears like a glossy white egg, with a big screen in the upper part to show expressive cartoon-like eyes. It resembles a robot, but does not move any element, except for the eyes on the screen. It has a camera, to fulfil its monitoring aims and users can interact with it through the screen with the eyes.

Asus Zenbo

Zenbo is a robot, that moves autonomously across the house and can follow the users. Presented as a smart little companion, Zenbo was born as a B2C product, but it is now available only in the B2B market. Its latest release is the Zenbo Junior II, a multipurpose robot that can find an application in several fields: education, health, tourism.

It is composed by two white plastic volumes connected by a thin junction that make them resemble a body with neck and head. To increase this perception, the head has a flat surface with a screen, showing cartoon-like eyes and mouth. Colored LED mark the wheels on the lower body. Gridded air intakes on the main body and cuts the sides of the head increase anthropomorphism, resembling a belly or ears. Zenbo can move using the two lateral wheels, and its head can turn on its axis and move up and down. For interacting with Zenbo, user can users their use their voice or click on the screen (where Zenbo eyes and mouth are).

Jibo

Jibo has been designed to be a friend, a social robot. It is characterized by three-axis motor systems that make its body segments move in a fluid and natural way, increasing its potential in communicating with users.

Its physical appearance is dominated by pure shiny white volumes that rotate one over the other. The top part of the robot is spheric with a flat surface presenting a touch screen.

Everything is designed to make this little device resemble a small robot, including its only cartoon-like eye moving across the screen. The touch screen, beyond voice interface, is the main way of interacting with Jibo, which has also two onboard cameras for panoramic view. The robot is also sensitive to the touch on its head.

Apple Home Pod

Apple is not one of the big players in the field of smart home assistants. Its first device was first marketed in 2018, four years after Amazon.

The original Home Pod is not in the market anymore, and the current product is the Home Pod Mini. It appears like a small sphere in gridded plastic, available in five colors (white, black, blue, orange, yellow). The top of the sphere is cut to host an interactive screen, that shows a Siri-like color animation and hosts two controls to adjust the volume. It does not have a specific companion app, but it is fully integrated in the Apple ecosystem (Apple Home app).

Emotech Olly

Olly was presented at CES in 2017 but it never saw an industrial production. Nevertheless, we consider it relevant to include it in the analysis because (i) it firstly introduced movement without going towards anthropomorphic robots, and (ii) focused on personality.

Olly is doughnut-shaped, with the top surface animated by color LED squares. The main toroidal body can rise from a point on the base an rotate 360° on it. It can be controlled with the voice and through a companion app, but it is highly proactive thanks to contextual awareness.

InGen Dynamics Aido

Aido does not differ from other products like Jibo. Its peculiarity is that the main body may be extended with add-ons that provide the device with mobility, making it close to the Asus Zenbo. From an esthetic point of view, it appears in white shiny plastic. Its head presents a big screen with cartoon-like eyes.

With its one-ball transport systems, it acquires a very distinct anthropomorphic shape, with a long body, neck and head. It does not differ from other similar products except for a video projector.

Like Jibo, it sensitive to the touch not only on the screen, while the main touch screen is the major touchpoint for interaction.

Home Connect Mykie

Mikie is still in its concept phase, and it very likely to stay in that condition. Nevertheless, it is a good example of specialization of a domestic robot. It is developed to serve in the kitchen and its characteristics are designed towards that scope.

It appears in white or black plastic, with troncoconical body and head. The aim it that of creating an anthropomorphic device, and the impression is reinforced by big, stylized eyes. A LED ring at the basis of the body provides feedback on its functioning. Beyond the typical functionalities described for the other devices, Mykie adds a projector to ease following recipes.

Samsung Galaxy Home

Samsung is not a big player in the field of smart assistants. The Galaxy Home announced in 2018 never reached the market. In 2020 Samsung released the Galaxy Home Mini in South Korea, but not as an independent product but rather as a companion to a smartphone. The Galaxy Home Mini follows Amazon and Google with a small dark grey ball in fabric. The top of the sphere is cut, and hosts four buttons to adjust the volume, mute the microphone, and prompt the assistant.

DEVICES		Amazon Echo	Google Home/ Nest	Mji Tapia	Asus Zenbo	Jibo	InGen Dynamics Aido	Apple Home Pod	Emotech Olly	Samsung Galaxy Home	Home Connect Mykie
Year of production		2014	2016	2016	2017	2017	2017	2018	2018	2020	tba
PHYSICAL APPEARANCE	Simple Shape	●	●	●				●			
	Assembled solids				●	●	●		●	●	●
	Main Color & Material	G, F	G, F	W, P	W, P	W+G, P	W, P	ALL, P	B, P	B, P	W, P
DISCOVERABILITY	Proactive			●					●		●
	Non Proactive	●	●		●	●	●	●		●	
INPUT	Voice	●	●	●	●	●	●	●	●	●	●
	App	●		●	●	●	●	●	●	●	●
	Buttons	●	●	●	●	●	●		●		●
	Touch Surface		●					●		●	
	Touch Display			●	●	●	●				●
	Vision (camera)			●	●		●		●		●
	Touch					●	●				
OUTPUT	Other Device Action	●	●	●	●	●	●	●	●	●	●
	Audio	●	●	●	●	●	●	●	●	●	●
	Video			●	●	●	●				●
	Movement			●	●	●	●		●		●
FEEDBACK	Lights	●	●		●	●		●	●	●	
	Voice	●	●	●	●	●	●	●	●	●	●
	Movement			●		●			●		
	Display			●	●	●	●				●

Colors: B=black; BL=blue; G=gray; O=orange; W=white; Y=yellow. Materials: F=fabric; P=plastic.

Table 1.2 – Comparison of virtual assistants for residential use.

1.3 Analysis of domestic virtual assistants

Physical Attributes

In terms of main colors and materials, domestic virtual assistants have a striking resemblance. They cover the grayscale, with white being the most popular color, maybe influenced by science fiction productions. Black and grey, on the other hand, are derived from classic hi-fi aesthetics, implying their primary purpose as speakers. Plastic is the most common material however fabrics are gaining traction for a better integration into the domestic environment.

The most important element is the shape. The review focuses on two main formal paths: on the one hand, there are smart objects that follow simple and primary regular shapes; on the other hand, assembled bodies are built using geometric addition of solids (Van Onck, 1994), with five cases looking for a characterization as a human/animal-like shape with recognizable head and body. The separation of formal outputs also emphasizes a different functional goal, demonstrating that the planned activities are related to the final shape: smart speakers are the first manifestation of AI, and thus a natural outcome for a speaking technology. For example, the Home Pod is solely focused on audio quality and appears to be nothing more than a speaker. Devices that try to develop a social connection, on the other hand, take on a more anthropomorphic shape. Aido, for example, is regarded as a butler, and its height and shape contribute to this perception. As a result, functions have a significant impact on the object's overall, formal configuration, as well as its dimensions. Speaker-shaped assistants are only decorative, whereas those with social capabilities grow to the size of tiny household appliances or grow larger, in the event of a more realistic human simulation. The ability to move may also add to the embodiment of function: Olly does not have a simple or human shape, but its movements foster a sense of social connection.

Discoverability and Behavior

A typical feature of AI assistants is that they are built on machine learning and then evolve according to their owners' preferences. The major objective of home virtual assistants, once again, determines their behavior, emphasizing the distinction between smart speakers and domestic robots.

Most of the assistants, including smart speakers (Amazon Echo, Google Home, Apple HomePod, Samsung Galaxy Home) and, surprisingly, some robot-like assistants (Zenbo, Jibo, and Aido), are non-proactive, answering when prompted.

Three assistants, on the other hand, are proactive and offer information, activities, or content to their users based on their habits, moods, or anticipated needs. They integrate a camera (or more) to improve their proactivity by relating on more data to support their suggestions: they not only analyze noises or routines, but they also read body language and can grasp what their users are doing. Furthermore, they can recognize and reacting to their users as they pass by. Olly represents the highpoint of empathetic engagement and proactivity, developing and manifesting its own personality in response to that of its interlocutor. Non-proactive objects have a limited discoverability, and most of their functions are unknown to the user. Alexa is an example of this: it contains thousands of skills, the majority of which were created by third-party developers, but they are rarely known or used (White, 2018).

Interaction

According to Saffer's Systems Design (Saffer, 2010), the inquiry of interactivity has been limited to input, output, and feedback modalities. The main inputs and outputs are vocal, highlighting one of AI's most significant contributions to more human contact. Indeed, Natural Language Processing (NLP) has progressed to the point where it can now understand and respond to human queries. This provides the foundation for the rise of digital assistants, which are the most visible manifestation of this technological achievement. A through-app interaction is another characteristic that all the devices share as an input: its functionality spans from basic setup to complete functions (especially for all the smart speakers).

Furthermore, they are provided with buttons for specialized functions, such as a mute-microphone button and volume up/down buttons or a touch surface. Robot-like assistants, on the other hand, just require the starting one because they have a touch display as a face. Proactive devices use their cameras to interpret body language and gestures as inputs. Finally, being snuggled causes some of them (particularly Jibo and Aido) to emotionally respond to tactile inputs.

As an output, all the assistants support interaction with other home appliances and the providing of audio content such as online searches, music, podcasts. Video contents could be supplied from devices with a display or a projector (Aido and Mykie), which can be a reproduction of online sources or an enriched characterization of what they are saying or doing. Movement can be a response to a request – such as dancing (Jibo) or moving across rooms (Aido and Zenbo) – or just a reinforcement of communication – such as moving up and down while counting push-ups (Olly).

The feedback mechanism, however, is more important in an interpersonal-simulated connection. Referring the analysis to a framework for studying human-product interaction (Wensveen, Djajadiningrat and Overbeeke, 2004), AI-infused assistants provide almost no inherent feedback because physical actions are only required in a limited way; thus, only functional feedbacks (corresponding to the described outputs) and augmented feedbacks characterize current domestic assistants. Lights, spoken utterances, movements, and displays indicate the object's internal status while the function is processing. Almost all the objects, notably the speaker-based assistants, employ lighting systems to indicate their present condition, and their choreographies are like those found on other devices (for example, it is the case of Google Home bouncing dots). Except for Galaxy Home, all the brands have chosen colorful lights, which are most expressive in Olly's custom-built circular LED display: it really emphasizes the effort of creating a patent communication system using lights.

Another common form of feedback provided by digital assistants is voice: whether in a strict or more private manner, with a robotic or person-like tone, these devices let their users know if and what they have understood before providing the desired content. This type of feedback is particularly useful when interacting with anthropomorphic helpers since it offers the appearance of having a real discussion with someone rather than just talking to a machine.

These are fantastic opportunities for feedback for devices with a display and the ability to move: natural and fluid movements can follow the action or stress the robot's awareness of being addressed to its user – for example, tilting its head towards the one who is speaking.

Instead, displays are employed to display the bot's abstracted eyes, which animate in response to the user's inputs. Those feedbacks are the most successful at creating a natural relationship since the machine looks to be more alive and expresses its own individuality.

Feedforward (Wensveen, Djajadiningrat and Overbeeke, 2004) is a feature that is closely related to feedback and has a significant impact on engagement. In the case of virtual assistants, users (and maybe also the designers and/or programmers) have no idea about what will happen once a request is made. In many circumstances, they can only guess or anticipate a particular outcome, although they can be sure about the more common interactions (through app or buttons) or the basic and routine commands that they had executed various times. Otherwise, the interaction's output is unpredictable and – more importantly – not immediate, two factors that may detract from the interaction's perceived quality.

1.4 Reflections on AI-infused Assistants' Embodiment

The evaluation of the formal outcomes is inextricably linked to the object's main functions. As previously said, the shape of the devices is determined by their function and the relationship they have with their context. The concept of home has been inadequately considered when developing the appearance of those goods, except from its incorporation in the names of some products. All of them are technologies that may be freely placed in any location but have yet to establish a link with our domestic reality, a situation that may lead to people perceiving them as strangers in their homes. Perhaps a closer relationship would make it easier for users to comprehend the benefits of the assistant. Meanwhile, they are understood in terms of their resemblance to other, well-known objects or the abstract expectations fostered by our culture's theories. The fact that the most prevalent use of speaker-shaped intelligent assistants is to play music (Sciuto *et al.*, 2018) is not by happenstance, despite the development of NLP features is opening up new possibilities. On the other hand, devices with a humanoid figure are commonly defined as companions or home managers.

Then, when it comes to their functionalities, it appears that customers prefer a few well-known commands that effectively turn intelligent assistants into regular executors, especially considering smart speakers. As a result, in most cases they rarely use their various skills, as explained by (Kinsella, 2018) and validated by (Sciuto *et al.*, 2018) through quantitative and qualitative analysis of Alexa usage. The study by (White, 2018) emphasizes that discoverability in smart speakers is a significant issue for designers, and it suggests alternative solutions, such as context awareness and proactivity, to address the issue. This lack of discoverability and affordance (Norman, 1988; Gibson, 2014) could be owing to a lack of skill embodiment into tangible things, which is typically associated with a rejection of sensory curiosity and pleasure (Jordan, 2000; Marti, 2010). In fact, consumers are precluded from perceiving the acts' fundamental effect when they do them. The framework designed by (Wensveen, Djajadiningrat and Overbeeke, 2004), might be used to pursue a more intuitive interaction. The authors propose reinstating natural couplings between actions and reactions based on six different parameters in order to improve its quality. At the present, the most promising and simple-to-integrate solutions are connected to interaction expression: the output modality might be a reflection of the conditions under which the request is made, such as the user's mood or the time of day (Pavliscak, 2018). Only three of the evaluated devices (Olly, Mykie, and Tapia) now have these features, but they might add real value to the assistants' general interaction and usability, as well as inform their shape.

One theory is that a humanized appearance and behavior can make an AI-infused device appear to be a true domestic assistant. Its likely relationship with a human person creates the conditions for a more natural contact, making it easy to envision that the object has a wide range of skills and will play a proactive part in our daily lives. Yet, as humans, we have a tendency to anthropomorphize everything (Pavliscak, 2018), and functions – made clear through proactivity –, shape, and movements seems to facilitate users in terms of discoverability and hence, engagement. Furthermore, the behavior of the devices provides them with well-defined identities.

In light of this reasoning, another issue to consider is that some anthropomorphic assistants have a characteristic known as mutuality of effects systems (Marti, 2010), which means they are sensitive to perceptual crossing. It is not just the user's responsibility to see the object of interaction; it is also the user's responsibility to sense the individual who will trigger it, preparing and announcing this awareness. Even without the use of an interface, the interaction becomes expressive, embodied, and responsive in this way.

Deeper linkages between shape, capabilities, and behaviors should be built based on a more reciprocal interaction between items and people, with the goal of figuring out what it is all about.

Faced with the complexity of these technologies and their significance in everyday life, designers must also consider the materialization of a large number of abilities, particularly if the only means of interface is dialogue. In fact, most of the examined assistants include a companion app that can be accessed via smartphone and allows for product customizing as well as traditional skill browsing. However, it opposes one of the principles of natural interaction: removing interfaces to make the artifact's mechanism directly available to system users (Dourish, 2001).

Devices that integrate a tablet as a head, which may be actively utilized for input and output, allow for a similar, if perhaps less obvious, engagement. A debatable solution that necessitates an obviously inadequate examination into the meaning of interaction.

Similarly, using voice as the primary concrete expression of domestic assistant intelligence does not automatically make them true conversational agents, nor does it make it easier to utilize their potential. Every maker has chosen to implement additional input mechanisms ranging from simple buttons to cameras and elaborate nested menus to be viewed through bespoke apps, demonstrating this condition. To put it another way, conversational agents are commonly conceived of as the tip of an iceberg that will make interaction more human and friendly, yet they still lack a clear definition.

1.5 Virtual Assistant Design Scenarios in the Future

As indicated in the introduction, the initial wave of AI materialization in the home sphere is still deeply immature in terms of function, language, and meaning (Kolko, 2011).

Designers are tasked with balancing form with ever-increasing functionality, finding a balance between conveying and obfuscating them. The current situation continues to reflect the contradiction that pervades

AI debate. On the one hand, smart speakers turn AI into aesthetically pleasing devices that we can utilize in our homes. Objects of various shapes, on the other hand, strive to be viewed as valuable humanoids who can assist in everyday living. A paradox that pervades all of the study's components – in terms of physical appearance, conduct, and interaction – and shows a still immature design reflection and the need to develop a unique language.

Despite the fact that the argumentation only addresses a few early factors and presents limits that are best addressed later, certain preliminary considerations may arise. The perception is that the sector is still going through the intoxicated period that comes with the debut of any new technology (Antonelli, 2018). Domestic assistants appear to be stuck in the toy phase, unable to make actual contact with reality (Levinson, 1977), and design should play a key part in directing a human-centered shift to meaningful products. As a result of current perspectives, AI discourse is being integrated into the discipline of Interaction Design. They all share the goal of enabling more natural contact, rather than relying on display-mediated interfaces.

Then, when it comes to converting AI into a real and domestic form, shape cannot be isolated from function and meaning from a human-centered and holistic perspective. According to what has been seen, AI-infused objects could simplify their functional structure so that interactions are more rapid and meaningful in terms of experience and utility. In this regard, there are three possible outcomes: (i) virtual assistants could evolve into self-standing objects, perhaps with more specific and limited functions that better translate into a clearly recognizable form; (ii) they could evolve into more accurate humanized robots, taking the role of actual people at the service of others; or (iii) they could be completely dematerialized and spread across other existing appliances and devices throughout the physical environment, with a sensibility for their location and proximity; or (iv) they could be completely dematerialized and spread across other existing This, in turn, may lead to more questions about the embodiment, its scale (product or environment), and how the enhanced functionality will be expressed.

Clearly, the study presents several limitations. It considers only a narrow number of products, even if they are representative of all those that responded to the selection criteria. A larger study may cover the whole spectrum of goods from the companies evaluated here, as well as third-party products that integrates AI agents. Furthermore, this initial argumentation begins with the fundamental characteristics of a product, but it may expand to include other design issues such as UX, or interdisciplinary reflections such as social/psychological implications, such as the emotional response they may foster through interaction, as well as experimental studies.

References

Antonelli, P. (2018). *AI Is Design's Latest Material*. Available at: https://design.google/library/ai-designs-latest-material/ (Accessed: 6 February 2019).

Brooke, J. (2013). "SUS: a retrospective". *Journal of Usability Studies*, 8(2), pp. 29-40.

Burr, C., Taddeo, M. and Floridi, L. (2020). "The Ethics of Digital Well-Being: A Thematic Review". *Science and Engineering Ethics* [Preprint]. doi:10.1007/s11948-020-00175-8.

Canalys (2020). *Global smart speaker market 2021 forecast*. Available at: www.canalys.com/newsroom/canalys-global-smart-speaker-market-2021-forecast?time=1623256196 (Accessed: 9 June 2021).

Dourish, P. (2001). *Where the action is: the foundations of embodied interaction*. Cambridge, Mass: MIT Press.

Dove, G. *et al*. (2017). "UX Design Innovation: Challenges for Working with Machine Learning As a Design Material". In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI '17), pp. 278-288. doi:10.1145/3025453.3025739.

Engelbart, D. (1962). *Augmenting Human Intellect: A Conceptual Framework*. PN.

Gibson, J.J. (2014). *The Ecological Approach to Visual Perception: Classic Edition*. 1st edn. Psychology Press. doi:10.4324/9781315740218.

Grudin, J. (2006). "Turing Maturing: The Separation of Artificial Intelligence and Human-computer Interaction". *Interactions*, 13(5), pp. 54-57. doi:10.1145/1151314.1151346.

Hao, K. (2018). *What is machine learning?, MIT Technology Review*. Available at: www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/ (Accessed: 6 February 2022).

Hassenzahl, M. (2001). "The Effect of Perceived Hedonic Quality on Product Appealingness". *International Journal of Human-Computer Interaction*, 13(4), pp. 481-499. doi:10.1207/S15327590IJHC1304_07.

Jordan, P.W. (2000). *Designing Pleasurable Products: An Introduction to the New Human Factors*. CRC Press. doi:10.1201/9780203305683.

Kaptelinin, V. and Nardi, B.A. (2006). *Acting with technology: activity theory and interaction design*. Cambridge, Mass: MIT Press (Acting with technology).

Kelly, K. (2016). *How AI can bring on a second Industrial Revolution*. Available at: www.ted.com/talks/kevin_kelly_how_ai_can_bring_on_a_second_industrial_revolution (Accessed: 11 July 2020).

Kinsella, B. (2018). *Amazon Alexa Now Has 50,000 Skills Worldwide*, works with 20,000 Devices, Used by 3,500 Brands, Voicebot.ai. Available at: https://voicebot.ai/2018/09/02/amazon-alexa-now-has-50000-skills-worldwide-is-on-20000-devices-used-by-3500-brands/ (Accessed: 8 February 2022).

Kolko, J. (2011). *Thoughts on Interaction Design* – 2nd Edition. Morgan Kaufmann. Available at: www.elsevier.com/books/thoughts-on-interaction-design/kolko/978-0-12-380930-8 (Accessed: 6 February 2022).

Levinson, P. (1977). "Toy, Mirror, and Art: The Metamorphosis of Technological Culture". *ETC: A Review of General Semantics*, 34(2), pp. 151-167.

Marti, P. (2010). "Perceiving While Being Perceived". *International Journal of Design*, 4(2).

Norman, D.A. (1988). *The psychology of everyday things*. New York: Basic Books.

Pavliscak, P. (2018). *Emotionally intelligent design: rethinking how we create products*. First edition. Sebastopol, CA: O'Reilly Media Inc.

Petersen, M.G., Hallnäs, L. and Jacob, R.J.K. (eds.) (2008). "Introduction to special issue on the aesthetics of interaction". *ACM Transactions on Computer-Human Interaction*, 15(4), pp. 1-5. doi:10.1145/1460355.1460356.

Saffer, D. (2010). *Designing for interaction: creating innovative applications and devices*. 2nd ed. Berkeley, CA: New Riders (Voices that matter).

Sciuto, A. *et al*. (2018). "'Hey Alexa, What's Up?': A Mixed-Methods Studies of In-Home Conversational Agent Usage". In *Proceedings of the 2018 Designing Interactive Systems Conference*. New York, NY, USA: ACM (DIS '18), pp. 857-868. doi:10.1145/3196709.3196772.

Van Onck, A. (1994). *Design: il senso delle forme dei prodotti*. Milano: Lupetti.

Weiser, M. (1991). "The Computer for the 21st Century". *Scientific American*, 3(265), pp. 94-104.

Wensveen, S.A.G., Djajadiningrat, J.P. and Overbeeke, C.J. (2004). "Interaction frogger: a design framework to couple action and function through feedback and feedforward". In *Proceedings of the 2004 conference on Designing interactive systems processes, practices, methods, and*

techniques – DIS ’04. the 2004 conference, Cambridge, Mass: ACM Press, p. 177. doi:10.1145/1013115.1013140.
White, R.W. (2018). “Skill Discovery in Virtual Assistants”. *Commun. ACM*, 61(11), pp. 106-113. doi:10.1145/3185336.
Winograd, T. (2006). “Shifting viewpoints: Artificial intelligence and human–computer interaction”. *Artificial Intelligence*, 170(18), pp. 1256-1258. doi:10.1016/j.artint.2006.10.011.
Yang, Q. *et al*. (2018). “Investigating How Experienced UX Designers Effectively Work with Machine Learning”. In *Proceedings of the 2018 Designing Interactive Systems Conference*. New York, NY, USA: ACM (DIS ’18), pp. 585-596. doi:10.1145/3196709.3196730.

2. User Experience and AI-infused products. A wicked relationship

Davide Spallazzo
Department of Design, Politecnico di Milano

The current panorama of AI-infused devices portrays a significant dominance of first-party smart speakers, which appear to be the first massive embodiment of AI in the domestic landscape. These devices are nothing more than discreet ornaments, looking at their simple physical appearance. Although, the simple appearance betrays a complexity determined by numerous features that make such products challenging to analyze from a UX point of view.

The main evident characteristic is that they are not just "simple products" but ecosystems consisting of several interfaces and touchpoints. Most of them integrate multiple interfaces – namely physical, digital, conversational – sometimes overlapping.

The second element of complexity resides in their technological core, based on learning algorithms. Therefore, the same device can provide different outputs at the same input over time, a condition that can affect the user experience.

To increase the complexity of these devices, at least from a UX standpoint, there is the fact that their real potential is rarely exploited by most users, which mainly uses routine actions such as reading news, weather forecasting, and controlling simple home appliances.

Accordingly, the chapter frames the wicked relationship between user experience and AI-infused products. Moving from the three identified elements of the complexity of AI-infused products, it advances reflection on how it could be possible to analyze these products from a UX standpoint.

2.1 Introduction

Continuous technological advancement has made it feasible to design ubiquitous systems capable of emulating human behavior, resulting in personal assistants that find noteworthy application in many sectors (e.g., medicine, security, transport, education). Conversational AI-based agents such as Amazon Alexa are spread in millions of houses, populating the domestic environment with products powered with Artificial Intelligence (AI) capabilities. AI-infused products can adapt autonomously to the context and users' needs, entertain a dialogue with them, recognize them, and track/anticipate their behavior.

This market has already known a great success. However, it is expected to exponentially grow in the next few years, reaching billions of products integrating smart assistants, being them first-party or third-party hardware (*Canalys Newsroom – Rise of Alexa, Google Assistant and Siri to drive US smart assistant-compatible device base to 1.6 billion in 2022*, no date). Even if not so significant in market share, we may also add to the count personal and domestic robots (e.g., the renowned robot Pepper) provided with AI-based conversational agents, which accompany users in daily tasks.

These goods, which can be considered emblematic of the first wave of AI embodiment in the domestic setting, drew much interest. Nonetheless, the HCI community and the design discipline at large have only recently developed a comparable interest.

The design world, and academia, in particular, has yet to address this issue thoroughly, although AI has been touted as a new material for designers (Holmquist, 2017; Antonelli, 2018). Some studies have analyzed machine learning as a design subject (Dove *et al.*, 2017), the use of virtual assistants in everyday life (Sciuto *et al.*, 2018) and their aesthetic-functional dimensions in the domestic realm (Spallazzo, Sciannamè and Ceconello, 2019). Similarly, other studies coped specifically with conversational interfaces (Vitali and Arquilla, 2019) or reviewed ways for evaluating the user experience enabled by voice-based interactions (Kocaballi, Laranjo and Coiera, 2018).

However, studies investigating the user experience enabled by AI-infused products are needed. Indeed, these devices are frequently perceived as gadgets or toys (Levinson, 1977) that tickle the users' craving novelties rather than playing a significant role in their lives.

Likewise, they are not exceptionally noteworthy in terms of interaction quality, sometimes generating frustration and an essential use unable to unlock their potential entirely (Sciuto *et al*., 2018).

To the best of our knowledge, no systematic attempt has been made to frame the UX entailed by AI-infused products, coping with these devices' complex nature holistically. Accordingly, we consider it relevant to understand and frame the User Experience (UX) they entail and review current UX assessment methods to understand their adequacy for AI-infused systems.

2.1.1 Understanding AI-infused systems

Personal assistants, robots, and self-driving cars are entirely changing the interaction paradigm we are used to. They are no more tools used by humans as extensions of their bodies or minds. AI-infused systems are perceived as *counterparts* (Hassenzahl *et al*., 2020). As such, they introduce a shift from the paradigm of embodiment (Dourish, 2001) to *alterity*, as stated by Hassenzahl and colleagues who defined these products *otherware* (Hassenzahl *et al*., 2020).

Perceiving interactive products as others and not as users' extensions entails introducing new interaction paradigms. We do not interact with voice assistants; we converse with them. Relying on the robustness of the system is no more enough. We must trust them since most of the operations (deep-learning) AI-infused products perform are frequently opaque to end-users, and their outputs are somehow unpredictable. Novel forms of interaction imply a different relationship between humans and machines and necessarily may impact the perception of machines and the user experience they enable.

AI-infused systems learn, reflect, talk, and clearly show their agency to end-users through proactivity. Referring to the three kinds of agency – conditional, need-based, delegated – proposed by Kaptelinin and Nardi (Kaptelinin and Nardi, 2009), these systems go beyond the delegated agency. Indeed, they can betray users' expectations or act autonomously, or, better, choose the best solution for a problem they have been delegated to solve.

Products integrating AI capabilities (e.g., Amazon Echo family) may appear nothing more than inobtrusive ornaments. However, it is evident they betray an intrinsic complexity, increased by features that make such products challenging to analyze from a UX point of view.

The primary distinguishing feature is that they are not only "simple goods", but rather ecosystems with several interfaces and contact points. Most of them combine various interfaces – physical, digital, and conversational – that occasionally overlap in terms of functionality. Being an ecosystem also entails communicating with – and even controlling – other devices. This is especially true for domestic devices, which frequently act – or are supposed to act – as a control hub for home entertainment and an ever-increasing number of connected objects in complex dynamic environments like our homes.

Secondly, their technical heart, based on learning algorithms and neural networks, adds a second layer of complexity. It implies that the same device might provide multiple outputs at the same input over time, a circumstance that could influence UX evaluation in traditional methods.

To add to the complexity of these products, at least from the standpoint of UX evaluation, their true potential is rarely explored by most users. They primarily perform ordinary behaviors such as reading news, weather forecasting, and operating simple home appliances (Sciuto *et al.*, 2018; White, 2018). A condition that may prompt, on the one hand, a conversation on discoverability, as advocated by White (2018), and, on the other, a more in-depth assessment of the role of such systems in our daily lives.

2.1.2 Aims and methodological approach

In line with the reflections expressed above, the hypotheses here advanced are that (i) the user experience enabled by such systems can hardly be described and framed with current assessment methods, and (ii) new qualities of the user experience must be considered. Consequently, a (iii) novel assessment method is required to evaluate the experience enabled by AI-infused systems holistically. Accordingly, the following sections aim to answer three main research questions: *[RQ1] Are current UX assessment methods enough for AI-infused products? [RQ2] Are new UX dimensions needed for these products? [RQ3] What characteristics the new method could have?*

Secondary research has been performed to answer these questions with the specific aims of (i) tracking the evolution of UX assessment and (ii) understanding current methods and their adequacy in evaluating AI-infused systems. The first analysis is confined to works published between 2000 and 2020 in the ACM Digital Library and Springer Link, retrieved through the query strings *UX AND evaluation* and *UX AND assessment*. It focused on the few articles tracing the evolution of UX and its assessment over time summarized in Figure 2.1.

Authors	**Sample of UX eval. meth.**
(Vermeeren *et al.*, 2010)	96
(Bargas-Avila and Hornbæk, 2011)	66
(Lachner *et al.*, 2016)	84
(Rivero and Conte, 2017)	227
(Pettersson *et al.*, 2018)	100

Fig. 2.1 – Most relevant articles tracing the evolution of UXEM and relative sample of UX studies.

A subsequent and more comprehensive study followed a snowball sampling approach moving from the articles listed above, to deepen specific methods. At the same time, the long list of methods collected by the researchers of the Allaboutux.org website served as a reference point to have a broad overview of UX evaluation methods.

2.2 User Experience and its assessment

Since the term User Experience (UX) was introduced around the turn of the millennium, academics have sought novel approaches to understanding and assessing the quality-in-use of interactive systems (Bargas-Avila and Hornbæk, 2011). Leading this change, which began in the HCI sector, was the perception that contemporary assessment techniques were overly focused on system usability and efficiency, missing more general but no less essential elements such as quality, pleasantness, and meaning.

This turn has been recalled by Hassenzahl, Burmester and Koller in a recent work (Hassenzahl, Burmester and Koller, 2021) that traces back twenty years of reflection on UX, moving from their seminal article (Burmester, Hassenzahl and Koller, 2002), which challenged the univocal association between usability and user experience.

Throughout the years, numerous studies have advanced alternative methodologies for understanding, conceptualizing, and supporting the creation of meaningful experiences with interactive systems (Hassenzahl and Tractinsky, 2006), awakening a discussion based on differing epistemological perspectives. Simultaneously, the scope of the studies steadily expanded from the exclusive instrumental and task-oriented evaluation of usability to encompass, to mention a few, pleasure (Jordan, 2000), positive emotions (Norman, 2004), and aesthetics (Tractinsky, Katz and Ikar, 2000).

The academic world actively adopted this unique perspective, either alone or in collaboration with companies and consulting firms, resulting in many publications dealing with UX evaluation of interactive systems, whether they are industrial products, digital interfaces, or services.

The proposed solutions ranged from specific tests (Schmettow, Noordzij and Mundt, 2013) to the deployment of evaluative tools (Lugmayr and Bender, 2016; Minge *et al.*, 2017; Sivaji, Nielsen and Clemmensen, 2017; Almeida *et al.*, 2018; Zhou *et al.*, 2019; Maguire, 2019) to more holistic and broad methods (Obrist, Roto and Väänänen-Vainio-Mattila, 2009; Kujala *et al.*, 2011; Jiménez *et al.*, 2012; Otey, 2017).

In 2010, Vermeeren *et al.* (2010) gathered 96 relevant and unique UX assessment techniques from academia and industry and classified them according to various criteria. The investigated approaches were generally used on digital interfaces (such as websites and mobile applications) and fully functional systems or working prototypes, interfering at an advanced stage of development.

Bargas-Avila and Hornbæk (2011) did a similar wide-ranging critical study concentrating on empirical methodologies for the evaluation of the UX, with an in-depth investigation of 51 publications out of 1254 arising from digital library research. Their research revealed a high frequency of digital interfaces as examined systems and a propensity to evaluate the overall UX and UX aspects such as affect, emotion, enjoyment, aesthetics, attractiveness and hedonic attributes. They also clearly demonstrated that researchers generally utilized self-developed questionnaires and that classic inquiry techniques like questionnaires, interviews, and diaries were the most often used to obtain qualitative data.

Rivero and Conte (2017) conducted a thorough review of UX evaluation methodologies for digital interfaces in 2017. Their findings are similar to those of the previous studies: the study methods are largely traditional (e.g., questionnaires, observation) and used with groups of final users in a controlled setting during or after the usage experience. Differentiating from the previous ones, this study emphasizes the importance of quantitative data as the primary output of the UX assessment.

One year later, Pettersson and colleagues performed a comprehensive review of 100 academic articles published between 2010 and 2016 describing empirical investigations on user experience evaluation focusing on reliability. The authors explicitly addressed the triangulation of techniques in user experience assessment (Pettersson *et al.*, 2018). The review of the state-of-the-art highlights findings that are comparable to those found in the studies mentioned above. The most usually addressed UX characteristics are overall UX and pragmatic features (usability), digital interfaces are the most commonly evaluated items, and self-developed surveys are frequently used as an inquiry method. Additionally, the authors highlight a current trend toward employing – and triangulating data from – various ways of inquiry to better comprehend and contextualize the results. The study emphasizes four open topics for future UX research, listing, among others, the necessity of adapting to ever-evolving technologies and non-human intelligence. The open question is how to perform this adaptation. Accordingly, in the following section, we analyze existing UX assessment methods that may adapt or not to the unique peculiarities of AI-infused systems and try to envision general traits of a novel method.

2.3 Coping with complexity

A first point to address is AI-infused systems' unique nature: ecosystems rather than single, self-standing products. Taking domestic smart speakers as a reference for the discussion, we may state that several devices – e.g., light bulbs, doorbells, cameras, and thermostats – can currently be linked and controlled, creating an integrated system with the smart speaker as its hub.

The user experience allowed by such an ecosystem (i) will evolve and necessarily (ii) involve a variety of touchpoints, whether physical, digital, or based on conversational interfaces.

Beyond this multiplicity of touchpoints in the ecosystem, the devices themselves can be considered complex artifacts from an interaction standpoint. Usually, smart speakers integrate physical buttons (e.g., volume, mute, activation) and are all equipped with conversational interfaces that duplicate more or less the same functions.

In addition, a companion smartphone app enables users to configure the device and conduct a variety of tasks, such as broadcasting music or controlling linked devices.

To add complexity from a UX standpoint, these products usually provide feedback on their current state through different means, whether with colored LED or the assistant's voice.

Looking at the current panorama of UX assessment methods, we could investigate different aspects separately. SUS (Brooke, 1995), Kansei (Schütte *et al.*, 2006) and AttrakDiff (Hassenzahl, Burmester and Koller, 2003) can be used to assess the product's usability/pragmatic or hedonic aspects on the hardware side. Similarly, we may use the well-known Nielsen and Molich heuristics (1990) to evaluate the companion app's user interface and eventually rely on specific methodologies to evaluate the user experience of conversational interfaces (Pyae and Joelsson, 2018; Maguire, 2019).

Still, it would be problematic to conduct a comprehensive user experience assessment to render a holistic view of the quality of the interaction enabled by the devices and their ecosystem.

This first insight may suggest the need for a novel approach to deal with the complexity that may consist in (i) a custom technique integrating existing ones or (ii) a generic method capable of capturing the spirit of the whole experience. The first approach may require the evaluation of the peculiar characteristics of distinct interfaces and touchpoints and, consequently, a broad and modular method to be tailored to each unique scenario under assessment.

The second option appears to be the most popular, based on critical reviews of existing UX rating methodologies. According to Bargas-Avila and Hornbæk's study, 41% of techniques evaluate generic user experience (UX) (Bargas-Avila and Hornbæk, 2011), and the ratio increases to 56% in a more recent examination by Pettersson *et al.* (2018).

The second point of complexity is that AI-infused products rely on learning algorithms. Accordingly, the same system might produce different outputs for the same input over time.

So, time acquires a great relevance in the user experience since the more users interact with the system, the more accurate it should become in anticipating/answering needs and requests.

As a result, assessing the experience over single episodes or even a whole day may produce unreliable data. In contrast, longitudinal research could presumably provide a more accurate picture of the user experience.

The study by Vermeeren and colleagues (2010) outlines a variegated picture of UX assessment methods regarding the periods of the experience objects of analysis. They range from unique snapshots/episodic activities to long-term usage through episodes/specific tasks. Looking at the percentage emerging from their analysis, 36% of the methods object of study already assesses systems' performances in everyday life over a long-term interaction. A different result is highlighted in the more recent work by Rivero and Conte, who point out a much more limited percentage – 6.6% – of methods assessing long-term interactions (Rivero and Conte, 2017). Regardless of the percentage, evaluating the user experience with a longitudinal approach is standard practice.

This second insight may indicate the need to choose/create a method to assess an experience over a long time and, eventually, track how the quality of the experience may evolve (Karapanos *et al.*, 2009).

The third level of complexity in assessing the UX of AI-infused systems is that they are relatively new and may present inherent flaws in the usage. Two recent studies focusing on Amazon's smart speakers highlight, for example, that users rarely acknowledge and understand the actual potential of such devices. This condition may be due to poor discoverability and a lack of proactivity (Sciuto *et al.*, 2018; White, 2018).

This condition raises further doubts on the ability of current UX evaluation methods to assess AI-infused systems holistically: if the user does not fully recognize the systems' agency and potential, how can she fully perceive her experience as unsatisfactory or limited?

Furthermore, an assessment approach focused exclusively on using existing systems/working prototypes, even if correctly developed, may disclose just a few issues that can be addressed through the iterations of the last stages of the design process.

Going back to the studies on Amazon's devices, proactivity can soothe some of the difficulties associated with discoverability (White, 2018). Nevertheless, adding proactivity to a system is not a simple task that can be performed in the last adjustments before the product's launch on the market. It is also true that using potentially powerful devices such as smart speakers for simple weather forecasts or as music players may reveal weaknesses in the reflection on the product's intended usage in the domestic environment. These reflections must be done early in the design process.

Looking again at the study by Vermeeren *et al.*, it emerges that only a low percentage – 25% – of current UX evaluation methods may be used in the early stages of product development. The majority of methods focus on working systems (Vermeeren *et al.*, 2010). Accordingly, a method to be applied at different stages of the design process may be considered a good option for AI-infused systems.

The last element of complexity here discussed relies on the very nature of AI-infused products, that of being *otherware* (Hassenzahl *et al.*, 2020).

The question is whether current UX evaluation methods are equipped to assess something not perceived as an extension of users but rather as a counterpart. Looking back at the comprehensive analyses of UX evaluation methods, they provide a very traditional portrait in terms of objects of study.

The analysis by Vermeeren *et al.* (2010) highlights an important focus on web services (81%), mobile software (77%) and PC software (76%). Very similar results emerge from the study by Bargas-Avila and Hornbæk (2011), which points out a tendency to assess products such as mobile applications and phones, audio, video, TV applications. Interestingly they also underline that 21% of the assessed methods regard art (e.g., audio photography, interactive canvas). These two studies come far before the spread of AI-infused products and do not track any methods evaluating specifically such systems.

A more recent study by Rivero and Conte (2017) highlights a significant percentage of methods broad enough to assess any type of interface (33.9%). Their results further portray attention to web (13.7%) and mobile (8.8%) applications.

One year later, Pettersson *et al.* (2018) provide a fragmented image of products studied in UX research. In their ranking the most assessed systems are mobile apps (15%), interactive games (13%) – that make their first appearance –, webtools (12%) and websites (10%). At the same time, a small percentage (4%) of methods address connected/IoT services.

The analysis of the criticalities of AI-infused systems against current UX evaluation methods suggests a clear answer to [RQ1] *(Are current UX assessment methods enough for AI-infused products).* So far, UX research has not explicitly addressed AI-infused systems and the creation of a novel, bespoke method to address such systems is explicitly needed.

2.4 Understanding UX dimensions

The reflections above indicate the necessity of a novel UX evaluation method to assess the experience enabled by AI-infused systems. Following this line, it is essential to understand if the dimensions of the user experience assessed so far are enough for these systems or if new dimensions must be elicited.

UX dimensions are at the heart of any evaluation process. More than any other characteristic, they evolve through time, reflecting an ever-changing form of comprehending and framing user experience.

Looking back at systematic reviews of UX evaluation methods (Bargas-Avila and Hornbæk, 2011; Pettersson *et al.*, 2018), we can assert that generic UX is by far the most researched dimension, despite its ambiguous description. According to Bargas-Avila and Hornbæk (2011), generic user experience is generally evaluated qualitatively and refers to an overall perception of the experience. Similarly, Pettersson *et al.* (2018) describe generic UX as a broad concept that may be evaluated holistically. Focusing on the whole experience may help manage complexity in multi-interface and multi-touchpoint ecosystems, like those we are studying. However, it may also be constraining and incapable of providing valuable inputs throughout the design process and go to the core of AI-infused systems.

Pragmatic aspects – namely, usability – rank second in the research by Pettersson *et al.* (2018). However, they were not even included as a

UX dimension seven years before in Bargas-Avila and Hornbæk (2011), underlining a deliberate early dissociation from usability concerns to distinguish the new methods. Nevertheless, it must be noted that some of the most popular AI-infused systems do not fully address usability concerns due to a lack of in-depth reflection on basic interaction design principles such as input, output, and feedback modalities. For instance, a recent analysis of AI-infused devices (Spallazzo, Sciannamè and Ceconello, 2019) shows that these devices provide almost no inherent feedback, as physical actions are required in a limited manner; thus, only functional and augmented feedbacks appear to characterize current domestic assistants. Accordingly, pragmatic questions cannot be given for granted and should be evaluated, particularly in the early phases.

Nonetheless, there is little question that usability, emotion/affect, and enjoyment play a significant role in the UX. Not surprisingly, these aspects are the most often examined after general UX (Bargas-Avila and Hornbæk, 2011; Pettersson *et al.*, 2018). Moreover, reflecting on the emotional reaction to objects/environments asks for an examination of the aesthetics/hedonic aspects of items.

Previous studies on the embodiment of AI capabilities in the domestic environment identify an immature perspective on the embodiment of AI in the domestic realm (Spallazzo, Sciannamè and Ceconello, 2019) and the resulting necessity to incorporate an aesthetic dimension into the evaluation technique, particularly during the early prototype phases.

As demonstrated in Bargas-Avila and Hornbæk (2011) and in Pettersson *et al.* (2018), the UX dimensions discussed above are more or less handled by the methodologies current assessment methodologies do not cover dimensions likely relevant to AI-infused systems.

At the same time the impression is that these dimensions, despite essential, seem unable to track the very nature of these devices. So, coming to *[RQ2] Are new UX dimensions needed for these products,* we may state that traditional UX dimensions must be included in the analysis, but they are not enough to assess the core qualities of AI-infused systems and their impact on the user experience.

Accordingly, in the next section, we outline the characteristics that a new tool could have in general terms.

2.5 Defining traits of a new UX evaluation method

The answer to *[RQ3] What characteristics the new method could have?* requires deep reflection and a structured research activity. Nevertheless, we consider it beneficial to stimulate a fruitful discussion advancing broad defining characteristics of a novel UX evaluation method for AI-infused products.

Following Bargas-Avila and Hornbæk' suggestion (Bargas-Avila and Hornbæk, 2011), we define a new evaluation method by describing the methodology to be used and the dimensions of the user experience and the types of goods to be studied.

Given for granted that the objects of study are AI-infused systems, we may firstly summarize some hints regarding the methodological approach.

The discussion on the complexity of such systems already highlighted four main traits for the novel method that we may summarize as:

1. Flexible and holistic, to capture the nature of ever-evolving ecosystems.
2. Able to assess an experience over a long time and, eventually, track how the quality of the experience may evolve.
3. Appliable at different stages of the design process, even at the early stages.
4. Capable of capturing the very core of AI-infused systems as *otherware* rather than users' extensions.

Defining a methodology means also taking a position in the long-lasting debate between the quantitative and qualitative approaches, that move from different epistemological standpoints, one considering UX as something quantifiable and the other considering simplistic the idea of measuring the experience (Bargas-Avila and Hornbæk, 2011; Pettersson *et al.*, 2018).

Vermeeren and colleagues highlight a fair distribution of approaches that use solely quantitative data, purely qualitative data, or both in their study (Vermeeren *et al.*, 2010). These percentages contrast with those obtained by Bargas-Avila and Hornbæk. They found a majority of pure qualitative approaches (50%) over quantitative ones (33%), as well as those combining the two (17%). The numbers are inverted in the research by Rivero and Conte, with around 58 percent of techniques collecting quantitative data, 14 percent collecting only qualitative data, and lastly, 28 percent collecting both forms of data (Rivero and Conte, 2017). Very varied results may be explained by the different methods used by the researchers to select the available papers for the study. These results do not provide an excellent basis to consider in the definition of the new method. Nonetheless, we can see the need for triangulation of different methods in the divergences of the studies mentioned above, as proposed by (Vermeeren *et al.*, 2010) and investigated by (Pettersson *et al.*, 2018). They highlight a growing trend toward using more than one method to increase the reliability of the results.

We believe that this is a potential direction to pursue to understand the complicated nature of the user experience generated by AI-enhanced home devices. Thus, it is unsurprising that one of the first research focused on Amazon Echo (Sciuto *et al.*, 2018) used a mixed-methods approach that included logs and qualitative interviews.

The second element to be considered in defining a new evaluation method regards the UX dimensions they assess. As stated in the previous section, the current UX evaluation methods mainly focus on traditional UX dimensions that appear unable to fully frame the experience enabled by AI-infused systems.

They do not assess the meaning of engaging with other intelligences (Hassenzahl *et al.*, 2020) embedded in AI-infused systems. Understanding meaning as a definable – and even quantifiable – element in the user experience could play a prominent role in products whose potential is not fully exploited (White, 2018). Indeed, the meaningfulness of a user experience requires reflection on both the physical environment, which is frequently overlooked in existing paradigms and the social context. Additionally, it entails a close relationship with the users' history and motivations and their ethical viewpoints.

At the same time, it becomes mandatory to reflect on the perception of the other intelligence itself and the quality of the relationship it may establish with the user. The study on the UX of conversational interfaces directly consequences the previous statement since voice interaction is becoming a common trait of AI-infused systems (Kocaballi, Laranjo and Coiera, 2018).

Furthermore, the ethical implications of systems that collect and manage a massive amount of data – even personal ones – to work correctly should be somehow considered. Instead of or in addition to reliability, assessing trustworthiness could be relevant to such systems.

2.6 Conclusions

AI-infused products are game-changing, and the UX research appears unprepared to cope with novel paradigms of interaction, entailing a very different way of understanding interactive systems and their agency.

The research discussed so far confirms the assumption that current UX evaluation methods are not entirely adequate to holistically assess the user experience enabled by AI-infused systems. Furthermore, the UX dimensions commonly assessed provide a general understanding of the user experience, but do not analyze the very core of such systems.

Accordingly, a novel UX evaluation method specifically addressed to AI-infused system is needed to provide guidance in the development and improvement of such system. In the last section we highlighted broad traits of a method to be, with the aim of stimulating discussion within the design discipline.

References

Almeida, R.L.A. *et al.* (2018). "Towards Developing a Practical Tool to Assist UX Evaluation in the IoT Scenario". *Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web* (WebMedia), pp. 91-95. doi:10.5753/webmedia.2018.4575.

Antonelli, P. (2018). *AI Is Design's Latest Material*. Available at: https://design.google/library/ai-designs-latest-material/ (Accessed: 6 February 2019).

Bargas-Avila, J.A. and Hornbæk, K. (2011). "Old Wine in New Bottles or Novel Challenges: A Critical Analysis of Empirical Studies of User Experience". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI '11), pp. 2689-2698. doi:10.1145/1978942.1979336.

Brooke, J. (1995). "SUS: A quick and dirty usability scale". *Usability Eval. Ind.*, 189.

Burmester, M., Hassenzahl, M. and Koller, F. (2002). "Usability ist nicht alles – Wege zu attraktiven Produkten (Beyond Usability – Appeal of interactive Products)". *i-com*, 1(1), p. 32. doi:10.1524/icom.2002.1.1.032.

Canalys Newsroom – *Rise of Alexa, Google Assistant and Siri to drive US smart assistant-compatible device base to 1.6 billion in 2022* (no date). Available at: www.canalys.com/newsroom/rise-of-alexa-google-assistant-and-siri-to-drive-us-smart-assistant-compatible-device-base-to-16-billion-in-2022 (Accessed: 14 January 2022).

Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. New Ed edition. Cambridge, MA, USA: The MIT Press.

Dove, G. *et al.* (2017). "UX Design Innovation: Challenges for Working with Machine Learning As a Design Material". In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM (CHI '17), pp. 278-288. doi:10.1145/3025453.3025739.

Hassenzahl, M. *et al.* (2020). "Otherware: how to best interact with autonomous systems". *Interactions*, 28(1), pp. 54-57. doi:10.1145/3436942.

Hassenzahl, M., Burmester, M. and Koller, F. (2003). "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität". In Szwillus, G. and Ziegler, J. (eds.). *Mensch & Computer 2003: Interaktion in Bewegung*. Wiesbaden, Germany: Vieweg+Teubner Verlag (Berichte des German Chapter of the ACM), pp. 187-196. doi:10.1007/978-3-322-80058-9_19.

Hassenzahl, M., Burmester, M. and Koller, F. (2021). "User Experience Is All There Is: Twenty Years of Designing Positive Experiences and Meaningful Technology". *i-com*, 20, pp. 197-213. doi:10.1515/icom-2021-0034.

Hassenzahl, M. and Tractinsky, N. (2006). "User experience – a research agenda". *Behaviour & Information Technology*, 25(2), pp. 91-97. doi:10.1080/01449290500330331.

Holmquist, L.E. (2017). "Intelligence on tap: artificial intelligence as a new design material". *Interactions*, 24(4), pp. 28-33. doi:10.1145/3085571.

Jiménez, C. *et al.* (2012). "Evaluating a Methodology to Establish Usability Heuristics". In *2012 31st International Conference of the Chilean Computer Science Society. 2012 31st International Conference of the Chilean Computer Science Society*, pp. 51-59. doi:10.1109/SCCC.2012.14.

Jordan, P.W. (2000). *Designing Pleasurable Products: An Introduction to the New Human Factors*. CRC Press. doi:10.1201/9780203305683.

Kaptelinin, V. and Nardi, B.A. (2009). *Acting with Technology: Activity Theory and Interaction Design*. Cambridge, MA, USA: The MIT Press.

Karapanos, E. *et al.* (2009). "User experience over time: an initial framework". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Boston, MA, USA: Association for Computing Machinery (CHI '09), pp. 729-738. doi:10.1145/1518701.1518814.

Kocaballi, A.B., Laranjo, L. and Coiera, E. (2018). *Measuring User Experience in Conversational Interfaces: A Comparison of Six Questionnaires*. doi:10.14236/ewic/HCI2018.21.

Kujala, S. *et al.* (2011). "UX Curve: A method for evaluating long-term user experience". *Interacting with Computers*, 23(5), pp. 473-483. doi:10.1016/j.intcom.2011.06.005.

Lachner, F. *et al.* (2016). “Quantified UX: Towards a Common Organizational Understanding of User Experience”. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. New York, NY, USA: Association for Computing Machinery (NordiCHI ’16), pp. 1-10. doi:10.1145/2971485.2971501.

Levinson, P. (1977). “Toy, Mirror, and Art: The Metamorphosis of Technological Culture”. *ETC: A Review of General Semantics*, 34(2), pp. 151-167.

Lugmayr, A. and Bender, S. (2016). “Free UX Testing Tool: The LudoVico UX Machine for Physiological Sensor Data Recording, Analysis, and Visualization for User Experience Design Experiments”. In *Proceedings of the SEACHI 2016 on Smart Cities for Better Living with HCI and UX*. San Jose, CA, USA: Association for Computing Machinery (SEACHI 2016), pp. 36-41. doi:10.1145/2898365.2899801.

Maguire, M. (2019). “Development of a Heuristic Evaluation Tool for Voice User Interfaces”. In Marcus, A. and Wang, W. (eds.). *Design, User Experience, and Usability. Practice and Case Studies*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 212-225. doi:10.1007/978-3-030-23535-2_16.

Minge, M. *et al.* (2017). “The meCUE Questionnaire: A Modular Tool for Measuring User Experience”. In Soares, M., Falcão, C., and Ahram, T.Z. (eds.). *Advances in Ergonomics Modeling, Usability & Special Populations*. Cham: Springer International Publishing (Advances in Intelligent Systems and Computing), pp. 115-128. doi:10.1007/978-3-319-41685-4_11.

Nielsen, J. and Molich, R. (1990). “Heuristic evaluation of user interfaces”. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Seattle, Washington, USA: Association for Computing Machinery (CHI ’90), pp. 249-256. doi:10.1145/97243.97281.

Norman, D.A. (2004). *Emotional Design: Why We Love (or Hate) Everyday Things*. New York, NY, USA: Basic Books.

Obrist, M., Roto, V. and Väänänen-Vainio-Mattila, K. (2009). “User experience evaluation: do you know which method to use?”. In *CHI ’09 Extended Abstracts on Human Factors in Computing Systems*. Boston, MA, USA: Association for Computing Machinery (CHI EA ’09), pp. 2763-2766. doi:10.1145/1520340.1520401.

Otey, D.Q. (2017). “A methodology to develop usability / user experience heuristics”. In *Proceedings of the XVIII International Conference on Human Computer Interaction – Interacción ’17. the XVIII International Conference*, Cancun, Mexico: ACM Press, pp. 1-2. doi:10.1145/3123818.3133832.

Pettersson, I. *et al.* (2018). "A Bermuda Triangle? A Review of Method Application and Triangulation in User Experience Evaluation". In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal QC, Canada: Association for Computing Machinery (CHI '18), pp. 1-16. doi:10.1145/3173574.3174035.

Pyae, A. and Joelsson, T.N. (2018). "Investigating the usability and user experiences of voice user interface: a case of Google home smart speaker". In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. Barcelona, Spain: Association for Computing Machinery (MobileHCI '18), pp. 127-131. doi:10.1145/3236112.3236130.

Rivero, L. and Conte, T. (2017). "A Systematic Mapping Study on Research Contributions on UX Evaluation Technologies". In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems*. Joinville, Brazil: Association for Computing Machinery (IHC 2017), pp. 1-10. doi:10.1145/3160504.3160512.

Schmettow, M., Noordzij, M.L. and Mundt, M. (2013). "An implicit test of UX: individuals differ in what they associate with computers". In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. Paris, France: Association for Computing Machinery (CHI EA '13), pp. 2039-2048. doi:10.1145/2468356.2468722.

Schütte, S. *et al.* (2006). "Developing Software Tools for Kansei Engineering Processes: Kansei Engineering Software (KESo) and a Design Support System Based on Genetic Algorithm". In *9th International Quality Management for Organizational Development (QMOD) Conference*, August 9-11, Liverpool, England. 9th International Quality Management for Organizational Development (QMOD) Conference, August 9-11, Liverpool, England. Available at: http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-15720 (Accessed: 7 February 2020).

Sciuto, A. *et al.* (2018). "'Hey Alexa, What's Up?': A Mixed-Methods Studies of In-Home Conversational Agent Usage". In *Proceedings of the 2018 Designing Interactive Systems Conference*. New York, NY, USA: ACM (DIS '18), pp. 857-868. doi:10.1145/3196709.3196772.

Sivaji, A., Nielsen, S.F. and Clemmensen, T. (2017), "A Textual Feedback Tool for Empowering Participants in Usability and UX Evaluations". *International Journal of Human-Computer Interaction*, 33(5), pp. 357-370. doi:10.1080/10447318.2016.1243928.

Spallazzo, D., Sciannamè, M. and Ceconello, M. (2019). 'The Domestic Shape of AI: A Reflection on Virtual Assistants". In *DeSForM 2019 Beyond Intelligence*, pp. 52-59. Available at: https://re.public.polimi.it/handle/11311/1121253#.XiheQBNKh0s (Accessed: 22 January 2020).

Tractinsky, N., Katz, A.S. and Ikar, D. (2000). "What is beautiful is usable". *Interacting with Computers*, 13(2), pp. 127-145. doi:10.1016/S0953-5438(00)00031-X.

Vermeeren, A.P.O.S. *et al*. (2010). "User experience evaluation methods: current state and development needs". In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*. Reykjavik, Iceland: Association for Computing Machinery (NordiCHI '10), pp. 521-530. doi:10.1145/1868914.1868973.

Vitali, I. and Arquilla, V. (2019). "Conversational smart products: a research opportunity, first investigation and definition". In DeSForM. *Beyond Intelligence*, pp. 60-68. Available at: https://re.public.polimi.it/handle/11311/1115471#.XhdPpBdKh0s (Accessed: 9 January 2020).

White, R.W. (2018). 'Skill Discovery in Virtual Assistants". *Commun. ACM*, 61(11), pp. 106-113. doi:10.1145/3185336.

Zhou, Z. *et al*. (2019). "ML-Process Canvas: A Design Tool to Support the UX Design of Machine Learning-Empowered Products". In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow, Scotland Uk: Association for Computing Machinery (CHI EA '19), pp. 1-6. doi:10.1145/3290607.3312859.

3. The qualities of AI-infused products. Reflections on emerging UX dimensions

Martina Sciannamè
Politecnico di Milano, Department of Design

Emma Zavarrone
Iulm University, Humanities Department

The flourishing moment of AI that yields the first waves of materialization of this technology offers a fertile ground for UX and interaction designers, who might bring added values to the currently technology-driven AI-infused products.

However, these artefacts present inherent complexity and distinctive features that significantly affect the user experience (UX) and should be considered for their development and evaluation. The dimensions commonly considered in current UX assessments, though, result insufficient and inconsistent for this task. For this reason, Meet-AI, a research project funded by the Design Department of Politecnico di Milano, focuses on identifying the most fitting qualities to describe AI-infused products and ultimately aims to create a specific UX evaluation method.

Building on the premises described in chapter 2, this contribution portrays the process that led to the delineation of seventeen qualities at the basis of such method. Carried out within the Meet-AI project, the preliminary research is divided into two phases. The first investigates more and less traditional UX dimensions coming respectively from a wide-range critical analysis of existing UX evaluation methods and a literature review on AI and humans. The second aims at assessing the feasibility of the assumptions from the first phase and explores new qualities through a survey submitted to advanced users, which responses have been analyzed in subsequent steps culminating in a workshop within the research group. This eventually produced the synthesis from which starting to build the UX evaluation method, final objective of the Meet-AI project.

3.1 The UX of AI-infused products: a challenge to take on

The evolution and democratization of personal computers and the Internet demonstrates that seemingly niche technologies can successfully spread among the lay public if less technical and more *designerly* factors are considered in their development. Hence, the current materialization of AI in everyday products and services is a great opportunity for UX and interaction design.

Between the hype for novelty and the disillusion caused by unfulfilled promises and incomprehension, AI-infused products offer several possibilities for experimentations that appeal to basic design principles like discoverability, proper exploitation of functions (Hekkert, 2006; Kinsella, 2018; White, 2018), and personal significance, to finally let this technology bring richness and enjoyment to people (Norman, 2004). The relevance of the limitations brought by a technology-driven concretization of AI is recognized and countered both in academia (Dove *et al.*, 2017; Yang, 2020) and by the biggest companies providing such products and services, who are themselves defining and publishing guidelines to support the design of AI-infused systems (Amershi *et al.*, 2019; Google PAIR, 2019).

However, as portrayed in chapter 2, the peculiar nature of these products requires additional, ad hoc preparation of interaction and UX designers, who primarily need to comprehend the key features they have to work with.

Framing the UX of this new generation of products, then, is a starting point for a deeper understanding of the limitations and the potentialities they entail. Moreover, in order to conceive, develop, and improve products and services integrating AI, some guidance is needed, and these are the main premises to the Meet-AI project: a one-year-long research project, funded by the Department of Design, Politecnico di Milano. Its objective is to build a new evaluation method that specifically addresses the UX of AI-infused products, comprehending and highlighting the peculiarities they have in relation to other interactive products of common use. Based on the principal hypothesis that current UX assessment methods cannot frame and analyze the UX enabled by such systems, appropriate UX dimensions must be detected.

The chapter discusses the project's preliminary research and findings, obtained with a multi-method approach framed in two main phases that share the common goal of identifying the most relevant qualities to describe the UX of AI-infused products, to build the premises for a UX evaluation method. The first phase includes extensive research on existing UX evaluation methods and a literature review attempting to define AI as a UX matter. The second aims to verify the assumptions from the previous inquiries and further expand the investigation, including advanced users through a survey analyzed according to a multi-step protocol.

3.2 Phase one: setting the ground to understand and assess the UX of AI-infused products

3.2.1 Research methods

As a first step, the research required acknowledging the state of the art in UX evaluation to get a comprehensive picture of the most considered qualities for describing interactions between people and products of various kinds and to understand which ones might characterize the unique relationship with AI-infused products. Hence, this preliminary phase of the investigation was twofold: it comprehended a wide-range critical analysis of qualitative and quantitative UX evaluation methods, and a literature review on the intersection of AI, interaction design, and HCI to explore possibly uncovered angles.

Firstly, the five researchers involved in the study independently identified and examined relevant scales and methods to assess the UX, both within the field of design and in related social sciences experimentations. The research was limited to articles published in the ACM Digital Library and Springer Link between 2000 and 2020, resulting from the entry of "UX evaluation" and "UX assessment" keywords. To spot and integrate potentially missing methods, the All About UX website – the largest repository of UX evaluation methods available at the time of the study – has also been used as a reference. In the end, a list of 129 UX evaluation methods emerged, and they have been analyzed according to various criteria (Spallazzo, Sciannamè, *et al.*, 2021). Central for the inquiry was to highlight the *UX dimensions* and *descriptors* addressed in each case, here respectively intended as the general qualities that significantly describe people's experience of products and the specific features explaining the nuances of such overarching qualities. Additionally, to understand how the existing evaluation methods are operationalized and can inspire the construction of a new one (the primary expected outcome of the Meet-AI project) other relevant pieces of information have been pinpointed, namely: the *collection method(s)* (tools and modalities used to retrieve UX evaluations); whether more than one method has been put in place (*triangulation*); the context (*lab/field*), and the *support materials* used; the nature of the investigation (*qualitative/quantitative/both*); the product's *development phase* (concept, early prototype, functional prototype, or market level), and associated *period of experience* (before use, after an episodic interaction, an accomplished task or long-term utilization) in relation to which the evaluation can be carried out; the kind of *object(s) of study*; the *evaluators* required (single user, groups, expert users), and the researchers' perceived *level of consistency* with AI-infused products. To complement this, also sources and personal notes were added.

Because one of the premises of the Meet-AI project is the alleged absence of UX evaluation methods able to capture the essence of artifacts integrating AI systems, a deepening on the theme was also necessary. To this end, the researchers collected insights and reflections from a literature review revolving around the relationship between people and AI. The facets of human-AI interaction have been initially investigated according to three main thematic strands: non-human intelligence, emotion, and meaning. These, later on, evolved to include other relevant aspects in the current debate on the topic, namely, conversational interaction and ethical implications.

At the time of the study, UX and interaction design interest in AI was in its infancy, therefore, not many references could be found in the disciplinary literature. This is why the research expanded into related fields, such as HCI, computer ethics and AI itself.

3.2.2 What can be gleaned from current UX evaluation methods?

The critical analysis of existing UX evaluation methods eventually resulted in an extensive table (Spallazzo, Sciannamè, *et al*., 2021), from which some inferences can be derived.

The most easily quantifiable considerations concern the framing of the assessment methods (Fig. 3.1). As expected, the most frequently

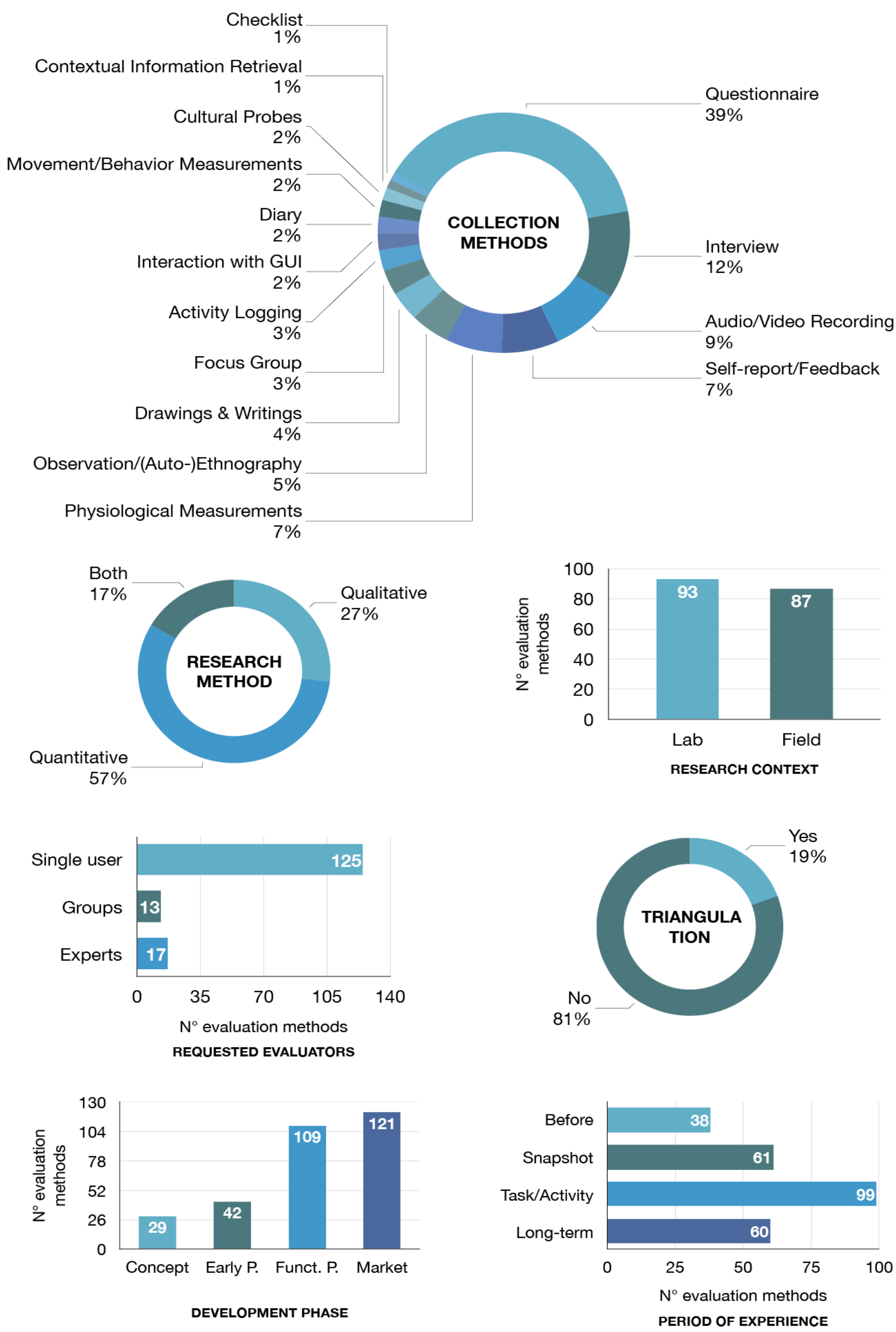


Fig. 3.1 – Synthetic overview of the UX evaluation methods analysis.

employed format for gathering evaluations is the questionnaire (69 methods), which has been interpreted both in traditional (questions and scales) and in more creative ways (with graphic, pictorial, auditory versions, and even randomly appearing when the phone screen is unlocked). Other prominent collection methods are interviews (21), video/audio recording (16), physiological measurements, and self-reports/feedbacks (13). These mostly reflect scientific approaches to evaluation, while modalities rooted in the design and social sciences fields, like diaries (4) and cultural probes (3), are less frequent.

As one can notice, the sum of collection methods exceeds the total of UX evaluation methods analyzed. 19% of them triangulate information retrieved in different ways to add soundness to more experimental or qualitative studies. The latter still represents a minority (27.6%) compared to quantitative practices (57%), although 16.4% opted for mixed methods.

Of course, digital devices (computers and mobile phones) stand out among support materials as they can easily process data coming from questionnaires, sensors, activity logs, and video/audio recordings, sometimes using custom software or apps.

The analyzed UX evaluation methods are equally submitted in a lab (93) and/or in actual contexts of use (87 occurrences) to test a wide variety of products and services. The majority is versatile and can encompass as many industrial products, as systems, environments, and events. Few specifically address reduced niches of interactive content such as visual interfaces or video games.

Additionally, this kind of investigation mostly require single non-expert users to evaluate the study objects when they are at an advanced design level – 109 methods can be applied to functional prototypes, and 121 when the product is already on the market (the same evaluation method can be submitted at different stages of development) – and after exploiting a precise task or activity (99). Debatably, the initial phases of the design process do not seem to be given much consideration for an early evaluation that may lead to a quick and rapid iteration of the product.

However, the wide-range analysis focuses on the qualities defining the UX dimensions and descriptors to evaluate products and services. Before getting into the issue, it is essential to state that literature reveals no agreement on terminology, and sometimes the exact words are used without careful attention to the semantic nuances or interchangeably as dimensions and descriptors.

If chapter 2 shows the recent dominance of generic UX in UX evaluation methods, our content analysis (Fig. 3.2) traces a slightly different story.

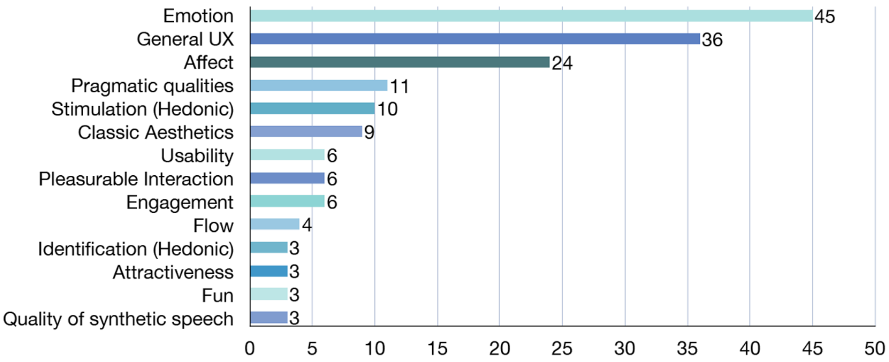


Fig. 3.2 – Prevailing dimensions in the mapping of UX evaluation methods.

The first relevant discrepancy concerning the previous studies is the prominent role that *emotion* and *affect* have come to play as UX evaluation dimensions: respectively, they appeared in 45 and 24 cases. *General UX*, instead, has been demoted to second place (with 36 occurrences). This is probably the result of the consistent number of experiments in the UX evaluation from the field of psychology.

Subsequently, *pragmatic* (17 occurrences if summed to usability), *hedonic* (10 stimulation + 3 identification) and *aesthetic qualities* (9) emerge, confirming the balance between practical and subjective sides of the overarching experience. More could be added to the list, that counts a total of 57 declared dimensions, like: *pleasurable interaction* (6), *engagement* (6), *flow* (4), *attractiveness* (3), *fun* (3), and so on. However, they mostly refer to soft characteristics that could be easily included in the previous, overarching, ones. Noteworthy is instead the *quality of synthetic speech* (3) – a new entry in the UX panorama – which is evidence that some methods are emerging to particularly address interfaces implying novel kinds of interaction modalities and they needed to introduce more specific attributes to assess the AI-infused products integrating them.

In the end, from the comprehensive work, the researchers highlighted the four most frequently assessed dimensions as the most significant for a holistic method, synthesizing the legacy of UX evaluation. They are the *pragmatic, hedonic, aesthetic,* and *affective* dimensions. As it refers to an overall perception of the user experience and lacks a clear measurability, *generic UX* has not been considered a valuable option for this study.

Afterwards, the descriptors collected from the UX evaluation methods have been organized according to these overarching dimensions to have the most recurrent ones depicting a nuanced picture of their dominant interpretations. It follows that the pragmatic dimension concerns the use of an artifact not just in terms of *helpfulness, efficiency,* and *functionality*, but also from more user-friendly characteristics such as *easiness, simplicity, clearness, navigation, learnability, reliability, and convenience.* Aesthetics presents multiple facets, probably due to its variability according to the object of the evaluation and subjective nature, but it can mainly be interpreted as *appearance* (where *clarity* and *sophistication* are two recurring themes), or *attractiveness* (*good* and *pleasant* qualities being frequently assessed). The hedonic dimension often overlaps with the affective one, as their borders are very subtle. *Enjoyability* and *excitement* are undoubtedly two factors that determine the pleasure of use, but also *creativity*, *inventiveness,* and *innovativity* appear to have their weight. Finally, the affective component is largely influenced by psychology. In fact, *valence* and *arousal* are by far the most common descriptors in this category, followed by the recognized basic emotions: *pleasure, fear, sadness, happiness, disgust, anger*, and *surprise.*

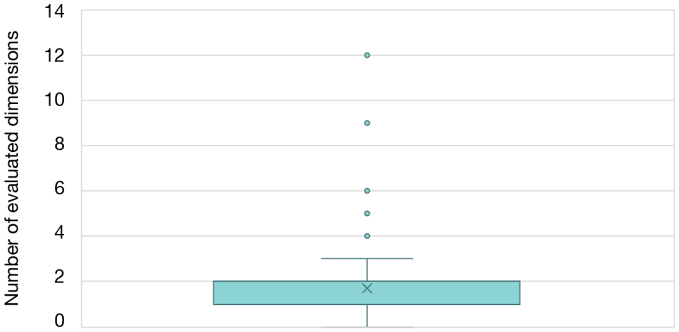


Fig. 3.3 – Variability of the UX dimensions assessed by each evaluation method.

The number of UX dimensions covered by each method represents another relevant information. The majority of the analyzed UX evaluation methods, in fact, moves within the boxplot (with a mean of 1.7 dimensions per method), underlining their limitation in dealing with complex, multi-faceted products, the interaction they imply is still to be fully comprehended. Just few outliers (showing high variability) demonstrated a more holistic perspective towards UX by exploring several dimensions, but they seemed more suitable references for the task at hand. Among the most distant from the mean are two methods notably referring to conversational interfaces: SASSI – Subjective Assessment of Speech System Interfaces (Pettersson *et al.*, 2018), 12 dimensions; SUISQ – Speech User Interface Service Quality (Polkosky and Lewis, 2003), 8; the UEQ (Polkosky, 2005), 6; and AttrakDiff (Laugwitz, Held and Schrepp, 2008), 4.

Overall, Meet-AI researchers' evaluations of the consistency of the analyzed UX methods with the subject matter confirmed the baseline hypothesis: no current UX evaluation method is sufficiently comprehensive and not too broad to adequately address AI-infused products. The majority (45) received an average score of 3 (on a 1 to 5 scale), 31 reached a 4, but none has been considered totally suitable to manage AI-infused products assessment. However, two methods have been flagged as particularly interesting for the purpose of the Meet-AI project. The first is the Affective Feedback Loop (Bruns Alonso *et al.*, 2013) which, although not strictly an assessment tool, suggests to design interactive objects in a way that they can sense users' behaviors and receive affective feedbacks as part of the human-computer interaction. This could be particularly valuable for products integrating machine learning (ML) systems – the predominant subset of AI – as they can learn and adapt over time. Translated into an evaluation method, this approach could be implemented in a functional prototype or a market phase of a product to gather rich and unmediated information and to respond with prompt iterations. While the second, Perceptive Sorting (Forlizzi, Gemperle and DiSalvo, 2003), is outside the timespan of the overview but still a relevant reference. It concerns the evaluation of unfamiliar products for target users – as in the case of AI-infused ones – and distinguishes three levels of assessments to infer design directions for the development of new artifacts: using normative qualities to describe the usability of displayed familiar objects, lifestyle indications for known objects with many unfamiliar models, and affective words for unfamiliar objects which function was explained by the researchers. In the latter, a significant scenario for products using natural language interactions emerged: personality traits and gender have been often attributed to the presented robots, as to define the unknown with more human qualities.

To conclude with the results of this analysis, despites the appealing alternative methods for assessment collection, the questionnaire format is not only the most frequently used, but, probably the most straightforward to elaborate and introduce in the design process of new products integrating AI. Thus, this will be the goal of the Meet-AI project.

3.2.3 A deeper dive into AI-infused products: what literature tells us?

As the investigation on existing UX evaluation methods confirmed the initial assumption of the Meet-AI project, a wide-ranging literature review on AI-infused products and AI-related issues seemed essential to identify possible latent but significant UX dimensions before starting the development of a specific evaluation method. Here, we will outline the main findings.

As anticipated, the inquiry started from three main strands: non-human intelligence, to clarify what artificial intelligence is and what are its dominant features; emotions, as they play a prominent role in the UX evaluation, they may help describe human-machine relationship; and meaning, to understand how people make sense or can interpret AI mediated interactions.

The basis of this argumentation retraces AI history to comprehend how this discipline has been conceived and evolved over time. Its very name already suggests a dichotomy in its nature, combining machines and a distinctly human trait: intelligence. In their world-wide recognized textbook on AI Russell and Norvig (2020) try to give an explanation of this concept as foundational feature of this subfield of computer science. They distinguish four connotations that have been attributed to artificial intelligence: intelligence as accurate simulation of *human* performance, as *rationality*, as an internal *thought* process or as an externalized *behavior*. The combination of such definitions yields four possible representations of AI systems. They can be considered intelligent if they (i) *act humanly* – i.e., they successfully communicate in a human language (natural language processing: NLP), store information they know or hear (knowledge representation), answer questions to draw new conclusions (automated reasoning), or adapt to new circumstances and identify patterns (machine learning: ML); if they (ii) *think humanly* – manifesting human-like thought processes as observed in psychology, cognitive and neural sciences; if they (iii) *think rationally* – applying irrefutable reasoning processes derived by logic or probability; or if they (iv) *act rationally* – meaning that AI systems, and predominantly ML systems, are rational agents that can operate autonomously (with no step-by-step program), perceive their environment and respond within it, adapt to change, improve over time by learning from past experiences, and pursue goals to achieve the best expected outcome. The latter is the prevailing approach in the field; hence, the listed qualities represent a major reference for the study.

Similar features of AI systems emerge from HCI, and ambient intelligence and they are regarded as the disruptive elements that force a change in the way we look at the UX they entail. Products with autonomy, adaptability, reactivity, multifunctionality, ability to cooperate, human-like interaction, personality (Rijsdijk and Hultink, 2009), and that can also be personalized on habits or preferences, context-aware and anticipatory (Aarts and Ruyter, 2009) are no regular products, and their UX cannot be defined by simple usability, utility and interaction aesthetics (Dove *et al.*, 2017). Which is why their *intelligence* (implying all the above) needs to be considered for a specific evaluation method, as it is corroborated by the work of Amershi *et al.* (2019) who have proposed 18 design guidelines for human-AI interaction, justifying them based on the unpredictability of behaviors that AI-infused products have.

This last aspect triggers several unprecedented implications, subject of a lively debate, especially among computer ethicists. The researchers investigated an additional domain and identified a second, essential UX dimension: *trustworthiness*.

As a matter of fact, for an effective user experience of autonomous and continuously evolving systems, certain issues are of utmost importance to gain users' trust. For instance, the values and objectives put into the machines should be aligned with people's ones (Russell and Norvig, 2020) to guarantee a beneficial impact; humans' role in the development and in the interaction should be well communicated as they are part of the AI systems themselves (Johnson and Verdicchio, 2017); and, above all, people should be able to understand why these systems make their decisions, how they function, what are their capabilities and limitations. Explainability is indeed a crucial aspect from different perspectives: for ethicists (Kulesz, 2018), designers (Yang, 2020) and also computer scientists, in whose discipline the specific branches of explainable AI – XAI (Confalonieri *et al.*, 2021), and interpretable ML (Molnar, 2019) emerged.

A great number of directions for developing beneficial AI systems emerged in a short span of time. For reference Algorithmic Watch (2020) represents the most up to date repository, while Hagendorff (2020) discusses them in his paper. However, the most comprehensive ones are those published by the European Commission (High-Level Expert Group on Artificial Intelligence, 2019). They include seven main guidelines for trustworthy AI: (i) human agency and oversight; (ii) technical robustness and safety; (iii) privacy and data governance; (iv) transparency; (v) diversity, non-discrimination and fairness; (vi) societal and environmental well-being; and (vii) accountability.

Looking at ethical issues also from an affective standpoint, they can be relevant for UX assessment as the presence or lack of concerns may affect people's responses, for example a lack of explainability can provoke uncertainty, frustration, doubt, mistrust (Fruchter and Liccardi, 2018).

While the thread of emotion studies dominates the scenario of UX evaluation methods as they fill the gap between people and products (Forlizzi and Battarbee, 2004) influencing their attitudes, behaviors, perceptions and assessments (Scherer, 2005), the researchers also delved into an intertwined matter: *meaning*. As Norman (2004) stated, both affect and cognition are related to an evaluation process: the first to determine the positive or negative impact that things surrounding us may have; the second to make sense of the world. Indeed, attributing a meaning to a product and understanding what it represents for the user is a relevant measure for the UX, even if outlining its defining traits can be tricky. Within the HCI, ethics and design communities, this has been variously debated, both as a cognitive issue (High-Level Expert

Group on Artificial Intelligence, 2019) and as a quality of the human-product interaction that satisfies psychological needs (Dourish, 2001; Hassenzahl *et al.*, 2013; Mekler and Hornbæk, 2019) like autonomy, competence, relatedness, popularity, stimulation, security (Hassenzahl, 2011). Not to perpetrate the path of devices showing off their technological novelties but confined to the gadget or toy dimension (Levinson, 1977), the design process should be steered by meaningfulness concepts. AI-infused products, as all human artifacts, might be developed with a predefined purpose, resonate as a personal significance, a shared and/or cultural significance, generate valuable experiences, communicate a symbol or exhibit a temporal quality, thus referring to some kind of meaning at a functional, ritual and/or mythical level (Ajovalasit and Giacomin, 2019).

The research finally closed the circle on possibly relevant UX dimensions for products integrating AI systems by investigating a very peculiar interaction modality that these activate as a result of the first conceptions of intelligence embodiment: *conversational* capabilities. Even though AI is not synonymous with conversational interface, voice assistants are among the most widespread manifestation of this technology, and they fail to be evaluated by traditional UX methods. In fact, as already brought out, specific methods have risen, and additionally, reflections and experiments emerged in HCI literature.

A good overview of the state of user interaction with speech systems has been given by (Clark *et al.*, 2019) who highlighted that commonly measured concepts include: user attitudes (towards the interface), task performance (total of dialogue turns, task completion, etc.), lexis and syntax choice, perceived usability, user recall (of specific aspects and outputs), and physiological qualities (like speech loudness and pitch). Moreover, heuristics (Maguire, 2019) and other experimentations (Bartneck *et al.*, 2009; Garcia, Lopez and Donis, 2018) outlined further essential aspects to keep in consideration when dealing with human-like interactions, such as accommodating conversational speech, ensuring high accuracy to minimize input errors and adequate system feedback (also in recovering from errors), but also more human features like anthropomorphism, animacy, likeability, and, above all, the personality of the agent.

3.3 Phase 2: Actively exploring the qualities of AI-infused products

3.3.1 Overview of the research methods

Not to limit the investigation to the subjective perspective of the researchers, they opted for introducing a novel element to the comparative review. In addition to the inferences from the literature review, a co-creation of the descriptors to compose the scale has been envisioned as a suitable way to engineer their selection and preserve as much objectivity as possible.

Hence, a protocol combining mixed methods and different steps was established. *Step 0:* the driving force behind the second phase of research is a survey intended to validate the assumptions concerning the dimensions to describe AI-infused products and to solicit a creative contribution on descriptors to expand the non-comprehensive set extracted from the literature review. While the survey provided immediate results on the dimensions (*step 1*), the analysis of the suggested descriptors was more articulated and composes *step 2*. A preliminary homologation of the responses proposing descriptors was independently conducted by two of the researchers and finally confronted to compile a shared list. The descriptors in the list were further analyzed through an affinity map to synthesize repetitions and filter out out-of-context responses. The consequent set of descriptors was then submitted to the researchers of the Meet-AI project for an inter-coder evaluation aimed at assessing the consistency of the descriptors with the related dimension and their relevance for AI-infused products, and finally at extracting the most significant ones. The last step (*step 3*) before the construction of the evaluation scale was an internal workshop within the Meet-AI team to define the elements (dimensions and descriptors) that would form its structure in light of the results of the whole research work.

Because of the articulated configuration of this second phase of the research, further details on the methods will be provided in the next paragraphs, in combination with the results of each step.

3.3.2 Step 0: a survey to expand the boundaries of AI-related qualities

To cross-examine the findings from the initial investigation (namely the eight identified dimensions: *pragmatic, aesthetic, hedonic, affective, intelligence, trustworthiness, conversational, meaningfulness*) and to further enrich the spectrum of attributes that might describe the target products, a group of advanced users has been involved through a digital survey. The selected population was composed of 110 students from MSc in Digital and Interaction Design and 47 young researchers from the Design Department of Politecnico di Milano, who are familiar with the type of products being studied and have a developed sensitivity and comprehension of the design of interactive objects. From the total, 42 responded, with a response rate of 26.75%.

The survey was meant to be clear and transparent on its purpose, therefore, this was openly stated in the introduction, and the AI-based smart speakers, learning thermostats and smart cams were presented both to give references of the artifacts to be addressed and to understand the respondents' level of experience with the exemplified objects. Instead, the core inquisitive part was twofold, focusing first on (i) seeking a new set of descriptors and then on (ii) acquiring feedbacks on the UX dimensions aimed at describing products integrating AI systems. (i) To avoid possible misunderstandings, the dimensions according to which the advanced users were asked to suggest new UX qualities were portrayed with a definition before getting to research request. As synthesized in Fig. 3.4, in some cases the respondents had to indicate at least three attributes, in others, a minimum of two positive and two negative features, to encourage heterogeneous answers, (ii) After the explorative contribution, the proposed dimensions have become the subject of critique for the researchers to understand how well they perform in the evaluation of AI-infused products, which are considered the most relevant, and whether may there be missing ones. To gather such information, direct questions have been posed. A profiling section closed the questionnaire.

DIMENSION	DESCRIPTION	QUESTION
Pragmatic dimension	Some qualities of products support users in achieving their concrete goals, such as performing specific tasks. They may include (but are not limited to) usability, intelligibility, efficacy issues.	Please write at least three attributes (adjectives, nouns, verbs) you consider peculiar and relevant to describe the quality of use of AI-infused products.
Aesthetic dimension	The aesthetic appearance of industrial products plays an essential role in our relationship with them. Despite being subjective, the appreciation of beauty may be affected by different aspects (e.g. shape, colour, material, finishing, behaviour, etc.).	Please write at least three attributes (adjectives, nouns, verbs) you consider the most relevant and unique to describe the aesthetic qualities of AI-infused products.
Hedonic dimension	Some qualities of products can make them attractive and engaging, and arise pleasant and satisfying sensations during use.	Thinking specifically of AI-infused products, please write at least three essential qualities (adjectives, nouns, verbs) that characterize them as pleasurable and attractive.
Affective dimension	While interacting with products, they often influence our emotional state by inducing subjective feelings. This can be particularly relevant with AI-infused products.	List a minimum of 2 positive and 2 negative affective responses you consider typically caused by AI-infused products.
Trustworthiness	A product can be defined as trustworthy when it is individually and socially acceptable and reliable, and it represents a well balanced trade-off between human principles and practical needs, benefits and risks.	Envisioning the possible positive and negative impacts of AI-infused products, write at least 2 essential features for them to be trustworthy and at least 2 unreliable.
Conversational dimension	Some AI-infused products like smart speakers (Amazon Echo, Google Home...) can use voice and text to interact with users. Voice can be used to do tasks, answer questions, control other products, and engage in conversation. A "conversational" product or system is able to use natural language in an interaction that lasts multiple turns of dialogue.	Reflecting on the most impactful features in the design and use of conversational systems, write at least 2 features (adjectives, nouns, verbs) that contribute to creating a positive and efficient interaction, and at least 2 features that may ruin the experience.
Intelligence	AI-infused products can autonomously learn to adapt their behaviour over time, and can proactively take action or propose suggestions to their users.	Write at least 2 relevant features (adjectives, nouns, verbs) an AI-infused product should have to be considered intelligent, and at least 2 features that lessen the perception of intelligence.
Meaningfulness	Some aspects of products can make them meaningful to their users in the sense that they may manifest a tangible purpose, a personal significance, a shared/cultural significance, generate past experience, communicate a symbol or exhibit a temporal quality.	Thinking specifically to AI-infused products, please write at least three attributes (adjectives, nouns, verbs) that make you perceive AI-infused products as meaningful.

Fig. 3.4 – Synthesis of the descriptors requests as they appeared in the survey.

3.3.3 Step 1: UX dimensions of AI-infused products according to advanced users

Moving backwards in the analysis of the survey responses, this paragraph outlines the assumptions related to the proposed dimensions, while more grained and qualitative considerations on the descriptors that the respondents attributed to each of them will be depicted in the following one.

As graphically portrayed in Figure 3.5, the overarching qualities listed in the survey received quite positive ratings. As might be expected, *trustworthiness*, *intelligence, conversational,* and *meaningfulness* – the dimensions stemming from the AI-focused literature review – found a strong consensus among advanced users (underlining their consistency with the target products), while the most frequently used in current methods – *pragmatic, aesthetic, hedonic* and *affective* dimensions – were mainly assessed as "important".

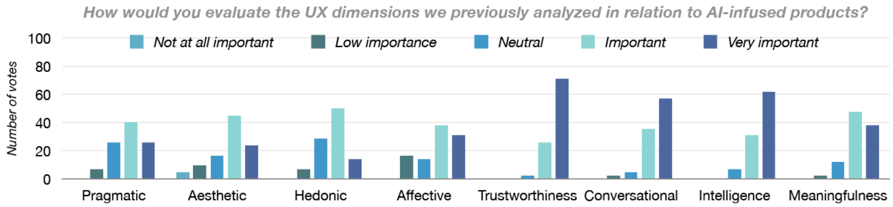


Fig. 3.5 – Survey results on the evaluation of the proposed UX dimensions for AI-infused products.

The next question, double-checking the relevance of the dimensions with respect to the UX of AI-infused products (Fig. 3.6), further confirmed these results. With the favor of 76% of the respondents, *trustworthiness* prevailed, and it was followed by *conversational* (59.5%), *intelligence* (50%), and *meaningfulness* (40.5%).

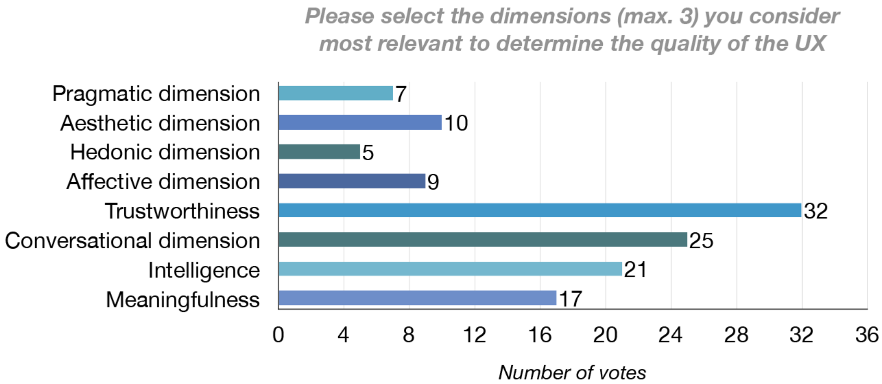


Fig. 3.6 – Survey results highlighting the most relevant UX dimensions for AI-infused products.

In contrast, the direct solicitation for additional dimensions essential to the project goal was not met with useful replies. In fact, the few comments received confirmed the proposed selection, contained features better identifiable as descriptors, or were off topic.

Other parameters may be looked at to find a confirmation of the results above and the researchers' initial hypothesis. The substance and the way in which people answered to the request for descriptors for each dimension is indeed significant.

For instance, the personal contribution, appropriateness, and coherence of the responses reveal the advanced users' perceptions towards the different dimensions. In some cases, they were straightforward (e.g., *pragmatic* dimension), in others, a proper assessment could be more difficult. *Hedonic*, *affective* and *meaningfulness* dimensions, in fact, received mostly incoherent and long-winded responses, showing respectively high subjectivity, shortcomings and both these issues merged, with some respondents openly stating their inability to answer at all. The inconsistency was also proved by valuable attributes provided within these categories, but that were more appropriate in other contexts.

The prevailing richness of the responses – in terms of amount (for *conversational* and *intelligence* dimensions), and articulation (*trustworthiness*) – their suitability with the overarching factors and the subjects of evaluation, and the pervasiveness of the related contents throughout the questionnaire made the appreciation of these qualities even more explicit.

Aesthetics, though, was the one dimension with poor-quality data. A low perceived relevance in relation to the research goal transpired (as confirmed in the explicit evaluation), with responses bearing some level of superficiality by addressing specific characteristics of the products on the market.

3.3.4 Step 2: insights from an intertwined analysis of AI-related descriptors

To properly derive all facets from the suggested descriptors and infer useful information for the construction of a new evaluation method for AI-infused products, the raw responses needed some preparation and further assessment by the Meet-AI research group.

Preliminarily, to make the survey responses consistent with the initial request, two of the researchers redacted a homogeneous list, translating sentences and Italian answers in single English words. The resulting lists of one-word descriptors were then confronted to compile a uniform one (Spallazzo and Sciannamè, 2021).

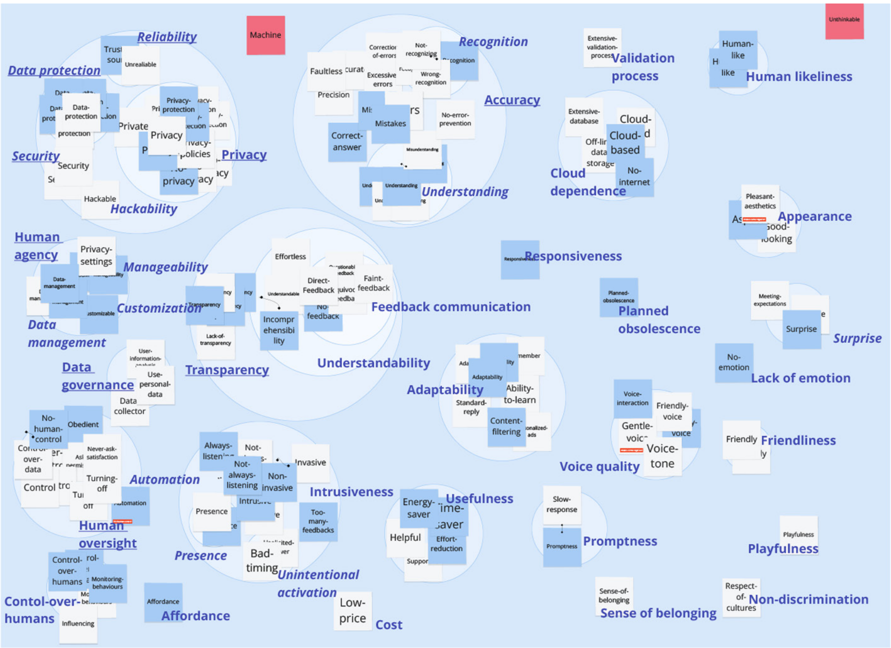


Fig. 3.7 – Affinity map of the descriptors from the Trustworthiness dimension.

Subsequently, all the entries were collected in a Miro board (Fig. 3.7), differentiated according to the study dimensions. For each, an affinity map was built to visualize semantic concentrations, extrapolate descriptors in a univocal way – blue words outside post-its (also referring to the language encountered in literature – underlined), and flag (in red) any incomprehensible or patently out-of-context concept (e.g., *lights* or *function* as descriptors of the *affective* dimension). Blue post-it notes highlight the terms modified by the two researchers to meet the one-word (English) format.

Ultimately, a synthetic list of unambiguous descriptors was then prepared and shared among the research group for crossed-evaluation, seeking intercoder agreement (Creswell, 2014) for each feature based on two parameters: (i) the consistency with the dimension, and (ii) the relevance for AI-infused products. Each descriptor was presented within the dimension for which it was originally proposed by the respondents, along with the frequency of its occurrences. To complete the picture, also descriptors from the literature review (L) were added in different sheets. The researchers, in the role of judges, had to assess all of them on a 1 (not consistent/relevant at all) to 4 (highly consistent/relevant) scale. The results are public at (Spallazzo, Ajovalasit, *et al.*, 2021), and they have been analyzed by calculating the mean and z score for each descriptor and according to each parameter (consistency and relevance). Then, they have been organized into quartiles to easily spot the most significant. Finally, a comprehensive overview (Fig. 3.8) was obtained by comparing the relevance z scores of the descriptors altogether. Also dividing these into quartiles, the segment >75% counts 134 descriptors. Hence, special consideration has been given to the 36 receiving the maximum score from all the researchers (the "golden" ones), and they have been operationalized in the last step (described in the next paragraph).

After this processing, an overview of which is offered in Fig. 3.9, some informed inferences could be derived. For the sake of synthesis, they are here discussed in relation to their overarching dimensions.

Pragmatic dimension. It was probably the easiest for the respondents, who answered consistently (one of the highest overall consistency scores from the researchers' evaluation), with one-word descriptors as requested, and collected a total of 46 descriptors with 134 submitted items (number of words originally suggested by the respondents and then synthesized based on their semantic affinity) – only two of which did not reach the list to be assessed. It did not present major innovations compared to the literature, yet some new aspects directly related to AI-infused objects emerged, e.g.: *smartness, customization, responsiveness, adaptability, connectivity, unobtrusiveness,* and different concepts linked to *trustworthiness*. In terms of relevance, it marked the second-best score in both the mean of evaluations and the overall "golden" descriptors, probably underlining that this basic dimension for evaluating UX is still essential or, at least, most of the relevant qualities of AI-infused products have been attributed to this dimension by the respondents.

Aesthetic dimension. Respondents' answers in this category presented a high influence of currently adopted practices in the industry of AI-infused products. Often, they were too specific in indicating characteristics of products on the market (e.g., white color, small size, rounded shapes, etc.) and a great work of generalization was necessary in order to submit the descriptors to the judges. In the end, these

SOURCE	DESCRIPTOR	R1 EV	R2 EV	R3 EV	R4 EV	R5 EV	R6 EV
CONV-L	Voice naturaleness	4	4	4	4	4	4
CONV-L	Voice pleasantness	4	4	4	4	4	4
CONV-Q	Accuracy	4	4	4	4	4	4
CONV-Q	Context awareness	4	4	4	4	4	4
CONV-Q	NLP quality	4	4	4	4	4	4
CONV-Q	Reliability	4	4	4	4	4	4
CONV-Q	Understanding	4	4	4	4	4	4
HED-Q	Empathy	4	4	4	4	4	4
INT-Q	Accuracy	4	4	4	4	4	4
INT-Q	empathy	4	4	4	4	4	4
INT-Q	Context awareness	4	4	4	4	4	4
INT-Q	Understanding	4	4	4	4	4	4
MEAN-Q	Usefulness	4	4	4	4	4	4
PRAG-L	Functionality	4	4	4	4	4	4
PRAG-L	Helpfulness	4	4	4	4	4	4
PRAG-L	Intelligibility	4	4	4	4	4	4
PRAG-L	Intuitivity	4	4	4	4	4	4
PRAG-L	Learnability	4	4	4	4		4

SOURCE	DESCRIPTOR	R1 EV	R2 EV	R3 EV	R4 EV	R5 EV	R6 EV
PRAG-L	Reliability	4	4	4	4	4	4
PRAG-L	Understandability	4	4	4	4	4	4
PRAG-Q	Customization	4	4	4	4	4	4
PRAG-Q	Ease of use	4	4	4	4	4	4
PRAG-Q	Transparency	4	4	4	4	4	4
PRAG-Q	Trustworthiness	4	4	4	4	4	4
TRUS-L	Access to data	4	4	4	4	4	4
TRUS-L	Human oversight	4	4	4	4	4	4
TRUS-L	Non-discrimination	4	4	4	4	4	4
TRUS-L	Privacy	4	4	4	4	4	4
TRUS-L	Quality of data	4	4	4	4	4	4
TRUS-L	Transparency	4	4	4	4	4	4
TRUS-L	Unfair bias avoidance	4	4	4	4	4	4
TRUS-Q	Accuracy	4	4	4	4	4	4
TRUS-Q	Data management	4	4	4	4	4	4
TRUS-Q	Data protection	4	4	4	4	4	4
TRUS-Q	Reliability	4	4	4	4	4	4
TRUS-Q	Transparency	4	4	4	4	4	4

Fig. 3.8 – List of the "golden" descriptors with the related dimensions.

DIMENSION	SUBMITTED ITEMS	RESULTING DESCRIPTORS	EXCLUDED ITEMS (S)	MEAN DESCRIPTORS CONSISTENCY	MEAN DESCRIPTORS RELEVANCE (S)	MEAN DESCRIPTORS RELEVANCE (L)	GOLDEN DESCRIPTORS
Pragmatic dimension	136 (S) + 49 (L)	46 (S) + 49 (L)	2	2.53	2.67	2.68	11
Aesthetic dimension	132 (S) + 43 (L)	37 (S) + 30 (L)	1	2.03	1.76	1.92	0
Hedonic dimension	133 (S) + 32 (L)	55 (S) + 30 (L)	2	2.00	2.05	2.31	1
Affective dimension	158 (S) + 219 (L)	49 (S) + 96 (L)	52	2.69	2.47	1.64	0
Trustworthiness	140 (S) + 41 (L)	39 (S) + 41 (L)	2	2.50	2.62	3.04	12
Conversational dimension	161 (S) + 28 (L)	60 (S) + 22 (L)	1	2.64	2.70	2.93	7
Intelligence	141 (S)	53 (S)	0	2.19	2.30	/	4
Meaningfulness	155 (S)	49 (S)	2	2.24	2.39	/	1

(S) from survey | (L) from literature

Fig. 3.9 – Descriptors performances, synthesized according to the related dimension, in the various steps of the analysis

received the lowest consistency, the lowest relevance mean of evaluation (even with a negative connotation: score of 1.76 out of 4), and they were the least represented among the overall most relevant with no one resulting as a "golden" descriptor. What stands out (with a relevance mean of 3.8), though, is a quality that diverge from the most traditional conception of aesthetic, merging with studies in the affective perception of products: *personality*. It is followed by the concept of *mimesis* (dear to the field of ubiquitous computing) which records a mean of 3 but occurs also in the nuances of *invisibility* and *unobtrusiveness*.

Hedonic dimension. Despites being at the second place for number of descriptors (55) from the questionnaire, its performances were among the lowest, with slightly sufficient thresholds both in the consistency and relevance evaluations. Here, qualities directly linked to AI-infused products emerged (e.g., *multifunctionality, responsiveness, voice interaction*) but most of them were considered not particularly significant or just more appropriate for other dimensions. This is also manifest in the overall ranking of relevant descriptors, where the hedonic dimension gives just a little contribution. Particularly noteworthy is instead the value of *empathy* which receives here the highest accreditation from the judges, immediately followed by *adaptability* (even if both occur in 6 over 8 dimensions).

Affective dimension. In stark contrast to the analyzed UX evaluation methods (counting 96 different descriptors and 219 occurrences from literature), the analysis depicted a lot of confusion among the respondents. The difficulty of correctly expressing one's emotions clearly emerges. Lots of articulated sentences (even if just single words were requested) appeared and most of the answers were aiming at the cause of emotions in the interaction and not at the affective responses themselves. For this reason, it recorded the highest, impressive number of exclusions for manifest inconsistency even before the judges' evaluation (around 1/3 of the descriptors coming from the responses were discarded). In the end, the affective descriptors from the questionnaire did not perform badly in the judges' opinions (also marking the best score for consistency), and actually relevant qualities for AI-infused products can be highlighted, like the empowering *feeling in control* and *feeling understood*, right before those emerging from direct interaction with such devices: *attraction, challenge, disappointment, frustration,* and *satisfaction*. However, the affective dimension was only second to the aesthetic one in terms of the least number of descriptors in the >75%, with no "golden" items as well, and the traditional qualities (coming from literature) proved not to be valuable when considering the UX of AI-infused products, with a negative mean of 1.64 (out of 4).

Trustworthiness. It marks its commonly agreed relevance in qualitative and quantitative ways. Firstly, even in this dimension, answers were quite articulated, but mostly they didn't highlight a misunderstanding or a difficulty in answering, but rather a desire for a better explanation. Secondly, qualities referring to this category emerged throughout all other dimensions, underlining their pervasive importance. As expected, this was also manifest in numbers: the amount of reported descriptors was rich, the judges' evaluations of trustworthiness descriptors were among the highest and this dimension is the one that contributes in the largest part to the top >75% overall relevance ranking with exactly 1/3 of the "golden" descriptors pertaining to it. The most successful ones quite echo the European guidelines and concern *accuracy, data management, data protection, reliability,* and *transparency.*

Conversational dimension. With 60 descriptors and 160 submitted items, it was the most prolific in absolute terms. It presented a good set of precise responses (maybe not from a literary point of view but with enough granularity), marking its perceived significance in relation to AI-infused products. The judges' evaluations also reflected this position in both consistency and relevance average values, as well as in the overall relevance ranking, where a lot of conversational descriptors appear in the >75% and "golden" ones count (only following *trustworthiness* and *pragmatic* dimensions). As stressed by the specificity of some of its descriptors (like *NLP quality, accent & dialect recognition, voice quality, character*, etc.), the conversational dimension is mostly relatable to a part of AI-infused products. Tough, others can also be generalized to a more comprehensive behavior. It is the case of *accuracy, context awareness, understanding, feedback quality*, but also *fluidity* and *naturalness.*

Intelligence. Even though it is undeniably difficult to define intelligence, highlighting the qualities that characterize a perceived intelligent behavior in AI-infused products proved to be a less heavy task. The responses in this dimension were satisfying: of the 141 submitted items, none was discarded in the first round of analysis (preceding the judges' evaluation). They also performed quite well, placing themselves in an average position in terms of evaluated consistency and relevance of the proposed descriptors, as well as in the overall relevance ranking. Here, again, characteristics like *accuracy*, *adaptability*, *context awareness,* and *understanding* stood out, in a mixed context that presents some traits reminding human capabilities (e.g., *learning, understanding needs, companionship*), and others strictly linked to the machine dimension (e.g. *data elaboration, connectivity*).

Meaningfulness. In conclusion, this was undoubtedly the toughest dimension to depict, and it is not by chance that this had the smallest number of items proposed (115). Nonetheless, respondents tried to answer according to the request – without long-winded digressions – but some of them openly expressed their inability to answer at all. Probably the difficulty to determine what belongs to this domain did not help to encounter some preferred qualities uniquely belonging to it. In fact, attributes with fuzzy boundaries recurred, like *trustworthiness, multi-purposeness, personality, empathy,* and *understanding*. Yet, those evaluated as most interesting (*usefulness*, *being beneficial*, and *helpfulness*) mostly appeal to the human-computer/product relationship.

3.3.5 Step 3: a summarizing workshop

Once all the preliminary research activities (from the literature review and the UX evaluation methods mapping to the survey submission and analysis) resulted in a synthetic portrait of the most relevant descriptors for the assessment of AI-infused products, a workshop within the research group seemed the most suitable way to collectively discuss the reached outcome and to pave the way to the construction of a specific evaluation method.

As anticipated, the so-called "golden" descriptors were extrapolated to understand their possible role within the UX assessment of products integrating AI systems. They were displayed on post-it notes on a Miro board with their related dimension. Then, the researchers categorized them according to their perceived likelihood of being part of the scale to be built.

Some (*data protection, quality of data, unfair bias avoidance, trustworthiness,* and *non-discrimination*) were labelled as "not usable" because of the lack of information and difficult measurability for a proper assessment. Instead, repeated descriptors pertaining to multiple and less coherent dimensions were considered "better to keep out", while weaker and too general attributes were left in the "could be in" category to leave space to the "must be in" ones. The ultimately selected descriptors are depicted in Fig. 3.10. While *empathy, understanding*, and *usefulness* refer to human-related qualities; *helpfulness, intuitiveness, reliability, accuracy, adaptability,* and *context awareness* are attributes properly belonging to the system itself. Lastly, *customization, human oversight, data management, privacy, transparency,* and *reliability* (as ethical concern) configure the product as a sociotechnical ensemble, merging human needs and system properties.

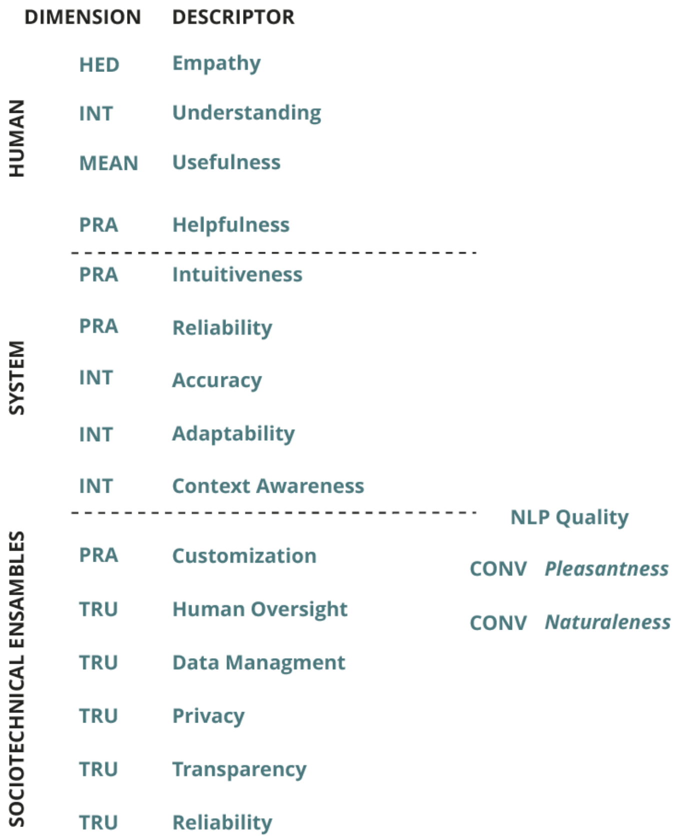


Fig. 3.10 – Ultimately selected descriptors to build a UX evaluation method for AI-infused products.

3.4 Conclusions, limitations, and future directions

The chapter describes all the phases and steps of the research conducted within the Meet-AI project that anticipate the construction of a UX evaluation method for AI-infused products. Moving from the initial assumption that current methods cannot frame the complexity and peculiarities of this novel products representing an opportunity for UX design, a first phase of the research resulted in eight possibly suitable dimensions to describe their UX: *pragmatic, aesthetic, hedonic, affective, intelligence, trustworthiness, conversational, meaningfulness.* After a second phase, starting from a survey to include perspectives external to the research team and comprehending subsequent steps of analysis, a final list of seventeen relevant descriptors (Fig. 3.10) was extracted as a basis on which to build the evaluation scale.

Indeed, the research presents limitations in terms of the number and context of people involved, as well as the subjectivity of methods, decisions and evaluations conducted by the researchers. However, future developments should balance the qualitative work here presented.

Specifically, the next steps should include a solid elaboration of a scale that will need to be tested by a large number of smart speakers' users (as they are the most widespread concrete products integrating the technology under study) to gain statistically valuable information for a definitive validation of a UX evaluation method for AI-infused products. After achieving quantitatively robust results, the method should be generalizable and disseminated to support the design and consequent assessment of devices or services integrating AI systems.

References

Aarts, E. and Ruyter, B. (2009). "New research perspectives on Ambient Intelligence". *JAISE*, 1, pp. 5-14. doi:10.3233/AIS-2009-0001.

Ajovalasit, M. and Giacomin, J. (2019). "Meaning of artefacts". *Conference Proceedings of the Academy for Design Innovation Management*, 2(1). doi:10.33114/adim.2019.02.266.

Algorithmic Watch (2020). *AI Ethics Guidelines Global Inventory*. Available at: https://inventory.algorithmwatch.org/ (Accessed: 27 July 2021).

Amershi, S. *et al.* (2019). "Guidelines for Human-AI Interaction". In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY: Association for Computing Machinery (CHI '19), pp. 1-13. doi:10.1145/3290605.3300233.

Bartneck, C. *et al.* (2009). "Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots". *International Journal of Social Robotics*, 1(1), pp. 71-81. doi:10.1007/s12369-008-0001-3.

Bruns Alonso, M. *et al.* (2013). *Measuring and adapting behavior during product interaction to influence affect*. Berlin: Springer. Available at: https://doi.org/10.1007/s00779-011-0472-3 (Accessed: 28 March 2020).

Clark, L. *et al.* (2019). "The State of Speech in HCI: Trends, Themes and Challenges". *Interacting with Computers*, 31(4), pp. 349-371. doi:10.1093/iwc/iwz016.

Confalonieri, R. *et al*. (2021). "A historical perspective of explainable Artificial Intelligence". *WIREs Data Mining and Knowledge Discovery*, 11(1), p. e1391. doi:10.1002/widm.1391.

Creswell, J.W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: SAGE.

Dourish, P. (2001). *Where the Action Is. The Foundations of Embodied Interaction*. Cambridge, MA: MIT Press.

Dove, G. *et al*. (2017). "UX Design Innovation: Challenges for Working with Machine Learning As a Design Material". In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York, NY: ACM (CHI '17), pp. 278-288. doi:10.1145/3025453.3025739.

Forlizzi, J. and Battarbee, K. (2004). "Understanding Experience in Interactive Systems". In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques. DIS '04 Designing Interactive Systems 2004*, Cambridge, MA: ACM Press.

Forlizzi, J., Gemperle, F. and DiSalvo, C.F. (2003). "Perceptive sorting: a method for understanding responses to products". In *DPPI '03*. doi:10.1145/782896.782922.

Fruchter, N. and Liccardi, I. (2018). "Consumer Attitudes Towards Privacy and Security in Home Assistants". In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal, QC: Association for Computing Machinery (CHI EA '18), pp. 1-6. doi:10.1145/3170427.3188448.

Garcia, M.P., Lopez, S.S. and Donis, H. (2018). "Voice activated virtual assistants personality perceptions and desires: comparing personality evaluation frameworks". In *Proceedings of the 32nd International BCS Human Computer Interaction Conference*. Belfast: BCS Learning & Development Ltd. (HCI '18), pp. 1-10. doi:10.14236/ewic/HCI2018.40.

Google PAIR (2019). *People + AI Guidebook*. Available at: https://design.google/ai-guidebook (Accessed: 5 July 2021).

Hagendorff, T. (2020). "The Ethics of AI Ethics – An Evaluation of Guidelines". *Minds and Machines* [Preprint]. doi:10.1007/s11023-020-09517-8.

Hassenzahl, M. (2011). "User Experience and Experience Design". In Soegaard, M. and Dam, F.R. (eds.). *The Encyclopedia of Human-Computer Interaction*. 2nd edition. The Interaction Design Foundation.

Hassenzahl, M. *et al*. (2013). "Designing Moments of Meaning and Pleasure. Experience Design and Happiness". *International Journal of Design*, 7(3), pp. 21-31.

Hekkert, P. (2006). "Design Aesthetics: principles of pleasure in design". *Psychology Science*, 48, pp. 157-172.

High-Level Expert Group on Artificial Intelligence (2019). *Ethics Guidelines for Trustworthy AI*. European Commission. Available at: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines.

Johnson, D.G. and Verdicchio, M. (2017). "Reframing AI Discourse". *Minds and Machines* [Preprint]. doi:10.1007/s11023-017-9417-6.

Kinsella, B. (2018). *61% of Alexa Skills Still Have No Ratings and Only 1% Have More Than 100 – Voicebot*. Available at: https://voicebot.ai/2018/10/05/61-of-alexa-skills-still-have-no-ratings-and-only-1-have-more-than-100/ (Accessed: 26 February 2019).

Kulesz, O. (2018). *Culture, platforms and machines: The impact of Artificial Intelligence on the diversity of cultural expressions. Information Document*. Paris: UNESCO. Available at: https://en.unesco.org/creativity/sites/creativity/files/12igc_inf4_en.pdf.

Laugwitz, B., Held, T. and Schrepp, M. (2008). "Construction and Evaluation of a User Experience Questionnaire". In Holzinger, A. (ed.). *HCI and Usability for Education and Work*. Berlin, Heidelberg: Springer (Lecture Notes in Computer Science), pp. 63-76. doi:10.1007/978-3-540-89350-9_6.

Levinson, P. (1977). "Toy, Mirror, and Art: The Metamorphosis of Technological Culture". *Learning Cyberspace*, 34(2), pp. 151-167.

Maguire, M. (2019). "Development of a Heuristic Evaluation Tool for Voice User Interfaces". In Marcus, A. and Wang, W. (eds.). *Design, User Experience, and Usability. Practice and Case Studies*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 212-225. doi:10.1007/978-3-030-23535-2_16.

Mekler, E.D. and Hornbæk, K. (2019). "A Framework for the Experience of Meaning in Human-Computer Interaction". In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow: Association for Computing Machinery (CHI '19), pp. 1-15. doi:10.1145/3290605.3300455.

Molnar, C. (2019). *Interpretable Machine Learning*. Available at: https://christophm.github.io/interpretable-ml-book/ (Accessed: 31 May 2021).

Norman, D.A. (2004). *Emotional design: Why we love (or hate) everyday things*. New York, NY: Basic Books.

Polkosky, M.D. (2005). *Toward a Social-Cognitive Psychology of Speech Technology: Affective Responses to Speech-Based e-Service*. Ph.D.

Polkosky, M.D. and Lewis, J.R. (2003). "Expanding the MOS: Development and Psychometric Evaluation of the MOS-R and MOS-X". *International Journal of Speech Technology*, 6(2), pp. 161-182. doi:10.1023/A:1022390615396.

Rijsdijk, S.A. and Hultink, E.J. (2009). "How Today's Consumers Perceive Tomorrow's Smart Products". *Journal of Product Innovation Management*, 26(1), pp. 24-42. doi:10.1111/j.1540-5885.2009.00332.x.

Russell, S. and Norvig, P. (2020). *Artificial Intelligence: A Modern Approach*. 4th edn. Hoboken, NJ: Pearson.

Scherer, K.R. (2005). "What are emotions? And how can they be measured?". *Social Science Information*, 44(4), pp. 695-729. doi:10.1177/0539018405058216.

Spallazzo, D., Ajovalasit, M. *et al*. (2021). "Assessment of Descriptors for UX Evaluation of AI-infused Products". *figshare*. doi:10.6084/M9.FIGSHARE.14387468.V1.

Spallazzo, D., Sciannamè, M., *et al*. (2021). "UX Evaluation Methods Mapping". *figshare*. doi:10.6084/M9.FIGSHARE.14350553.

Spallazzo, D. and Sciannamè, M. (2021). "UX Descriptors for AI-infused Products". *figshare*. doi:10.6084/M9.FIGSHARE.14345498.V1.

White, R.W. (2018). "Skill Discovery in Virtual Assistants". *Communications of the ACM*, 61(11), pp. 106-113.

Yang, Q. (2020). *Profiling Artificial Intelligence as a Material for User Experience Design*. Carnegie Mellon University. Available at: http://reports-archive.adm.cs.cmu.edu/anon/hcii/abstracts/20-100.html.

4. The role of design in the era of conversational interfaces

Ilaria Vitali, Alice Paracolli, Venanzio Arquilla
Department of Design, Politecnico di Milano

With the spread of voice assistants such as Alexa, Google Assistant, and Siri we are witnessing the multiplication of products we can interact by talking using natural language. This chapter refers to these products as "Conversational Smart products", or ConvSP for brevity.

The term *conversational* is used to describe them since they embed and embody Conversational User Interfaces (CUIs). The suffix "Smart" is earned because they possess particular technical features, as they are cyber-physical (merging digital-software and physical-hardware components), networked, and have computational intelligence capabilities.

The first half of the chapter explores the specific traits of Conversational Smart Products, while the second revolves around how conversational interfaces influence products and product design practice. It offers a descriptive analysis framework and strategic design guidelines for novice and expert product designers interested in approaching the design of ConvSP[1].

4.1 What are Conversational User Interfaces? And what does "conversational" mean?

Conversational User Interfaces (or CUIs) are digital interfaces where the primary mode of interaction is a conversation – a repeating pattern of replies and responses (Ashri, 2020). The term CUI implies an inter-

1. This chapter synthesizes and re-elaborates some elements of "hidden for blind review" Ph.D. Dissertation.

face that can use natural language following the structure of a human-like conversation. To be considered "Conversational", an interface must offer the possibility to interact for several dialogue turns, going beyond the first round of interaction and beyond the simple question/answer pairs (Pearl, 2016).

CUIs simulate natural dialogue by applying human conversation patterns (Moore and Arar, 2018) and using spontaneous, informal language (McTear *et al.*, 2016), endowed with greater human likeness.

There are different types of Conversational Interfaces, classified depending on how humans can interact with their system. CUIs include **Chatbots**, **VUIs**, **VPAs**, and **ECAs**.

Chatbots are computer programs that process a user's natural-language input in text form and generate a textual response. Therefore, they converse with users through text messages.

A Chatbot presents itself as a virtual character (a Conversational Agent), with a name and a defined personality, expressed by its writing style.

Chatbots were the first examples of conversational interfaces, and their history goes back to 1966 with ELIZA, the first chatbot developed at MIT by Joseph Weizenbaum.

Chatbots are commonly integrated within websites and applications, including messaging platforms and social media. They communicate through text, including gifs, emoticons, media, and interactive elements. Chatbots are used to assist, complement, or replace human-provided services, especially by companies with large volumes of user interactions. Bots can take the role of routers between humans in a service context, sorting customers with the most appropriate human service provider.

An example is Vodafone's TOBI, which provides 24/7 support and immediate question answering for the telecommunication brand. Bots can also be used for productivity, coaching, alert notification, and facilitating business workflow (Shevat, 2017). Chatbots sometimes replace mobile apps for information retrieval, conversational commerce, companionship, brand engagement, gaming, and entertainment purposes (Klopfenstein *et al.*, 2017; Gentsch, 2019).

The term ***Voice User interface*** **(or VUI)** refers to interfaces that operate **voice as input and output modality** (Pearl, 2016). The term VUI is often used in contrast to GUI (Graphical User Interface) because it specifies a different interaction paradigm **since there are no visual affordances**; its design should follow separate guidelines. Not all VUIs are strictly "Conversational": using speech does not necessarily mean that the interface can simulate a conversation i.e. a smart trash can be opened and closed through voice commands but has no personality nor the ability to dialogue with users on multiple turns.

VUIs are CUIs that can afford a conversation with the user. Indeed, a VUI is conversational when it becomes a Conversational Agent (CA) such as **Virtual Personal Assistants (VPAs)**. Conversational Agents are dialogue systems that conduct natural language processing and respond automatically using human language.

VPAs (Like Alexa, Google Assistant, and Siri…) are powerful **cloud-based** dialogue systems that can be **integrated into various devices.** They usually allow both voice and chat-based interactions. VPAs support a broad set of functionalities and become platforms for creating conversational applications and services. Third-party developers can create applications that are compatible with VPAs. i.e., Amazon Alexa calls these apps *Alexa skills*, while Google calls them *Actions*. These apps can be found, downloaded, and added to the assistant's capabilities at any time.

VPAs can be defined as super bots that facilitate and manage multiple services (Shevat, 2017). **Compared to chatbots, VPAs can carry out a broader set of tasks.**

Although there is the ambition to build VPA companions, most VPAs remain transactional tools, not yet able to create long-term relationships with users (Luger and Sellen, 2016), and aimed at facilitating precise tasks rather than being digital friends.

However, any Conversational Agent is characterized by features designed to communicate its personality and character (e.g., tone of voice, specific wording and use of language, use of emojis, visual elements, etc.). Among those aspects, particular attention should be given to the choice of the tone of voice assigned to the virtual assistant, because it could contribute to strengthening stereotypes. Voice assistants nowadays mainly adopt female voices reinforcing gender prejudices, since these objects use female voices and have a subdued tone and role. The female voice could be associated with a submissive figure, negatively stereotyping the woman's figure, unconsciously linking the woman's skills to the housekeeper (Hall, 2018, p. 83). At the same time, artificial voices should be more varied, including male and neutral tones of voices. The designer should use this opportunity to represent individuals who have not yet been considered in the world of artificial voices, such as non-binary individuals (UNESCO, 2019). However, voices without clear gender markers or discordances between voice and personality are perceived as less clear (Nass and Brave, 2005). Therefore, the designer needs to pay special attention to the tone of voice and investigate how voice assistants can be designed not to perpetuate gender biases, valuing the brand identity.

Embodied Conversational Agents (ECA) are the last category of CUIs. They are agents in the form of animated characters on screens. ECAs combine verbal and non-verbal communication signals synchronously. They can use body posture, gaze, and hand gestures, all synchronized with the verbal dimension, enhancing and complementing it.

ECA research aims to transport the richness of human face-to-face communication to the interaction with computers to get a more intuitive and engaging interface, able to recognize and reproduce emotions and expressivity (Cassell *et al.*, 1999). Embodied characters range from human-like characters, i.e., news anchor-man, to more stylized virtual companions such as animated animals. ECAs may exist virtually as software or become part of physical objects (Mctear *et al.*, 2016).

So far, we have defined ConvSP as conversational because they embed and embody Conversational User Interfaces (CUI). Moving forward, we also described them as "Smart".

4.2 But, what does it mean for a product to be smart?

Nowadays, it is common to use the word Smart to indicate physical objects that process information and showcase a certain degree of intelligence.

Smart's suffix is used with technological significance: it represents the intelligence obtained through embedded IT technology. We refer to Smart Products as a category including internet-connected consumer electronics that possess three main technical characteristics: cyber-physical, networked, and computational intelligence.

Smart products blend hardware and software. They are physical objects with a digital representation. Material things need a digital counterpart to represent them on the network and be part of the IoT (Internet of Things).

The term cyber-physical (Abramovici, 2014) has been used to describe this dualism.

Globally unique IDs identify smart products and make them accessible for remote control and communication during their product lifecycle (Gutierrez *et al.*, 2013; Kärkkäinen *et al.*, 2003). Interaction is a significant aspect of smartness and can occur through multiple interfaces on physical and digital touchpoints. Maass and Janzen (2007) stress that smart products' interfaces are dynamic: they can offer real-time communication and use data (about the product itself or its surroundings) to be localized in time and physical spaces. Various modalities can be used as input and output (Sabou *et al.*, 2009). For example, a product could synchronize a physical UI and an external web application.

The second common characteristic of Smart Products is the connectivity: Smart Products are networked. Thanks to unique identification numbers and an Internet Protocol (IP), they can connect to the Internet and other products using different wired and wireless communication technologies (Abramovici, 2014; Greengard, 2015).

Internet connection makes them part of a larger network of things, people, and services. They can communicate, bundle, and interoperate with other devices (Maass and Janzen, 2007; Gershenfeld *et al.*, 2004). They can connect with the encompassing environment (i.e. a smart car that connects to a bigger infrastructure at the city level) and with peer products that dynamically become available (i.e. a smart home product that detects a new smartphone to connect to) (Mühlhäuser, 2007). Machines can automatically interact with other machines in bidirectional exchanges of information (Boswarthick *et al.*, 2012).

Connectivity gives devices access to services and capabilities external to the physical product, such as cloud services. It can also be used for interaction purposes (i.e., remote control of a connected device) and to communicate to users, even proactively (i.e., a smart product sending a notification, Kärkkäinen *et al.*, 2003).

In addition, the capability to connect enables smart products to be revised thanks to "over the air" updates. In time, a smart product can expand with new functionalities and evolve. This ability brings implications for product designers building new conceptions of products.

Moreover, smart products can collect data about themselves and their environment (Gutierrez *et al.* 2013; Lyardet and Aitenbichler, 2007). Once shared, collected, and aggregated, data become a valuable asset at the economic level and for the development of future products and services (Greengard, 2015).

Lastly, Smart products have computational intelligence as they embed electronic "brains" (processors and microprocessors) able to process data and perform programmed behaviors.

Their "intelligence," as in the ability to handle information or carry out decision-making, that can be located not necessarily at the device level (Meyer *et al.*, 2009). Intelligence could occur inside the physical product or even outside, such as in the cloud, or be unloaded to a different device like a smartphone (McFarlane *et al.*, 2003; Kärkkäinen *et al.*, 2003). Smart products collect, process, and produce information (Rijsdijk and Hultink, 2009). They display autonomous and proactive behaviors and can operate independently, performing tasks without the need for direct user interaction (Sabou *et al.*, 2009). They can often learn from experience and infer patterns and high-level events from data (e.g. understanding the preferences of a specific user). This enables smart products to display forms of awareness and evolve their performance in time.

4.3 How do Conversational interfaces impact the product design of smart objects?

In this vast category of devices, the impact of conversational interfaces on physical products is little explored by academic research, especially in the design field; most existing research focuses on the user experience and interaction with smart speakers and social robots. It is less explored how CUIs physically get embodied in products outside these two categories. From the analysis of 40 Conversational Smart Products (Vitali, 2020), it emerged that there are three qualities common to any ConvSP: they are related to a conversational agent (by embedding, embodying, or being able to connect to a CA), have their own physical shape, and tangible parts. Each of these qualities manifests differently depending on the kind of ConvSP.

4.4 Different ways in which a Conversational agent can be infused into a physical product

- It is **physically embedded** when the CA is built-in in the product, but it has an overall low impact on its shape, as the product design doesn't try to deeply be "the body" of the CA. For example, a built-in Alexa washing machine: the product should integrate a speaker and have available internet connectivity to relate to Alexa and answer user's requests. Still, the washing machine design remains pretty traditional.
- When the CA is **physically embodied**, the product makes it more evident that there is a conversational intelligence built-in inside and communicatively becomes the "body" of the agent. The products can interact with the user through more complex and expressive feedback. For example, an expressive smart speaker, whose design clarifies the presence of a conversational agent and uses interactive feedback to be more explicit.
- The embodiment can even be **remote** when the input/output of the conversation happens on another device. It is the case of those products that are compatible with assistants but don't have them built-in. *E.g.* a TV remote is used to record the user input, yet the agent interacts on the TV.

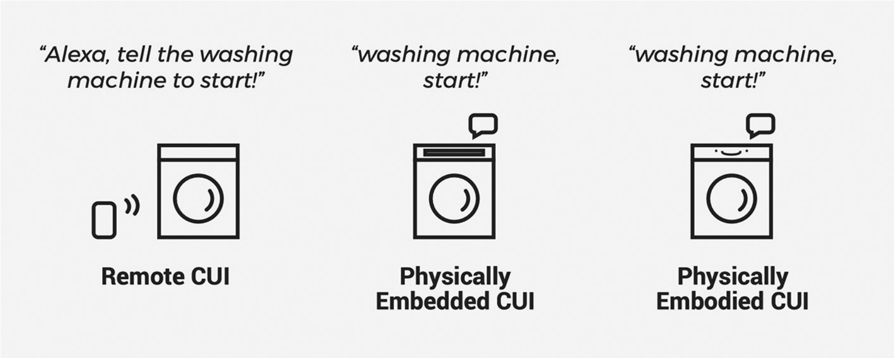


Fig. 4.1 – Different ways CAs can be infused into a product.

On the market, ConvSP primarily embed and embody the virtual assistants developed by large telephony and e-commerce companies such as Amazon's Alexa, Google's Assistant, Apple's Siri, Samsung's

Bixby, and Alibaba's Aligenie. The strategy of these large companies revolves around integrating the assistants' functionality into any product. Their VPAs are integrable, powerful, and generalist, because the same assistant should be able to be integrated into a variety of products and perform different tasks. It also means that there are very few Conversational Agents with "different voices" and personalities.

Indeed there are less specialized voices, typically present in domestic robots, toys, specific assistants (e.g., a VPA to teach a particular language), and products with a defined purpose (e.g., a washing machine with a chatbot assistant to set the washing cycle). Specialized voices own proper names and personalities, with precise capabilities and are not ideal for generalist assistants.

4.5 How does the presence of a CUI affect the shape of a product?

First, the impact of the CUI's presence manifests in the physical shape with the need to display the conversational agent's presence through status indicators and talking/listening feedback. The product shape must communicate that it contains an active intelligence able to listen and respond; it is a right for whoever enters a "monitored" space, as the presence of agents who can listen to us extensively worries users who care about their privacy.

Common Feedback components such as light effects, animations on screens, or simple icons can concur to make it evident. The Feedback components can also become branding elements that make the presence of a particular assistant recognizable on different devices (e.g., Alexa-based products use blue light rings).

The agent's presence should always be visible, even through feedback that appears only when a request awakes the assistant.

It was observed that the presence of the CUI doesn't have a substantial impact on the physical design of the product, as more abstract, anonymous shapes and compact dimensions are preferred, as they easily fit into the home environment. More anthropomorphized or biomorphic shapes are less common, as they are less flexible for commercial applications.

Similarly, this happens for objects where the CUI is not the primary mode of interaction. In these cases, products generally maintain a shape more linked to their primary function (i.e., a monitoring camera with an integrated assistant may not make evident the presence of the assistant).

Among the Conversational Products analyzed in the study (Vitali, 2020), only a limited number employed shapes with biomorphic or anthropomorphic features. It was the case in which products tried to create a deeper connection with users (i.e., robot companions, toys, voice assistants to teach languages, Smart speakers intended for kids…). In this case, the choice of these shapes is used to give a deeper embodiment to the Conversational Agent.

Deeper embodiment impacts users' perception, who tend to attribute more intelligence, animacy, and empathy to the agent, increasing the user's expectations of its ability to dialogue. Indeed, the level of embodiment and anthropomorphism in the design needs to be evaluated according to the device's capabilities, because it may cause a mismatch in user expectations.

4.6 Tangible aspects of interaction and feedback

ConvSP usually integrates tangible controls, and screen-based interaction, and produces listening and talking feedback when interacting.

Tangible controls are a shortcut to interact with the agent, to provide shortcuts to control the agent into its most frequent operative functions. E.g., wake up the agent, control volume, move music forward and backward, and "mute" button to interrupt the agent's listening capabilities. Indeed, users tend to use conversational functions only when it is the most convenient, fast, hands-free way to perform a task (Luger and Sellen, 2016). Therefore, offering tangible alternatives to voice for the frequent and standard functionalities increases the interaction's efficiency, resulting in a fundamental part of the user experience.

The integration of touch screens can increase efficiency. It is visible from the market that displays are the future add-on to ConvSP that will simultaneously possess voice and touch screens.

Proof of this trend can be seen by following the evolution of the smart speaker category, which has gone from being objects without a screen to full-fledged 'tablets' with speaker capabilities (e.g., Google Nest Hub, Amazon Echo Spot, Amazon Echo Show, Lenovo Smart clock).

The combination of VUI and screen allows mitigating some of the negative aspects of voice interaction, such as:

- The difficulty of discovering the CA's abilities.
- The cognitive load for recalling information instead of recognizing them as in GUIs (Murad *et al.*, 2018).
- The slowness of the conversation compared to other modes.
- The difficulty in performing predominantly visual tasks such as managing lists or multiple choices.
- It also provides flexibility in updating the interface and the product.

Smart screens can upgrade the visuals and screen-based content that complement speech. Screens also allow additional functions such as becoming a clock or making video calls (e.g., Amazon Echo Spot).

Whether the ConvSP has a screen or not, Listening and Talking feedbacks are their primary expressive cues. Listening (and thinking feedback) frames the product's interactive behavior when the user is speaking and when the system processes a response. While talking, feedback concerns the device's behaviors while generating a response to the user's request.

This feedback is generally obtained in the products through animated lights and on-screen animations. As already anticipated, they become branding elements that communicate the agent's presence and let users feel how advanced the solution is.

Deepening the non-verbal communication skills of CUIs & ConvSP is one of the intervention areas for designers. In particular, a greater synchronization of speech and visual feedback makes speech more expressive (i.e., using paralanguage, prosody, and a greater variety of responses).

Few ConvSP employ Embodied Conversational Agents (ECA) on screens in the form of animated characters or partially animated features (e.g., faces, expressive eyes). ECA is also common in robotics to create Expressive face robots (Breazeal, 2003). In products, the same screens used to simulate the character often double as interactive surfaces to interact with the product. In this way, the screen takes the role of both the product's face and an input modality for the user.

4.7 Defining five categories of Conversational Smart Products

Embodiment, Tangibility, and Shape are the three common characteristics of any Conversational Smart Product, but they can manifest differently into the devices. All the possible variables have been grouped in the descriptive framework proposed by Vitali and Arquilla (2019), composed by five units of investigation.

1. **Type of CUI**: it specifies which kind of conversation interface is used by the product (i.e., **VUI, VPA, Chatbot, or ECA**).
2. **Type of Conversation:** a conversation with a digital agent can be defined as ***task-led*** (transactional – single user request, the CA focuses on accomplishing a concrete task) or ***Topic-led*** *(user* discusses in a multi-turn interaction, the CA discusses, exchanges ideas on a set of subjects, and maintains positive relationships with users).
3. **Conversational Level**: describes the complexity of the CA's abilities. **Low** (based on a script with limited command) ***Limited*** (support user's initiative on a specific topic) **Moderate** (CUI supports *user initiative and mixed-initiative on a few topics)* ***High*** *(can manage multi-turn interactions and follow-up questions and support topic-led conversations).*
4. **Type of CUI Embodiment**: how the CA is physically **integrated** and **embodied** into a **tangible** object.
5. **Details of tangible interaction and feedback**: the way ConvSP communicate their conversational capabilities in the form of inputs, outputs, and feedback. Does the agent have **Tangible controls?** Does it give any **feedback** when **listening** and **talking?** Does it support **Screen-based interaction?**

The descriptive framework successfully portrays and categorizes the currently existing ConvSP, distinguishing their functionalities and common traits. It is also a fundamental guide for designers of ConvSP.

The descriptive framework identifies **five types of ConvSP:**

1. **Conversational Companions**: this ConvSP category has deeper embodiment and higher conversational capabilities and expressive feedback: they can support topic-led dialogues and multi-turn inter-

actions. This category includes social robots and toys, usually used as companions, teachers, and personal assistants.

i.e., ***Lily speaker by Maybe*** is a smart speaker designed to teach the Chinese language and practice simulated conversations in different settings. Lily uses speech recognition to understand what users say and speech synthesis to respond with a human voice; it can also display expressive facial expressions to better communicate with users. The product has an accompanying app where users can chat with Lily and exercise their grammar.

2. **Products with Physically Embodied Agents**: their primary function is to give a body to a conversational agent with transactional purposes (like Amazon Echo or Google Home). The product is the shell of the agent, displaying its presence by using communicative feedback (e.g., specific light patterns) concurrent with conversational output. The limit of this category is that it usually lacks a clearly defined "killer functionality", possessing an unclear purpose instead. They are currently used for playing music, hands-free search, IoT control, and entertainment. Therefore, when designing this typology of ConvSP, it is necessary to define the agent domain of expertise to frame the product's purpose and benefits correctly.

   i.e., ***Amazon Echo Dot***[2] (3rd generation) (visible at Amazon.it) is a compact smart speaker that integrates Amazon Alexa. It can play music, answer questions, control compatible devices, make calls, and use Alexa skills. It has four built-in microphones to listen from all directions and gives visual feedback through a luminous ring. The light ring produces different effects depending on the state of your request (e.g., it blinks if it doesn't understand the query). Alexa can be muted with a corresponding button, and there are physical controls for the volume.
3. **Products with Physically Embedded Agents,** their CUI is not the only mode of interaction, nor is the primary. They tend to maintain a more linked shape to the object's primary function and yield more simplified talking and speaking feedback. The limit is that their

2. Although chapter 1 already takes Amazon Echo as an example to analyse the landscape of all Smart Domestic Assistants, in this chapter, we instead provide the same products to position a specific type of smart assistant within the landscape of all smart conversational objects. Indeed, Amazon Echo is one of the most famous and familiar "Products with Physically Embodied Agents", ideal to represent this category of ConvSP.

ability to listen and converse is not always evident. The first question for the product designer is to understand if the conversation skills are coherent with the product's primary purpose. If so, how to declare the agent's presence and conversational ability (light effects are commonly used).

i.e., ***Nest Cam IQ*** (visible at store.gogle.com) is a smart security camera with a built-in Google Assistant. The cam uses infrared LEDs to see in the darkness and automatically recognize familiar faces. It supports a monthly monitoring service 24/24h. When the assistant recognises the wake word, a light ring hidden on the cam perimeter fades with a standard animation. The light remains on when the product is talking and switches off afterward.

4. **Conversational Screens** are smart screens with conversational capabilities that present outputs in both spoken and written/displayed modalities. Displays can support conversational interactions and provide additional information on the screen. It benefits from both VUI and GUI interactions adding considerable flexibility, it reduces the amount of training required for the user. Screens can enrich conversational interaction by providing additional content synchronised with speech. This doubles the design effort but has the advantage that it can be adapted flexibly**.** i.e., Amazon Echo Spot and Google Nest Hub (visible at store.google.com).
5. **Conversationally** Enabled products have partial conversation capabilities. Inputs and outputs may happen via an external interface (e.g., a smartphone) or require the support of other products to work (e.g., a conversational remote control for a smart TV). Depending on the application, these products can have limited, more simplified conversational abilities. These products only work when paired with another device, but they are compatible with several products on a positive note. The downside of this product category is the discoverability of their conversational ability. If the product doesn't communicate this capability adequately, e.g., through dedicated icons on the product, the functionality may remain unnoticed. Examples for this category are Alexa Fire TV remote and Apple Air Pods (visible at apple.it).

This categorization of Intelligent Conversational Products is a starting point for exploring an area still little covered, especially for product design researchers and professionals.

4.8 How to design Conversational Smart Products

With the birth of smart speakers and the pervasive diffusion of virtual assistants within products, we witness the first waves of conversational products (ConvSP) at home. This is a relevant change, challenge, and occasion for product designers that continuously need to be updated. Indeed, for designers that wish to explore this category of objects, this is a precious design guideline for any kind of conversational smart product.

Considering the different insights on how conversational interfaces impact products in terms of embodiment, shape, and tangibility, it is possible to identify eight critical aspects in the design of ConvSP grouped into two categories. Three more strategic dimensions should be considered in the early stage of the project, and five dimensions related to product design and the physical and formal aspects of ConvSP.

Strategic Dimensions

1. Assess the conversation's purpose and its desirability: Why should the product become conversational? This is the first big question for ConvSP designers. Not every product should have a built-in conversational agent, especially those whose main feature is not conversation (i.e., a toothbrush). Thinking critically about this aspect allows designers to identify the most promising configurations for integrating conversational interfaces into physical products. This means understanding which type of CUI would fit better and the ideal embodiment scenarios. For example, would it be better to create a proprietary conversational agent or use an available one like Alexa? Will the product communicate through voice or text? Should the product physically integrate the agent (built-in agent) or need another device to work (remote agent embodiment)? This phase opens up different directions for product development, impacting the product shape, interaction, and overall user experience.
2. Assess the Conversational level: the strategic decision consists of understanding how advanced the conversational capabilities will be (low, limited, moderate, or high). The scale considers initiative (the ability of both user and system to lead the conversation and ask questions), the breadth of the domain of expertise, and the ability to follow up responses and go beyond the first round of discussion. It

is an essential technical requirement that defines consistent forms and interactions for the ConvSP aligned with user expectations and doesn't oversell the agent's capabilities.

3. Assess embodiment necessity: which will be the correct configuration in terms of embodiment: remote, physically embedded, or embodied?
   The need for embodiment (achievable in terms of anthropomorphic/expressive shapes, and through the complexity of product behaviors, feedforward, feedback, and movements) depends on functions, technical limitations, and on which will be the assistant's role. Not every product needs to embody the presence of the conversational agent deeply. i.e., an object that wants to teach a language may need to communicate its presence strongly, while if it is an intelligent microwave, it may not be necessary. It needs to be evaluated depending on the product's core functionalities and the relationship with the user.

Physical Dimensions

4. Visibility of agent's presence: how to physically communicate the agent's presence. The visibility of the agent's presence contributes to raising user trust: a little evidence of the agent's presence makes ConvSP be perceived as intrusive because they listen "in secret". (i.e., lights feedback for listening and speaking behaviors of the agent).
5. Antrophomorpic or machinelike shapes? The existing user experience scales to evaluate robots differentiate between anthropomorphic/biomorphic or machinelike/abstract forms (Bartneck *et al.*, 2009; Bartneck and Forlizzi, 2004), and they can also be applied to ConvSP. The designer should consider that the more the level of anthropo/biomorphism is present in the product, the higher the users' expectations for intelligence and conversational capabilities will be.
6. The complexity of listening and talking feedback: this is a relevant dimension for product design because they are the expressive behaviors perceived by the users as manifestations of the agent's intelligence. The user perceives ConvSP with more complex talking feedback as more capable and intelligent than those with simpler

limited feedback. Examples of simple listening and talking feedbacks are static signals such as a LED light switching on and off; more complex behaviors could be performing various messages or synchronizing with the content and form of speech.

7. The need for physical controls: As users strive for efficiency, they will use voice only when it is the best way to interact. Physical controls can still prove valuable for users for more frequent interaction (such as on/off and volume).
8. Branding value of CUI: The static and dynamic elements of the device, such as listening and speaking feedback, are not only seen as signals of agent embodiment; they are also branding elements. They make recognizable and identifiable the presence of a specific VPA agent. That's why the designer's job is not just to find expressive solutions for each conversational product. Still, it is also to define or apply a design language that allows its expansion to other products. As the same VPA could be embedded and embodied in multiple devices, it is necessary to design rules for coordinating the agent's presence, making it distinctive and compelling. For example, many products with Alexa built-in use a blue light ring, a light bar, or the assistant's logo as visible trademark elements, and Amazon provides style guidelines with rules to follow for the agent's physical integration.

4.9 Conclusion

The constant increase in products' processing capabilities and connectivity facilitates the diffusion of smart products with conversational interfaces. Starting from the reference of this chapter and looking at the birth of an important field called Conversational User Experience Design (Moore and Arar, 2019), it is always more evident that product designers need to improve and incorporate more skills. Designers interested in specializing in this path must become curious conversational experts and draw knowledge from different disciplines that study human conversation.

Doubtless, technology has changed and will keep changing the way we live, transforming our daily objects into devices. Therefore, designers should continually be updated on new technologies to adequately embed them in everyday objects and change their ordinary meaning.

Thinking from a conversational point of view does not only mean that an object uses Voice Control, but that the products themselves acquire conversational logic that transforms the object and its usefulness and functionality. The designer must think about how the user will be recognized, through what technology.

How will the device react to events and adjust coherently to the environment and context? Will it be able to learn by the user and adapt its behavior over time? Will it anticipate users' plans and intentions?

These features make smart products such an opportunity for designers to build the next generation of daily objects and user interactions.

They raise new challenges, such as designing interactions distributed across multiple devices, in which the focus of the user experience is the service and not the product itself (Vitali and Arquilla, 2019).

Intelligent products can perform multiple functions that often are not evident, traceable, or geometrically comprehensible in the artifact's shape. This leads to anonymous bodies that do not communicate or invite users to interact. Many smart products can't even be manipulated, since interaction occurs through external interfaces and smartphones (Vitali *et al.*, 2019).

Therefore, the ability of the designer is to make the design comprehensible to the receiver/user by presenting qualities that will cause and fulfill certain expectations (Kazmierczak, 2003).

The guidelines given in this chapter are a precious starting point that foster innovation and creative thinking around possible variables for Conversational Smart Product.

The last but fundamental prerequisite for designers is sustainability, which must be considered a constraint in the design of ConvSP, and in general, for any type of smart object. Any user interaction requires a vast planetary network, powered by the extraction of (currently) non-renewable materials, labor, and data. This type of object creates an ever-growing network of data-consuming objects that stay on forever, using incredible amounts of energy. The way users interact with products leads to substantial environmental impacts (Tang and Bhamra, 2009). In this regard, the designer must take into consideration two further aspects: the amount of energy consumption invested in the life of the designed product, given the quantitative aspect of its energy demand; usefulness for people, to understand when a product significantly improves people's lives and if it is worth it to consume energy.

Given that design usually focus only on the end users, Circular Design approach instead considers the whole system in which the design will exists and its consequences. The designer needs to foresee the design impact. Thus it is fundamental to assess both user's needs and the systemic implications of the product in society and on earth, continuously analyzing these two equally critical perspectives (Circular Design Guide by Ellen Macarthur Foundation). The designer should ponder this trade-off when designing ConvSP.

This is still an open question that reminds us that the professional practice of design is changing in two ways. The first is hybridization; the product designer must now be able to handle knowledge from different disciplines and stay up to date on a technological level. The second is the integration of product design by digital designers who could find themselves interfacing with physical products, in a transversal exchange of knowledge.

These dimensions of hybridization and integration also refer to a cultural background that should not be forgotten. Indeed, the designer of the future should handle three elements: acknowledge traditional design practice, owing marketing skills with a design thinking perspective, and have computational skills to comprehend

smart objects. In his book "How to speak machine", Maeda (2019) proffers every designer to have this type of skill.

References

Abramovici, M. (2014). "Smart products". *CIRP Encyclopedia of Production Engineering*, 1-5.

Agrò, L., (2018). *IoT designer: Progettare oggetti e servizi relazionali*. Milano: FrancoAngeli.

Ashri, R. 2020). *AI Is the New UI. In The AI-Powered Workplace* (pp. 83-92). Berkeley, CA: Apress.

Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S. (2009). "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots". *International journal of social robotics*, 1(1), 71-81.

Bartneck, C., and Forlizzi, J. (2004). “A design-centered framework for social human-robot interaction”. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication*. pp. 591-594. IEEE.

Breazeal, C. (2003). “Emotion and sociable humanoid robots”. *International journal of human-computer studies*, 59(1-2), 119-155.

Brown, T. (2022). *Circular design*. [online] https://archive. ellenmacarthurfoundation.org. (Accessed 1 December 2021).

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H., and Yan, H. (1999). *Embodiment in conversational interfaces: Rea*. 520-527.

Garcia, D.M.P., Lopez, S.S., and Donis, H. (2018). *Voice-activated virtual assistants*.

Gentsch, P. (2019). “Conversational AI: How (Chat) Bots Will Reshape the Digital Experience”. In *AI in Marketing, Sales and Service*. Cham: Palgrave Macmillan, pp. 81-125.

Gutiérrez, C., Garbajosa, J., Diaz, J. and Yagüe, A. (2013, April). “Providing a Consensus Definition for the Term Smart Product”. In *2013 20th IEEE International Conference and Workshops on Engineering of Computer Based Systems (ECBS)*. IEEE, pp. 203-211.

Kärkkäinen, M., Holmström, J., Främling, K., and Artto, K. (2003). “Intelligent products – a step towards a more effective project delivery chain”. *Computers in industry*, 50(2), 141-151.

Kazmierczak, E.T. (2003). “Design as meaning making: from making things to the design of thinking”. *Design Issues*, 19(2), 45-59.

Klopfenstein, L.C., Delpriori, S., Malatini, S., and Bogliolo, A. (2017). “The rise of bots: A survey of conversational interfaces, patterns, and paradigms”. In *Proceedings of the 2017 conference on designing interactive systems* (pp. 555-565).

Luger, E., and Sellen, A. (2016). “Like Having a Really Bad PA”. *The Gulf between User Expectation and Experience of Conversational Agents*. 5286-5297.

Maass, W., and Janzen, S. (2007). *Dynamic product interfaces: A key element for ambient shopping environments*.

Maeda, J. (2019). *How to speak machine: Laws of design for a digital age*. Penguin UK.

McTear, M., Callejas, Z. and Griol, D. (2016). *The Conversational Interface: Talking to Smart Devices*. Berlin: Springer International Publishing. https://doi. org/10.1007/978-3-319-32967-3.

Moore, R.J., and Arar, R. (2019). *Conversational UX Design: A Practitioner’s Guide to the Natural Conversation Framework*.

Murad, C., Munteanu, C., Clark, L., and Cowan, B.R. (2018, September). "Design guidelines for hands-free speech interaction". In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (pp. 269-276).

Nass, C.I., and Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship*. Cambridge, MA: MIT Press, p. 9.

Pearl, C. (2016). *Designing voice user interfaces: Principles of conversational experiences*. Sebastopol, CA: O'Reilly Media, Inc.

Sabou, M., Kantorovitch, J., Nikolov, A., Tokmakoff, A., Zhou, X., and Motta, E. (2009). "Position paper on realizing smart products: Challenges for semantic web technologies". In *CEUR workshop proceedings* (vol. 522, pp. 135-147).

Shevat, A. (2017). *Designing bots: Creating conversational experiences*. Sebastopol, CA: O'Reilly Media, Inc.

Tang, T. and Bhamra, T.A. (2009). *Improving energy efficiency of product use: An exploration of environmental impacts of household cold appliance usage patterns. The 5th International Conference on Energy Efficiency in Domestic Appliances and Lighting EEDAL'09*, Berlin, Germany, 16-18 June, 2009.

UNESCO (2019). *Explore the Gendering of AI Voice Assistants*. UNESCO. en.unesco.org/EQUALS/voice-assistants.

Vitali, I. (2020). *A conversation on Conversational Smart Products: how the diffusion of Conversational Interfaces impacts Products and the Design Practice*, Ph.D. Thesis, XXXIII Cycle, Politecnico di Milano.

Vitali, I., and Arquilla, V. (2019). "Conversational smart products: a research opportunity, first investigation and definition". In *DeSForM 2019. Beyond Intelligence* (pp. 60-68).

Vitali, I., Arquilla, V. and Rifino, I. (2019). "Design for Meaning of Smart Connected Products". In *Academy for Design Innovation Management 2019* (pp. 1189-1202). Academy for Design Innovation Management.

5. Understanding meaningfulness in AI-infused artefacts

Marco Ajovalasit
Department of Design, Politecnico di Milano

Eternal truths will be neither true nor eternal unless they have fresh meaning for every new social situation.

Franklin D. Roosevelt, 1940

Creating meaningful artefacts translates into creating meaningful experiences for the consumers going beyond the artefact's functional features to reach people intrinsic motivation and their symbolic values. This process would require acquiring an understanding of how people come to understand in their own terms and for their own reasons the meaning attached to the designed artefacts. Despite the advances on the market of AI-infused products, the sophistication of modern technologies of these designed artefacts and the complexity of modern social behaviors would make simplistic to continue to consider the user experience (UX) mostly characterized by the performance of its users or in terms of interaction quality. Relatively little empirical work has instead focused on understanding why designed artefacts are seen as salient choice criteria, that is, understanding the reason why do consumers perceive AI-infused artefacts to be personally relevant for their needs. The understanding of the concept of meaningfulness discussed in this chapter suggests that the consideration of target values and meanings is important as a strategy towards defining the relational role of AI-infused artefacts in the lives of their owners for which artefacts are not simply functional tools, but are also relational mediators which shape the long term aims, objectives and behaviors of an individual or of a group. The research presented in this chapter aims to set the basic considerations regarding the term "meaning" used by commercially active designers and introduces the concept of "meaningfication" for the purpose of designing AI-infused artefacts based on new meanings.

5.1 Introduction

Nowadays many everyday products are equipped with some functionality enabled by artificial intelligence (AI) (Follett, 2015). Examples span from mobile devices in the form of activity trackers, mail filters, autocomplete, and social network feed ads to home assistants like Amazon Echo with Alexa, Apple HomePod with Siri, Google Home with Google Assistant that help users to accomplish simple tasks (Sciuto *et al.*, 2018; Spallazzo *et al.*, 2019; Vitali *et al.*, 2019).

Previous research on AI-infused artefacts has predominantly focused on the utilitarian attributes and usability dimensions (Amershi *et al.*, 2019), on the use of conversational agents (Sciuto *et al.*, 2018) and on the tangible form and appearance as well as the interaction modalities (Spallazzo *et al.*, 2019).

While on one hand the advances in the AI domain (Follett, 2015) and the growing uses of AI technologies in human-facing applications provide opportunities for user interface design (Amershi *et al.*, 2019) proposing generally applicable design guidelines for human-AI interaction, on the other hand there are current challenges of how the use of such technology may bring value to their users and how to design them in order to mean something in their users' world and therefore to be meaningful to them.

Although Weiser's (1991) ubiquitous computing vision of embedding information processing and network communication as key components in the design of everyday objects and human environments to make familiar tools and environments do their jobs better and help people in their ordinary activities (Follett, 2015; Kaptelinin and Nardi, 2009) may have existed thirty year ago, the complexity of technology of the algorithmic reasoning (intelligence) and the design of the embedded systems may have overshadowed nearly all the consideration of the user experience and motivation, whether the designed artefact resulted useful to full fill desires, to be enjoyable or meaningful to their users. As noted in Gartner's hype cycle (Blosch and Fenn, 2018) product innovations like, self-driving cars and personal assistants which use AI-infused artefacts, follow "a typical progression of innovation, from overenthusiasm through a period of disillusionment to an eventual understanding of the innovation's relevance and role in a market or domain". The use of such AI-infused designed artefacts is rarely the most important activity in someone's life, but the artefacts form part of a larger flow of needs, desires and meaningful experiences (Kuniavsky, 2010).

In the context of this exposition, the word "meaningfulness" is described by the definitions found in standard dictionaries of the English language as "the fact of having a serious or important meaning", and "the quality of having great value or significance". Significance, as a concept in design, explains how forms assume meaning in the ways they are used, or the roles and meaning assigned to them, often becoming powerful symbols or icons in patterns of habit and ritual. In contrast to the emphasis on efficiency or experience significance has more to do with expression and meaning (Ravasi and Rindova, 2008). Also in this exposition, AI-infused artefacts refer to artefacts that have features harnessing AI capabilities that are directly exposed to the end user (Amershi *et al.*, 2019). AI-infused artefacts leverage computers and machines to mimic the problem-solving and decision-making capabilities of the human mind (IBM Cloud Education, 2010).

The aim of this chapter is to draw attention to the most obvious and influential issues which affect the meaningfulness of a designed artefact and to note the implications for the design of AI-infused products. The objective is to place into perspective the basic considerations regarding the term "meaning" used by commercially active designers. The reflections conclude by introducing a "design for meaning" framework which organizes one possible sequence in which the various considerations might be dealt with, and which provides a tool for identifying key questions which should be answered by commercially active designers. The concept of "meaningfication" is introduced for the purpose of designing AI-infused artefacts based on the new meanings that the AI-infused artefacts are intended to provide or to facilitate for the consumer.

5.2 Conceptualising the term "meaning"

Within market-driven economic systems the commercially active designers must consider the forms of value and meaning which a product, system or service may hold for its customers. It is in fact frequently claimed that the value and meaning of a commercial offering is the actual basis of the business (Verganti, 2009).

Given the importance of “*meaning*” and “*value*” in design it is worth noting the definitions of the concept of two words which are often used interchangeably, with important practical consequences in terms of possible misunderstanding.

According to standard dictionaries of the English language, the word “meaning” can express at least three possible concepts:

- the sense or signification of a word or sentence;
- the significance, purpose or underlying truth of something;
- the motive or intention of something.

The word “value” can express instead:

- the amount of money that can be received for something;
- how useful something is;
- the importance or worth of something for someone.

Neglecting the purely linguistic sense of the word “meaning” and the economic aspect of the word “value”, the questions can be asked of how design artefacts assume a purpose or intention and why they have value to people. According to Baudrillard (1968) he views value as a meaning, that is, people value objects not for what they do, or what they are made of, but for what they signify. According to Richins’ (1994), based on the measurement of the value of physical possessions, “an artefact’s value derives from its meaning within the cultural system”. Meaning should be then considered as the source of value. Both anthropologists and neuroscientists (Diller *et al.* 2005) agree that “meaning is the sense we make of reality. Assigning meaning to experience is how each of us creates the story of our life and its ultimate value and purpose”. In the context of designed artefacts, researchers such as Diller *et al.* (2005) have highlighted a main difference between values and meanings claiming that “values involve preferences; they represent our choices between opposing modes of behavior, and they are shaped not only by ourselves, but also by those around us”. Whereas “meaning provides a framework for assessing what we value, believe, condone, and desire”.

In conceptualizing “meaning”, this exposition doesn’t follow the ontologizing of meaning as it were an entity that could be attached to objects implying that the meaning would be the same for everyone. It also avoids the representationalism of semiotic discourse for which artefacts would signify something unrelated to its use (Williamson, 1978). The concept of meaning discussed in this exposition is assumed to be "the full set of interactions and experiences associated with a specific spatial-temporal event". It doesn't focus on the semiotic studies that concentrate on the artefact itself as the most important site of its meaning, but it focusses on social semiotics that emphasizes the social effects of an artefact's meaning placing the emphasis on the social modality of meaning-making (Hodge and Kress, 1988). According to Desmet and Hekkert (2007) the experience of meaning involves our ability to assign personality or other expressive characteristics and to assess the personal or symbolic significance of products". Giving more attention to the ways the meanings of artefacts are made socially is also reported by Theo van Leeuwen:

> in social semiotics, the focus has changed from the "sign" to the way people use semiotic "resources" both to produce communicative artefacts and events and to interpret them in the context of specific social situations and practices (van Leeuwen, 2005).

The focus of social semiotics put thus the emphasis firstly on the understanding the social context where the artefact's meaning is taking place; secondly on the wide range of modes in which meaning is made. "Mode" here is referred to something like the medium of the communicative act in question. Kress (2010) describes key modes as: image, writing, layout, music, gesture, speech, moving image, soundtrack, 3D objects.

The working definition of the word "meaning" that underlies the ideas put forth in this chapter exposition is that of "the sense of purpose that makes a person feel that his/her life is valuable" (e.g., "my car means a lot to me"). As also stated by Verganti (2011, p. 384), "meaning" represents "the profound psychological and cultural reasons people use a product". "Meaning" thus specifically refers to the "reason why" a designed artefact has value for the person in the operational and social context of its use.

5.3 The role of meaning in design practice

In the context of this exposition, the term "designed artefacts" refers to the intentionally designed objects that designers implement through a conscious, goal-oriented activity as opposed to "non-intentional design products" (Brandes *et al.*, 2013) which offer a broad spectrum of potential ulterior uses and types of repurposing driven by the consumers' motivation to use an object for a purpose other than that for which it was professionally intended.

In the context of designed artefacts (Douglas and Isherwood, 2021), designers typically follow a process that leads from the concept to the designed artefact as shown in Figure 5.1. When using products, however, the consumers participate in the opposite dynamics: they are looking for a particular concept that make sense to them, and on the way they will find artefacts that best match this concept.

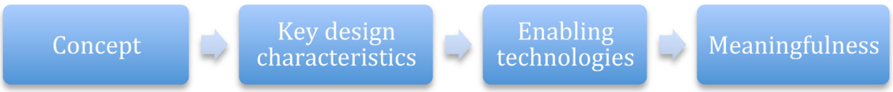


Fig. 5.1 – The designer's process: from conceptual ideas to meaningfulness of artefacts.

The idea that design is a manner for making sense of things (Krippendorff, 1989) is frequently discussed in professional circles, and debated in the design literature (Giacomin, 2017; Knudsen and Haase, 2018). For many practicing designers, the activity of design cannot be separated from the intended values and meanings of the artefact which is being designed. As noted by Krippendorff and Butter (2007) "it is a truism that we surround ourselves with objects that we are comfortable with and experience as meaningful. To design artefacts for use by others is to design them to be or to have the chance to become meaningful to these others – not merely in their designers' terms, but according to these others' own and often diverse conceptions".

Sociological research performed by Csikszentmihalyi and Rochberg-Halton (1981) have demonstrated how people construct patterns of meaning from the objects surrounding them. Most research to date provide an increasing evidence that many products and services evoke meanings and are symbolic stimuli for consumers (Levy, 1959; Holman, 1986, Csikszentmihalyi and Rochberg-Halton's, 1981; Park *et al.*, 1986; Fournier, 1991; Smith and Colgate, 2007; Almquist *et al.*, 2016) and an increasing interest in the process of how new meanings are constructed (Krippendorff and Butter, 2007; Verganti, 2009; Knudsen and Haase, 2018). All these previous studies suggest that although consumers seek functional benefits, underlying this there is a more profound significance composed of deeper meanings and instinctive ways of defining and shaping themselves and the world around them (Ravasi and Rindova, 2008). It further suggests that consumers are likely to select those alternatives in their decision-making that are more useful to their needs, that is, those that have personal relevance for achieving consumers' goals and values.

Adding further sophistication to the analysis of the meaning of designed artefacts, Krippendorff and Butter (2007) have suggested that "key to our conception of meaning is the recognition that humans create their own worlds and distinguish among their artefacts not in physical terms but according to what they mean to them, including how they enter the communications about them. Our concept of meaning involves a second-order understanding of how others come to understand and interact with our designs".

Boradkar (2010) suggested that "design's core mission is to fashion things so that we may have meaningful interactions with the world. Meanings are neither inherent properties of the things themselves, nor are they total fabrications of the human mind; they are suspended in the spaces between us and all that is around us. Meanings emerge and change continuously as people and things travel through their lives, constantly bumping into each other".

If something or a situation has a meaning for a person this in turn implicates that it would be necessary for a designer to understand what constitute the meanings in his/her life as he/she interprets it in order that the designer could anticipate much about his/her desires and behaviour (Siefkes, 2012).

Key to the conception of meaning by Krippendorff and Butter (2007) is to consider a human centered design approach for which "meanings cannot be separated from how people interact with the technologies that their culture creates and renders meaningful, with each other, and with how we – for example as designers or researchers – describe, conceptualize, and enact our conceptions of these meanings". It does seem reasonable to consider human centered design as ways of conceptualizing, designing and evaluating the symbolic meanings involved.

As described by Giacomin (2014), human centered design involves a hierarchy of considerations which places the value or meaning of the product (see Figure 5.2) at the crucial position at the top. Giacomin (2014) has described human centered design as the use of techniques which communicate, interact, empathize and stimulate the people involved, obtaining an understanding of their needs, desires and experiences which often transcends that which the people themselves actually realized.

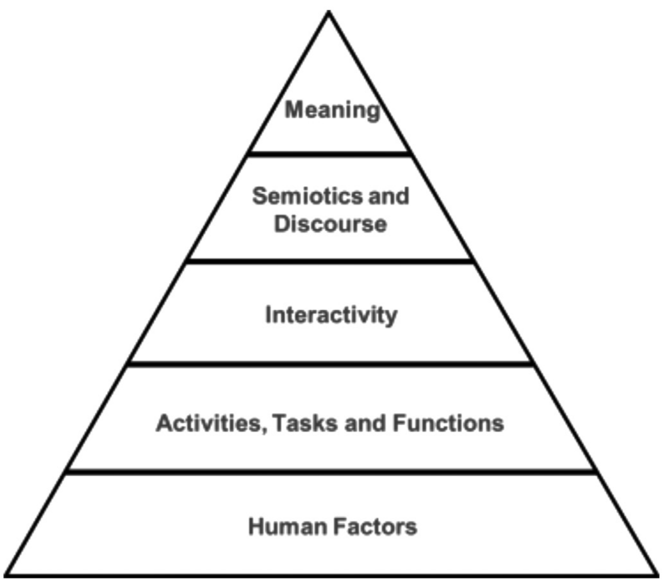


Fig. 5.2 – The human centered design pyramid (adapted from Giacomin 2014).

5.4 Meaning is a social, culture-based construct

The claims of sociologists regarding the constructed nature of meaning and its possible relativity to a given culture at a given point are supported by several studies of the meaning of artefacts. For example, Dewey (1934) has noted that "every person, as an individual, responds to different aspects of an artwork and experiences it differently; they also give it a meaning in different ways. In a similar vein, it can be assumed that every user experiences products differently". Extensive research by Csikszentmihalyi and Rochberg-Halton (1981) has shown that meaning can change substantially as a function of age, gender or other demographic descriptors of the people involved. Further, studies such as those of McCracken (1986) and Wallendorf and Arnould (1988) have shown that "the meaning associated with an artefact can change substantially as a function of the general cultural context in which the artefact is emerged".

Krippendorff and Butter (2007) suggest that "meaning is made collectively across society via culture using language. Desmet and Hekkert (2007) report that at the level of meaning, cognitive processes, like interpretation, memory retrieval, and associations enter into account to be able to recognize metaphors, and as such these cognitive processes depend onto individual and cultural differences.

Ravasi *et al.* (2012) noted that "the cultural perspective on value creation calls for a fundamental rethinking of the system of activities a firm engages in, e.g. how to involve designers in product development, how to manage new types of knowledge stocks...it also draws attention to the fact that the collective construction of the meanings that surround symbols and artefacts makes some of the cultural resources that producers use only partly under their control. Value creation and appropriation in such resources may depend more on skillful use and dynamic updating than on control and protection of intellectual property".

Battistella *et al.* (2012) have articulated the meaning as a stakeholder dependent property emphasizing the relativistic nature of the construct and highlights the need for commercially active designers to satisfy at least three potentially divergent world views (culturally constitute world, the company business model and the individual customer/stakeholder).

The fact that people respond socially to artefacts, it opens the door for AI technology to apply for a host of ethnographic research and persuasion dynamics that are collectively described as social influence: the type of influence that arises from social situations (Fogg, 2003).

5.5 Meaning is context-based, salient, collective, personalized, and dynamic

Meaning is perceived uniquely by individual customers; it is conditional or contextual (depending on the individual, situation, or product); it is relative (in comparison to known or imagined alternatives); and it is dynamic (changing within individuals over time and with age).

Decoding meaning from the product to the consumer is likely to be a more reinforcing and ongoing process for personalized objects than it will be for cultural objects (Fournier, 1991). The temporal quality exhibited by many objects with personalized sources of meaning (e.g., the favorite sweatshirt from college, the china that grandmother used, the ring received on wedding day) encourages on-going reflection by the owner/user, resulting in knowledge structures that are more dynamic and evolutionary. The forces that drive the meaning creation process may also differ by source of meaning (collective or personalized). Advertiser-constructed messages may be more dominant in the creation of meaning for objects with a cultural center while empathic responses and the generation of self-referent ties may be more likely to govern the creation of personalized meaning.

Because of these ties with the self, objects with personalized meaning center may enjoy higher levels of enduring involvement, greater salience and evocation potential, and stronger motivations for processing and elaboration.

5.6 Towards the need of a design for meaning framework

In the conceptualization of the meaning, Osgood (1952) suggests that meaning is composed by a "bundle of components" which include both objective and subjective elements, such as experiences, feelings and images. This allows to visualize these components as the basic structural elements of the construct meaning. Friedmann and Lessig, 1986 report that the meaning that products elicit in the consumer's mind is shown to be a function of the bundle of attributes found in the product, the consumer's perceptual mode, and the context in which the perceptual process takes place. More specifically, Friedmann (1986) and Hirschman (1980) have characterized the meaning of an artefact in three dimensions: tangibility, emotionality and commonality.

- Tangibility concerns whether the attribute at the basis of meaning is tangible, in the sense that is objective and verifiable through the senses, instead of being subjective. It reflects if the meaning is located in the object itself or in the mind of the user.
- Emotionality concerns the emotional responses to products and it can range from low to high intensity, going from simple affective reactions to true emotional experiences such as enjoyment and excitement.
- Commonality identifies the degree to which meaning is shared by members of the culture, or if it has a more individualized and personal character. It is linked to the source that is more responsible of the assignment of meaning to the object. For example, it is a cultural source if the meaning is assigned by advertising and shared media opinion. It is a personal source if meaning is created by the user through time, experience and interaction, such as an object received as a gift from a dear friend, your most comfortable pair of shoes, or your favorite toy when you were a child.

Holman (1986) proposes a typology of products that captures the exact nature and character of the emotional experiences beyond the utilitarian/experiential dichotomy to present five categories of products: background props, mediators, enhancers, self-expression and emotional objects.

Fournier (1991) acknowledges that all products contain degrees of both hedonic and utilitarian elements, which allows for the placement of objects along a hedonic/utilitarian continuum. Eight categories of objects are defined based on the nature of the meaning perceived by the consumer: objects of utility, action, appreciation, transition, childhood, ritual enhancement, personal identity, and position and role. Fournier (1991) has further categorized the meaning of objects using the following dimensions: utilitarian, symbolic, emotional, shared, personalized.

In terms of customer values, Park *et al.* (1986) has defined three brand value categories: functional, experiential, symbolic. Smith and Colgate (2007) have defined four categories of value: functional/instrumental, experiential/hedonic, symbolic/expressive and cost/sacrifice.

Creusen and Schoormans (2014) distinguished six roles of product appearance for consumers based on a literature review: aesthetic, symbolic, functional, ergonomic product information, attention drawing and categorization. In a large qualitative study (N = 142) Creusen and Schoormans (2014) tested whether these roles indeed existed in consumers' process of product choice, and whether they were sufficient to describe the way in which product appearance plays a role for consumers. In addition, they gained qualitative insight into these roles. The aesthetic and symbolic appearance roles were far more salient to consumers, and the appearance influenced perceived ergonomic value for one-third of the subjects. The functional role of the appearance was mentioned less by the customers.

Giacomin (2017) has developed a design for meaning framework which was further developed by Ajovalasit and Giacomin (2019) who capture the main forms of meaning of designed artefacts into three primary categories of meaning covering a spectrum from the purely instrumental to the purely symbolic: function, ritual and myth. Giacomin (2017) further suggests that "while the choice of other semantics is of course possible, the categories of function, ritual and myth align closely with three of the four categories of value defined by Smith and Colgate (2007) of functional/instrumental, experiential/hedonic, symbolic/expressive and cost/sacrifice. The categories of function, ritual and myth also align closely with the three brand value types suggested by Park *et al.* (1986) of functional, experiential and symbolic".

Vitali *et al.* (2019) have recently proposed a toolkit that consider three kinds of meaning that are relevant for smart connected products: the meaningful identity of the object as product category, the meaning of the product in relation to its shape and functionality, and in relation to a phygital ecosystem.

Mekler and Hornbæk (2019) have defined a framework of experience of meaning in human-computer interaction which outlines five distinct senses of the experience of meaning: connectedness, purpose, coherence, resonance, and significance. They report that their "framework focuses on the moment-to-moment experience of meaning".

Table 5.1 summarizes the different categories of meaning as found in various frameworks from the literature.

Authors (year)	Hirschman (1980)	Holbrook and Hirschman (1982)	Friedmann (1986)	Friedmann and Lessig (1986)	Park et al. (1986)	Fournier (1991)	Diller et al. (2005)	Smith and Colgate (2007)	Siefkes (2012)	Creusen and Schoormans (2014)	Giacomin (2017)	Vitali et al. (2019)	Mekler and Hornbæk (2019)
Meaning category	tangibility	-	tangibility	heuristics	functional	utilitarian	accomplishment	functional/ instrumental	function or frame	functional	function	shape and functionality	purpose
	emotionality	aesthetic	emotionality	fun	experiential	shared	beauty	experiential/ hedonic	style	ergonomic product information	ritual	meaningful identity	connectedness
	communality	hedonic	communality	elation	symbolic	personalised	creation	symbolic/ expressive	Iconicity or metaphor	aesthetic	myth	phygital ecosystem	coherence
		symbolic		hedonic pleasure		emotional	community	cost/sacrifice	individual experiences	symbolic			resonance
						symbolic	duty		cultural allusions	attention drawing			significance
							enlightenment		connection to social groups	categorization			
							freedom		specific contexts				
							harmony						
							justice						
							oneness						
							redemption						
							security						
							truth						
							validation						
							wonder						

Table 5.1 – Categories of meaning of artefacts found in the literature

The work done by the different researchers, as presented in Table 5.1, to attempt framing and categorizing diverse kind of meaning which the artefact is anticipated to provide or facilitate for the customer lead to a series of reflections and implications regarding the need of a framework of design for meaning.

Firstly, most research to date tends to address the topic of product meaning from the product attribute perspective, whereby meaning is tied only to the physical, observable characteristics of the product (Gutman, 1982; Kleine and Kernan, 1988; Reynolds and Olson, 2001). Moreover, current research seems to consider forms of meaning to be fixed taxonomic categories (Park *et al.*, 1986; Fournier, 1991; Smith and Colgate, 2007; Almquist *et al.*, 2016). However, given the complexities and divergences in forming the construct of meaning by the individual consumer's perceptual mode (Dewey, 1934), it can be argued that not all artefacts or actions which hold meaning for people are likely to fall into easily predetermined categories.

In addition, most of the research about meaning performed to date seems to consider the taxonomy to be self-evident, or at least easily identified. Given the complexities of human language and the infinite variations in social constructs, it is not currently obvious what words should be used and what ideas should be raised when working with people to establish the taxonomy and parameters of meaning. Existing research does not suggest how to talk about meaning with people, or how to ensure that the approach, language and concepts adopted do not lead to irreparably biased and misleading conclusions.

If commercially active designers are therefore expected to clarify, decide upon and communicate the meaning which an artefact is anticipated to provide or facilitate for the consumer (Giacomin, 2017), it would be then necessary for the designer to acquire an understanding of how people come to understand the meaning attached to a designed artefact in their own terms and for their own reasons. This is what Krippendorff (2006) refers to as second-order understanding.

5.7 The framework of "design for meaning"

To assist to the limitations of current research approaches, Giacomin (2017) has developed a "design for meaning" framework which provides basic considerations which should be considered if meaning is a fundamental characteristic of the product, system or service which is being designed. The framework, shown in Figure 5.3, is based on a systematic review of several key semantics involved in design for meaning literature. The framework of design for meaning is subdivided into two sections in relation to the fundamental consideration of whether the artefact should adhere to an existing technological or societal stereotype or, instead, whether there is the opportunity or the need to define a new stereotype due to technological or societal change. The considerations of ideology, meaning, function, ritual, myth, meaningfication and metaphor constitute the basic checklist of questions to ask and clarifications to achieve.

The concept of "design for meaning" suggests that the three categories of pre-existing meaning of function, ritual and myth can provide a bridge between the global meaning of an artefact and the specific metaphor which is deployed by the designer.

The category of "function" is meant to reflect all those situations in which a physical or informatic use is acting as the focus of attention, with less attention being paid to the psychological or sociological considerations. The category of "ritual" is meant to reflect all those situations in which the meaning of the artefact is closely related to action of a symbolic nature (Rook, 1985). The category of "myth" is meant to reflect all those situations in which the meaning of the artefact is mainly symbolic, thus not necessarily requiring dedicated externally visible activity on the part of the consumer (Barthes, 1973).

Further, each type of meaning is implicitly a word which describes a complete spatial-temporal event. A function or a ritual is not as such if it is not fully performed, and a myth is not a myth if most of the story is not fully known to the person.

When a designer identifies an opportunity, which interconnects several previously unrelated technological and cultural codes, and articulates one or more product, system or service concepts which address the opportunity, the process can be described as one of "meaningfication" which is defined as:

> The use of data, design ethnography, real fictions and co-creation for the purpose of designing artefacts based on new meanings which emerge from the interconnection of evolving patterns of technology, experience, personal identity, societal identity, value assignation and consumption (Giacomin, 2017).

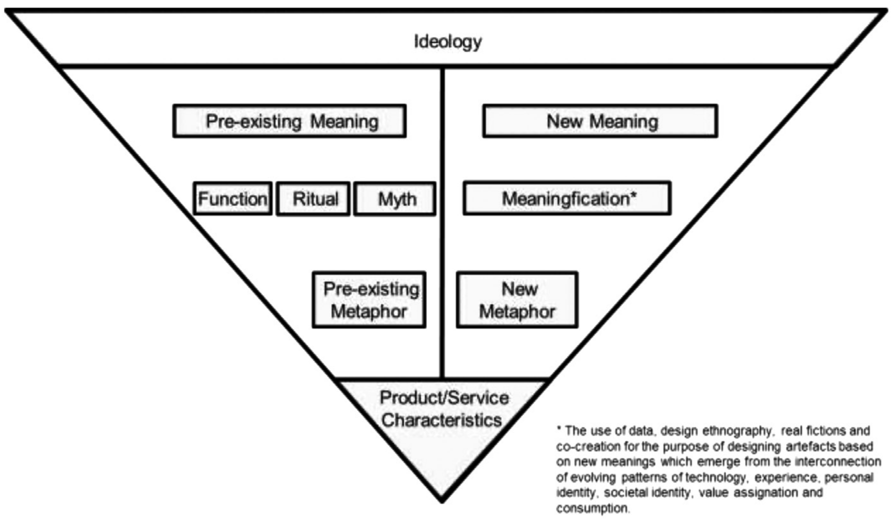


Fig. 5.3 – Framework of design for meaning (adapted from Giacomin, 2017).

The usefulness of the framework has been further explored by Ajovalasit and Giacomin (2019) demonstrating that three categories of pre-existing meaning, "function", "ritual" and "myth", are commonly encountered in practice, either occurring individually or be co-present to some degree covering a spectrum from the purely instrumental to the purely symbolic. While it is possible to find designed artefacts of many kinds defined solely in terms of function related to the quality of appropriateness in use, to the performance of specific tasks or to effectively execute them, there are situation for which a meaning stems from a personal evocation of relationship and ritual or that is mainly symbolic which is intrinsic and not dependent upon any specific affordance (Ajovalasit and Giacomin, 2019).

Further research is needed to identify the elements that constitute and define the semantic dimensions and the microstructure of the three categories of meaning to assist the designers in the construction of meaning and the required words and ideas which are adequate to capture the main forms of meaning.

5.8 Meaningfication of AI-infused artefacts

The growing uses of AI technologies interfacing with people and their advances of the AI-infused artefacts into the societal domain and everyday activities (Follett, 2015; Amershi *et al.*, 2019) lead to the consideration that a business opportunity can be achieved by exploiting a new technological capability or a new cultural code which Giacomin (2017) defines it as "meaningfication" in which the interconnection of previously unrelated technological and societal codes lead to new ways of thinking by the consumer. Fundamental to the concept of meaningfication is that the "consumer is implicitly invited to think differently about the opportunities and value propositions which are on offer" (Giacomin, 2017). The design implication for AI-infused artefacts is therefore that innovative aspects of such artefacts should be framed with new meanings, or it should be clarified if the meanings of the AI-infused artefacts still align with the pre-existing meanings linked to previous people's experiences. As observed by Giacomin (2017), "in time the new business opportunities can develop into societally recognized functions, rituals or myths, but start off initially as hybrids involving previously unrelated technological and societal codes".

AI-infused artefacts can augment people natural interactions that are already happening in the world, recoding them as data or interpreting them as input and taking action (King and Chang, 2016) leading to new meanings for the consumer. Current research deploying the design for meaning approach for AI-infused artefacts has outlined why AI-infused artefacts could be considered as a result of the concept of meaningfication. Vitali *et al.* (2019) for example suggested that "a successful example of design for meaning approach is represented by the Nest learning thermostat[1] (…), that can detect whether a person has left the house and turn down the temperature. They suggested that "through technology the Nest was able to shift the identity/meaning of the thermostat from a static, passive, and hard to personalize element of the house, towards an active and reactive element, that automatically answers to user's presence and habits. This shift from the pre-existing

1. TheGoogle Nest learning thermostat: https://store.google.com/gb/product/nest_learning_thermostat_3rd_gen?hl=en-GB.

values led to a new vision for the market of thermostats".

Another example concerns the creation of new visions in the market of Smart TVs. Samsung for example reflects on the role in the room of large, wall-mounted TVs, trying to suggest new meanings. Vitali *et al.* (2019) reported that "Samsung proposes 'The Frame' a TV that when switched off becomes a personal piece of art displaying selected high-resolution images from a curated art gallery. In this way, the TV acquires meaning in the room, even when is not switched on. Another innovation is the Samsung Ambient Mode, which lets the TV screen disappear, simulating the texture of the wall on which it is mounted. In these examples the new meanings and visions redefine the architectural role of the TV, trying to anticipate future user wishes". Vitali *et al.*, (2019) suggested that AI-infused artefacts should have a strong identity that differentiates them from, or relates them to, the pre-existing meanings of their unconnected, "less smart" counterparts. The meaningful identity of the artefact should be considered as a product category.

Other commercial artefacts which are the result of meaningfication include numerous products from the wearables sector (Follett, 2015) which combine the initially unrelated topics of "micro-computer", "fashion accessory" and "health awareness". Recent examples of smartshoe, fitness tracker and hearable devices are artefacts of meaningfication which are currently developing into established stereotypes. Bosch Sensortec (2022) for example has developed a revolutionary self-learning motion sensor that adds AI to portable devices: the BHI260AP self-learning AI sensor[2] which enables personalized solutions for every user.

By adding self-learning AI capabilities, Bosch Sensortec's new movement sensor changes how users interact with their fitness devices from a mere one-way approach and activity tracker to an interactive way of training. For fitness trackers the shift in demand to home-based "trainerless" solutions offers substantial potential to benefit users by accurately informing and assisting them with their solo exercise programs (Kenez *et al.*, 2021). The self-learning AI leads to new ways of thinking for fitness tracking making exercising more automated, personalized, and upgradeable. This makes exercising more motivating

2. The Bosch BHI260AP self-learning AI sensor: www.bosch-sensortec.com/products/.

and rewarding – whatever a user's level of fitness or experience (Kenez *et al.*, 2021).

5.9 Conclusions

AI-infused artefacts can be regarded as consumer goods which have a significance that goes beyond their utilitarian character and commercial value. This significance rests largely in their ability to carry and communicate meaning for their users.

The fact that people respond socially to AI-infused artefacts has significant implications for intrinsic motivation and persuasion. The design of AI-infused artefact can then leverage the principles of social influence to motivate, persuade and generate meaning (Fogg, 2003).

Common characteristic of AI-infused artefacts is that unlike other artefacts, they have the possibility to persuade and influence people' attitude and behaviors, or to adapt and learn over time. They may also react differently according to lightning conditions, ambient noise, human accents, and other contexts. Their effect and unpredictability if not designed well may confuse users, dwindles their trust, and may ultimately lead to abandonment of the artefacts for its lack of meaningfulness.

For such AI-infused artefacts, a design for meaning framework based on human centered design approach has been proposed here to help the designer to careful articulate the intended meaning by either referencing an existing meaning or articulating a new meaning through the process of meaningfication for which "customers are implicitly invited to think differently about the opportunities and value propositions which are on offer" (Giacomin, 2017).

The "design for meaning" framework provides basic considerations and checklist of questions to ask and clarifications to achieve which should be considered if meaning is a fundamental characteristic of the product, system or service which is being designed. The concept of "meaningification" is not prescriptive in terms of the methodologies to adopt, but it leaves ample room for the deployment of approaches such as real fictions (Dunne, 2008), projective techniques (Soley and Smith, 2008), collaborative organizations (Manzini, 2015) and crowdsourcing (Brabham, 2013). These and other ethnographic approaches typically used for social science and business research (Khoo-Lattimore *et al.*, 2009; Porr *et al.*, 2011) would help to uncover feelings, beliefs, attitudes and motivations that consumers otherwise find difficult to articulate. They can provide a greater depth of understanding into what people truly think and feel about an artefact.

The reflections made in this chapter suggest that the consideration of target values and meanings that surround artefacts are important as a strategy towards defining the relational role of AI-infused product in the lives of its owners for which artefacts are not simply functional tools, but are also relational mediators which shape the long term aims, objectives and behaviors of an individual or of a group. The meaningfulness factor and whether artefacts are there to motivate people to set and achieve their own goals or develop better habits make AI-infused artefacts one of the most promising frontiers in the use of AI-technology interfacing with people.

References

Ajovalasit, M., Giacomin, J. (2019). "Meaning of artefacts: interpretations can differ between designers and consumers". In *Conference Proceedings of the Academy for Design Innovation Management (ADIM2019)*, 2(1), 1178-1188. https://doi.org/10.33114/adim.2019.02.266.

Almquist, E., Senior, J., and Bloch N. (2016). "The Elements of Value Measuring – and delivering – what consumers really want". *Harvard Business Review*, pp. 46-53, September.

Amershi, S., Weld, D.S., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S.T., Bennett, P.N., Quinn K.I., Teevan, J., Kikin-Gil, R., and Horvitz, E. (2019). "Guidelines for Human-AI Interaction". In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. doi:10.1145/3290605.3300233.

Barthes, R. (1973). *Mythologies*, London: Paladin.

Battistella, C., Biotto, G., and DeToni, A.F. (2012). "From Design Driven Innovation To Meaning Strategy". *Management Decision*, 50(4), 718-743.

Baudrillard, J. (1968). *The system of objects*. New York, NY: Verso Books.

Blosch and Fenn (2018). *Understanding Gartner's Hype Cycles*. Available at: www.gartner.com/en/documents/3887767/understanding-gartner-s-hype-cycles (Accessed on 18 January 2022).

Bosch Sensortec GmbH (2022). www.bosch-sensortec.com/news/ai-revolutionizes-fitness-tracking.html (Accessed on 20th April 2022).

Bower, J.L., and Christensen, C.M. (1995). “Disruptive Technologies: catching the wave”. *Harvard Business Review*, January-February, pp. 506-520.
Brabham, D.C. (2013). *Crowdsourcing*. Cambridge, MA: The MIT Press.
Brandes, U., Stich, S., and Wender, M. (2013). *Design by Use: The Everyday Metamorphosis of Things*. Boston, MA: Birkhäuser. https://doi.org/10.1515/9783034609128.
Clayton, C. (1997). *The Innovator’s Dilemma: when new technologies cause great firms to fail*. Boston, MA: Harvard Business School Print.
Creusen, M.E.H., Schoormans, J.P.L. (2014). “The Different Roles of Product Appearance in Consumer Choice”. *Journal of Product Innovation Management*, 22(1), 63-81.
Csikszentmihalyi, M., and Rochberg-Halton, M. (1981). *The Meaning of Things*. Boston, MA: Cambridge University Press.
Dahlin, K.B., and Behrens, D.M. (2005). “When Is An Invention Really Radical?: defining and measuring technological radicalness”. *Research Policy*, 34(5), 717-737.
Desmet, P.M.A., and Hekkert, P. (2007). “Framework of product experience”. *International Journal of Design*, 1(1), 57-66.
Dewey, J. (1934). *Art as Experience*. New York, NY: Capricorn Books.
Diller, S., Shedroff, N., and Rhea, D. (2005). *Making Meaning: how successful businesses deliver meaningful customer experiences*. Berkeley, CA: New Riders Publishing.
Douglas, M., and Isherwood, B. (2021). *The World of Goods*. London: Routledge.
Dunne, A. (2008). *Hertzian Tales: Electronic Products, Aesthetic Experience, and Critical Design*. Cambridge, MA: The MIT Press.
Fogg, B.J. (2003). *Persuasive Technology. Using Computers to Change What We Think and Do.* 1st edn. Burlington, MA: Morgan Kaufmann.
Follett, J. (2015). *Designing for Emerging Technologies.* 2nd edn. Sebastopol, CA: O’Reilly Media, Inc.
Fournier, S. (1991). “Meaning-Based Framework For the Study of Consumer-Object Relations”. *Advances in Consumer Research*, 18, 736-742.
Franklin, D. Roosevelt (1940). *Address at University of Pennsylvania.* Online by Gerhard Peters and John T. Woolley, The American Presidency Project Available at: www.presidency.ucsb.edu/node/210464 (Accessed: 12 January 2022).
Friedmann, R. (1986). “Psychological Meaning of Products: Identification and Marketing Applications”. *Psychology and Marketing*, 3 (Spring), 1-15.
Friedmann, R., and Lessig, V.P. (1986). “A Framework of Psychological Meaning of Products”. In Lutz, R.J and Provo, UT (eds.). *North American*

Advances in Consumer Research, vol. 13, *Association for Consumer Research*, pp. 338-34.
Giacomin, J. (2014). "What is human centred design?". *The Design Journal*, 17(4), 606-623.
Giacomin, J. (2017). "What is Design For Meaning". *Journal of Design, Business & Society*, 3(2), 167-190.
Google Nest Learning Thermostat. Google Nest Labs. https://store.google.com/product/nest_learning_thermostat_3rd_gen?hl=it (Accessed on 26th April 2022).
Gutman, J. (1982). "A means-end chain model based on consumer categorization processes". *Journal of Marketing*, 46(2), 60-72.
Gutman, J. (1990). "Adding meaning to values by directly assessing value-benefit relationships". *Journal of Business Research*, 20(2), 153-160.
Hirschman, E. (1980). "Attributes of Attributes and Layers of Meaning". *Advances in Consumer Research*, vol. 7, Jerry Olson (ed.), Ann Arbor, MI: Association for Consumer Research, 7-11.
Hodge, R., and Kress, G. (1988). *Social Semiotics*. Cambridge, MA: Polity Press.
Holbrook, M.B., and Hirschman, E. (1982). "The Experiential Aspects of Consumption: Consumer Fantasies, Feelings and Fun". *Journal of Consumer Research*, 9, 132-140.
Holman, R. (1986). "Advertising and Emotionality". In Peterson, R., Hoyer, W. and Wilson, W. (eds.). *The Role of Affect in Consumer Behavior*. Lexington: D.C. Heath and Co, pp. 119-140.
Holt, D., and Cameron, D. (2010). *Cultural strategy: using innovative ideologies to build breakthrough brands*. Oxford, UK: Oxford University Press.
IBM Cloud Education (2010). *What is Artificial Intelligence (AI)?*. Available at: www.ibm.com/cloud/learn/what-is-artificial-intelligence (Accessed on 16 December 2021).
Kaptelinin, V., and Nardi, B.A. (2009). *Acting with Technology. Activity Theory and Interaction Design*. Cambridge, London: The MIT Press.
Kenez, N., Shubhabrata, P., and Gandhi, K. (2021). "Me, myself and AI – How the new self-learning AI sensor personalizes your home workout". *Bosch Sensortec White Paper*, January 2021 www.bosch-sensortec.com/news/ai-revolutionizes-fitness-tracking.html.
Khoo-Lattimore, C., Thyne M., Robertson, K. (2009). "The ZMET method: using projective technique to understand consumer home choice". *The Marketing Review*, 9(2), 139-154. doi: 10.1362/146934709X442674.
King, S., and Chang, K. (2016). *Understanding Industrial Design: Principles for UX and Interaction Design*. 1st edn. Sebastopol, CA: O'Reilly Media, Inc.

Kleine, R.E., and Kernan, J.B. (1988). “Measuring the Meaning of Consumption Objects: an Empirical Investigation”. In Micheal, J. (ed.). *NA – Advances in Consumer Research*, vol. 15, Houston, Provo, UT: Association for Consumer Research, pp. 498-504.

Knudsen, L.S., and Haase, L.M., (2018). “The construction of meaning in design-driven projects: a paradox initiated process”. *International Journal of Design Creativity and Innovation*. doi:10.1080/21650349.2018.1501281.

Kress, G. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London, UK: Routledge.

Krippendorff, K. (1989). “On the essential contents of artifacts or on the proposition that design is making sense (of things)”. *Design Issues*, 5(2), Spring.

Krippendorff, K. (2006). *The Semantic Turn: a new foundation for design*. Florida, US: Taylor&Francis.

Krippendorff, K., and Butter, R. (2007). “Semantics: meanings and contexts of artifacts”. In Schifferstein, H.N.J., and Hekkert, P. (eds.). *Product Experience*. Amsterdam, NL: Elsevier.

Kuniavsky, M. (2010). *Smart Things, Ubiquitos Computing User Experience Design*, 1st edn. Burlington, MA: Morgan Kaufmann.

Manzini, E. (2015). *Design, When Everybody Designs: an introduction to design for social innovaiton*. Cambridge, MA: The MIT Press.

McCracken, G. (1986). “Culture and Consumption: A Theoretical Account of the Structure and Movement of Cultural Meaning Of Consumer Goods”. *Journal of Consumer Research*, 13 (June), 71-84.

Mekler, E.D., Hornbæk, K. (2019). “A Framework for the Experience of Meaning in Human-Computer Interaction”. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI ’19*. Association for Computing Machinery, Glasgow, Scotland Uk, pp. 1-15. https://doi.org/10.1145/3290605.3300455.

Norman, D.A., and Verganti, R. (2014). “Incremental and radical innovation: design research versus technology and meaning change”. *Design Issues*, 30(1), 78-96.

Osgood, C.E. (1952). “The nature and measurement of meaning”. *Psychological Bulletin*, 49(3), 197-237. https://doi.org/10.1037/h0055737.

Park, C.W., Jawarski, B.J., and MacInnis, D.J. (1986). ‘Strategic Brand Concept-Image Management”. *Journal of Marketing*, 50, October, pp. 135-145.

Porr, C., Mayan, M., Graffigna, G., Wall, S., and Vieira, E.R. (2011). “The Evocative Power of Projective Techniques for the Elicitation of Meaning”. *International Journal of Qualitative Methods*, March, 30-41. https://doi.org/10.1177/160940691101000103.

Ravasi, D., and Rindova, V. (2008), "Symbolic Value Creation". In Barry, D., and Hansen, H. (eds.). *The Sage Handbook of New Approaches to Organization Studies*. London, UK: Sage.

Ravasi, D., Rindova, V., and Dalpiaz, E. (2012), "The cultural side of value creation". *Strategic Organization*, 10(3), 231-239.

Reynolds, T.J., and Olson, J.C. (2001). *Understanding Consumer Decision Making: The Means-end Approach To Marketing and Advertising Strategy*, 1st edn., US: Psychology Press. https://doi.org/10.4324/9781410600844.

Richins, M.L. (1994). "Valuing Things: the public and private meanings of possessions". *Journal of Consumer Research*, 21, 504-521.

Rook, D.W. (1985). "The Ritual Dimension of Consumer Behavior". *The Journal of Consumer Research*, 12(3), 251-264.

Samsung Ambient Mode www.samsung.com/sg/support/tv-audio-video/what-is-ambient-mode-and-how-to-use-it/ (Accessed on 26th April 2022).

Samsung The Frame www.samsung.com/it/lifestyle-tvs/the-frame/highlights/ (Accessed on 26th April 2022).

Sciuto, A., Saini, A, Forlizzi, J., and Hong, J.I. (2018). "Hey Alexa, What's Up?": A Mixed-Methods Studies of In-Home Conversational Agent Usage". In *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 857-868). ACM.

Siefkes, M. (2012). "The Semantics Of Artefacts: how we give meaning to the things we produce and use". Themenheft zu Image 16, *Semiotik*, pp. 67-102.

Smith, J.B., and Colgate, M. (2007), "Customer value creation: a practical framework". *Journal of Marketing Theory and Practice*, 15(1), 7-23.

Soley, L.C., and Smith, A.L. (2008). *Projective Techniques For Social Science And Business Research*. 1st edn, US: The Southshore Press.

Spallazzo, D., Sciannamè, M. Ceconello, M. (2019). "The Domestic Shape of AI. A Reflection on Virtual Assistants". *Desform 2019, Beyond Intelligence*. Boston, MA: MIT Design Lab.

Van Leeuwen, T. (2005). *Introducing Social Semiotics*. London, UK: Routledge.

Verganti, R. (2009), *Design-Driven Innovation: changing the rules of competition by radically innovating what things mean*. Boston, MA: Harvard Business Press.

Verganti, R. (2011). "Radical design and technology epiphanies: A new focus for research on design management". *Journal of Product Innovation Management*, 28(3), 384-388.

Vitali, I., Arquilla, V., Rifino, I. (2019). *Design for Meaning of Smart Connected Products*, ADIM2019 Conference 19-21 June 2019, London.

Wallendorf, M., and Arnould, E.J. (1988). "My Favorite Things: a cross-cultural inquiry into object attachment, possessiveness, and social linkage". *Journal of Consumer Research*, 14(4), 531-547.

Weiser, M. (1991). "The Computer for the 21st Century". *Scientific American*, 3(265), 94-104.

Williamson, J.E. (1978). *Decoding Advertisements: Ideology and Meaning in Advertising*. London, UK: Marion Boyars.

Conclusions

Davide Spallazzo, Martina Sciannamè
Department of Design, Politecnico di Milano

Notoriously, the field of AI has been characterized by multiple interpretations and perspectives, fueling diverse research strands and approaches that marked the evolution of such an ambitious and disruptive technology. At the same time, however, this results in definitions that are not universally agreed upon, already within computer scientists' circles. The ambiguity in understanding is inevitably amplified when AI theories move from algorithm experimentations in academic and research contexts to spread in the broader and more complex socio-technical reality, with implications ranging from everyday pragmatism to the more abstract making sense of what AI-infused artifacts are and represent.

The collection of essays represents an initial exploration of how the design discipline, with its mediating role between people and their artificial world, can face the embodiment of AI capabilities within industrial products.

Of course, the book cannot be exhaustive: the design issues surrounding AI-infused products are manifold, and there can be many standpoints from which to approach the subject. Nevertheless, the essays clarify that the design discipline is looking for an approach to deal with AI and the complexity of embedding intelligence into everyday products.

Machine learning, in particular, introduces abilities beyond the traditional interaction model "input-elaboration-output," where the same input always has an identical output. It paves the way for more flexible, reactive, adaptive, and even proactive artifacts, challenging our perception of objects as mere instruments and suggesting their intelligent entity. This disorienting paradigm shift found designers – practitioners and researchers – unprepared.

The first chapter photographs this disorientation, evident in industrial products that, driven by the push of the technological hype, have forgotten decades of usability studies, with several leaks in terms of smooth basic interaction.

Uncertainty and confusion – both for users and designers – are even increased when the ambivalence of machines mimicking human behaviors, rooted in the origins of the AI discipline, comes into play. The introduction of objects that seemingly have or emulate human capabilities profoundly affect interaction modalities, bringing to light relevant controversies. For instance, today, talking to objects is an expected feature, but it is not without dark sides. As suggested in chapter four, conversational interfaces are an open field of experimentation for designers. However, their implementation often overwhelms traditional interfaces – at least in the imagination of users – but they are still not reliable enough to allow a smooth experience.

The essays portray AI-infused products as devices with a high potential for end-users but underexploited. At least in the materialization of AI capabilities, the fascination for this groundbreaking technology has not allowed going further than employing CUIs to control connected devices and services. Accordingly, they may play the role of mere gadgets that do not go beyond their toy phase.

From these considerations, an urge for mapping and systematizing the characteristics of AI-infused artifacts emerges as a way to facilitate designers' understanding of the current panorama. The approach explored in the volume relies on UX assessment as a pillar for the design discipline to appraise and make sense of things. Chapters two and three show the lack of methods to analyze the complexity of such products and present a first attempt to introduce new UX dimensions specific to this category of appliances.

Traditional qualities of the UX experience (e.g., usability, aesthetics) are losing ground to more sophisticated dimensions (e.g., trustworthiness, intelligence), which simultaneously indicate people's attraction to the new possibilities unlocked by the thriving technology but also their rising awareness of the complexity and lack of transparency of such products. For instance, end-users are beginning to question the ethical implications of employing massive data to provide a service, at least from a UX standpoint. Furthermore, their perception and expectations of the intelligence of these devices affect the quality of the interaction as much as the capabilities of AI systems to understand users' commands and needs.

Indeed, questioning the meaning and meaningfulness of AI-infused products in our daily life (see chapter 5) entails opening a debate on the very core of these devices and their perceived utility and quality.

After all, assessing the user experience of these innovative products becomes an urgent need not only for providing end-users with a better UX but also to guide designers and companies in developing more helpful, satisfactory, and trustworthy products.

In line with this reasoning, developing a UX evaluation method specific to AI-infused systems is a consistent and necessary follow-up. In fact, it is the core output of the Meet-AI research project, which created AIXE (AI user eXperience Evaluation), a statistically validated questionnaire for assessing the UX of AI-infused systems.

However, evaluating is just one way of coping with poor UX in such systems and understanding their limits in terms of UX does not automatically entail their solutions.

For that, more radical actions should be implemented. For example, there is a pressing need to form a new generation of designers that master (at least) basic knowledge of AI and ML. They may help envision novel solutions that fully exploit the technological potential and include human needs, desires, relations, and rights in a broader, systematic perspective.

Authors

Marco Ajovalasit. Associate Professor in Human-Centered Design in the Design Department at the Politecnico di Milano, Italy, where he teaches modules in User Experience Design, and Digital Interaction Design. His research expertise is in the field of design for meaning and human-centred design methods leading to products, systems and services which are physically, perceptually, cognitively and emotionally intuitive to their users. Research focuses on the "design for meaning" approach defining human-centered design methods for designing products, systems and services based on new meanings for the consumers, including the use of data, designing ethnography, real fiction and co-creation. A particular emphasis of the research is on the development of a toolkit for designers for organizing the consideration of the intended meaning of the designed artefacts. Focus is placed on the thinking process, dialogue and use of semantics consumers typically associate to the understanding of the designed artefact. He has been the Principal Investigator (PI) (2013-2016) of a £3.2M FP7 EU Grant funded project "Light.Touch.Matters: Design-driven development of touch sensitive luminous flexible plastics for applications in care & well-being", which put together a multi-disciplinary team of designers and material scientists and industry leaders involving four Universities and thirteen SMEs designers from 9 different countries. Within the EU project he has developed design guidelines for meaningful interactions for the consideration of the effects of product properties on the perceptual experiences which occur at a product interface.

Venanzio Arquilla. Designer, Co-Founder of the StartUp Narvalo, Spin-Off of the Politecnico di Milano, is an Associate Professor at the Design Department, Politecnico di Milano. He is the Coordinator of the BSc in Product Design and the MSc in Integrated Product Design at the POLIMI

Design School. His research activities deal with user experience, strategic and service design, smart and connected products. He is the founder and coordinator of the Experience Design Academy that comprise the UX Design and the APP Design and Development higher education courses at POLI.design. He is the co-director of the Master of User Experience Psychology held by Università Cattolica del Sacro Cuore and Politecnico di Milano/POLI.design.

Mauro Ceconello. Architect and Associate Professor, Product Design, School of Design, Politecnico di Milan. He focuses his research activity on tools and apps to enhance cultural heritage and tourism using mobile technology, serious games, and interactive systems. He held courses about augmented reality and mobile experience. He is a scientific coordinator of many research projects concerning the valorization of culture through digital technologies and interaction tools.

Alice Paracolli. Ph.D. student in Design at Politecnico di Milano. She is specialised in User Experience, and her research focuses on Sustainable UX applied to AI-infused Objects. Ubiquitous computing elicits an endless range of internet-enabled devices, offering the potential to the user to be more diligent in energy use, except that it creates an ever-growing web of data-consuming objects that stay on forever. The Ph.D. aims to set design guidelines for sustainable, connected objects, proposing a framework to build continuously incremental sustainable innovation in technology.

Martina Sciannamè. Ph.D. student in Design at Politecnico di Milano. She investigates how AI and ML can be meaningfully included in Design education with a multidisciplinary approach. Her aim is to produce theoretical knowledge and practical tools to enable designers to exploit the opportunities offered by this technology and to communicate with ML experts, being aware of its actual capabilities and limitations. Her hope is that the intersection of Design and ML, within the frame of Ethics, will finally steer the development of ML systems towards real-world problems and responsible innovation for the flourishing of humans.

Davide Spallazzo. Ph.D. in Design and Associate Professor at the Department of Design of Politecnico di Milano. Active in Interaction Design and Human-Computer Interaction, his main research interest is the human-centred approach to digital innovation and meaning-making, applied to diverse fields of inquiry. Over the years, he coordinated and took part in several national and international research projects concerning digital technologies in the Cultural Heritage field, serious gaming, and AI. He is now coordinating the

Meet-AI research project, aimed at creating a new UX assessment method for AI-infused products. He is deputy coordinator of the Ph.D. program in Design and secretary of the MSc in Digital & Interaction Design.

Ilaria Vitali. Product designer and Ph.D. in Design, with honors, at Politecnico di Milano with a thesis titled "Conversation on conversational smart products", graduated with a Master's degree in Product Design for Innovation and a dual honors degree from Alta Scuola Politecnica. Her research focuses on smart connected products and devices with conversational interfaces and explores how to design them, creating guidelines and tools for didactic and professional activities. In particular, she developed the Mapping the IoT Toolkit (mappingtheiot.polimi.it), an accessible kit to aid in the design of IoT devices.

Emma Zavarrone. Ph.D. in Strategy, Management and Quantitative Methods in Business (Univ. Bicocca), she is Associate Professor in Social Statistics and Demography at Iulm University. She teaches statistics and textual analysis laboratory for social media. She is the Rector's delegate for Third Mission and a member of the Evaluation Committee (Nucleo di Valutazione). She is the coordinator of two European projects on the issues of sustainability and dissemination of STEM among young women. She is the scientific director of Humanities Analytics Lab, structure based on textual and AI language analysis. Her main research interests are related to structural models and textual methodologies. She is the author of several international publications.

Artificial intelligence is more-or-less covertly entering our lives and houses, embedded into products and services that are acquiring novel roles and agency on users.

Products such as virtual assistants represent the first wave of materialization of artificial intelligence in the domestic realm and beyond. They are new interlocutors in an emerging redefined relationship between humans and computers. They are agents, with miscommunicated or unclear properties, performing actions to reach human-set goals.

They embed capabilities that industrial products never had. They can learn users' preferences and accordingly adapt their responses, but they are also powerful means to shape people's behavior and build new practices and habits. Nevertheless, the way these products are used is not fully exploiting their potential, and frequently they entail poor user experiences, relegating their role to gadgets or toys.

Furthermore, AI-infused products need vast amounts of personal data to work accurately, and the gathering and processing of this data are often obscure to end-users. As well, how, whether, and when it is preferable to implement AI in products and services is still an open debate. This condition raises critical ethical issues about their usage and may dramatically impact users' trust and, ultimately, the quality of user experience.

The design discipline and the Human-Computer Interaction (HCI) field are just beginning to explore the wicked relationship between Design and AI, looking for a definition of its borders, still blurred and ever-changing. The book approaches this issue from a human-centered standpoint, proposing designerly reflections on AI-infused products. It addresses one main guiding question: what are the design implications of embedding intelligence into everyday objects?

FrancoAngeli
La passione per le conoscenze