Proceedings of the Future Technologies Conference (FTC) 2018

Volume 1

Springer

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-030-02685-1 ISBN 978-3-030-02686-8 (eBook)
https://doi.org/10.1007/978-3-030-02686-8

Library of Congress Control Number: 2018957983

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Ten Guidelines for Intelligent Systems Futures

Daria Loi(✉)

Intel Corporation, Hillsboro, OR, USA
daria.a.loi@intel.com

Abstract. Intelligent systems – those that leverage the power of Artificial Intelligence (AI) – are set to transform how we live, travel, learn, relate to each other and experience the world. This paper details outcomes of a global study, where a multi-pronged methodology was adopted to identify people's perceptions, attitudes, thresholds and expectations of intelligent systems and to assess their perspectives toward concepts focused on bringing such systems in the home, car, and workspace. After background details grounding the study's rationale, the paper first outlines the research approach and then summarizes key findings, including a discussion on how people's knowledge of intelligent systems impacts their understandings of (and willingness to embrace) such systems; an overview of the *domino effect of smart things*; an outline of people's concerns with, flexibility toward and need to maintain control over intelligent systems; and a discussion of people's preference for *helper usages*, as well as insights on how people view Affective Computing. Ten design guidelines that were informed by the study findings are outlined in the fourth section, while the last part of the paper offers conclusive remarks, alongside open questions and a call for action that focuses on designers' and developers' moral and ethical responsibility for how intelligent systems futures are being and will be shaped.

Keywords: Intelligent systems · Design guidelines · Ethics of AI

1 Introduction

This paper discusses outcomes of a global study, in which a multi-pronged methodology was adopted to identify people's perceptions, attitudes, thresholds and expectations of intelligent systems and to assess their perspectives toward concepts focused on bringing such systems into the home, car, and workspace. The paper is divided into five sections. Background details to ground the study's rationale are first offered, followed by an outline of the study approach. Key study findings are summarized in the third section of the paper and the fourth highlights ten design guidelines that were informed by the study findings. Conclusive remarks are finally offered, including open questions and a call for action.

K. Arai et al. (Eds.): FTC 2018, AISC 880, pp. 788–805, 2019.
https://doi.org/10.1007/978-3-030-02686-8_59

2 Background

Intelligent systems – those that leverage the power of Artificial Intelligence (AI) – are set to transform how we live, travel, learn, relate to each other and experience the world. While these systems so far proved beneficial through scripted automation and transactions (e.g. sensor-based factory automation or phone-based health or financial transactions), serious challenges emerge when they are designed to have unscripted, autonomous, active roles. Challenges increase further when these systems include Affective Computing abilities [1] or become integral part of the environments we inhabit daily (e.g. home, office, school, vehicle). As these systems continue to be developed, overlapping concerns are accelerating and becoming mainstream – from fears of jobs replacement [2] to the emergence of surveillance [3] or deeply unequal societies [4], to name a few. At the core of such concerns is the realization that AI and intelligent systems may challenge, if not threaten, the fundamentals of human and social behaviour and the very foundations of our society. As Bostrom and Yudkowsky [5] point out, "although current AI offers us few ethical issues that are not already present in the design of cars or power plants, the approach of AI algorithms toward more humanlike thought portends predictable complication."

The fascinating part of the AI debate is that negative focus is often directed to the systems, as if they had the ability to come into existence autonomously. While AI will be able to independently design and develop another AI [6], as of now intelligent systems have one common trait: they are designed by people – typically data scientists, often assisted by designers, social scientists, and business experts that make decisions on what to design, how, why and what data to feed into systems, making them smart over time, *training them*. An challenging aspect of such a process is that people are not perfect nor fully predictable – in other words, not only human biases may play a key role while systems are designed, human imperfections (i.e. traits of humanity) may also have repercussions on systems themselves, once they start interacting with the world.

A good example is offered by Tay, a chatbot designed as an experiment to "experiment with and conduct research on conversational understanding" [7] capable of getting smarter by engaging with people in casual conversations. In less than 24 h from its launch, Tay was morally corrupted (and subsequently shut down) as its users did something not accounted for: they fed the system all sorts of misogynistic, racist remarks and Tay simply learned – "repeating these sentiments back to users, proving correct that old programming adage: flaming garbage pile in, flaming garbage pile out" [8]. The Tay example shows how an intelligent system may channel or enable unwanted, unpredicted behaviors because of at least three aspects that may have been underestimated during the design process:

- people can be unpredictable,
- unpredictability implies a potential for unpredictable outcomes, and
- unpredictable outcomes may damage initial design intentions as well as the context surrounding the system.

Stanford researchers recently made public work that utilized Deep Neural Networks to detect sexual orientation from facial images. While in their paper Wang and Kosinski

[9] explained their rationale, mentioning that "findings expose a threat to the privacy and safety of gay men and women", their work turned against them, infuriating LGBT advocacy groups [10], attracting AI experts' criticism and disapproving readers' email threats, and becoming the center of an ethical review by the American Psychological Association (cleared, see [11]). In this example, while scientists appeared to have a benevolent agenda, their work negatively impacted the very cohort they allegedly intended to protect – as well as themselves. Again, it is clear that an intelligent system may end up channeling or enabling unwanted, unpredicted behaviors, likely because of key aspects that were underestimated during the design process.

AI and Intelligent Systems are in desperate need for ethical as well as design guidelines. While AI has greatly evolved from a technical point of view, it is in its infancy as far as ethics and design process goes. The challenge is not only a technical one, it is first and foremost a social, cultural, political, ethical one. Jake Metcalf articulates this issues when he states that "more social scientists are using AI intending to solve society's ills, but they don't have clear ethical guidelines to prevent them from accidentally harming people (…). There aren't consistent standards or transparent review practices" [12].

The AI ethics debate is palpable yet not novel, given the number of organizations focused on the topic (e.g. Partnership on AI; Leverhulme Centre for the Future of Intelligence; Data & Society) and publications [5, 13–15]. It is clear that there are a number of unaddressed social, behavioural, decisional and moral questions and that great responsibilities are on the shoulders of those in charge of designing and developing intelligent systems. As Bostrom [14] puts it, while we could build a "superintelligence that would protect human values", the "problem of how to control what the superintelligence would do" looks rather difficult – within this context, designers and technologists have key roles, agencies and responsibilities.

Ackerman [16] proposed that when AI lets us down it is not due to its creators' lack of care, it is due to the social-technical gap that exist between "what we know we must support and what we can support technically". Agreeing with Ackerman's [16] view, this paper additionally proposes that a future enriched and enabled by intelligent yet trustworthy, ethical systems requires careful implementation of guidelines that govern the actions of designers, technologists, social scientists, and business experts that decide what to design, how, why and what data to feed into a given system. The study here reported was motivated by a need to contribute to conversations on such guidelines.

The ethics debate surrounding intelligent systems has been and still is dominated by two forces: data science on one side, social science and humanities on the other: on both sides sit experts. Barocas and Boyd [15] well discuss such a polarization, adding that "the gaps between data scientists and critics are wide, but critique divorced from practice only increases them". Adding to Barocas and Boyd [15] perspectives, in this paper it is proposed that "practice" should be extended to include end users' everyday life expertise. In other words, end users should actively participate in this debate and related decision making.

This perspective is at the core of why the study described in this paper focused on identifying design guidelines that are inspired from, supported by and grounded in everyday people's perspective, attitudes, thresholds and expectations toward intelligent

systems. Derived from people's everyday practices, the guidelines focus on empowering designers and developers to shape human-centric AI futures. The study did not aim at creating the ultimate guideline list – rather, it was conducted to identify practical people-centric recommendations that will hopefully spark a healthy debate on the processes used to develop intelligent systems and the agency that designers and developers should have in such processes.

3 Study Approach

The study at the center of this paper adopted a multi-pronged approach that mixed four very diverse techniques: a large scale market analysis; a multi-country survey; 18 in-home qualitative interviews; and one participatory workshop with 8 participants.

The market analysis, conducted at the start of the project, focused on intelligent systems from a landscape perspective, with emphasis on nine verticals (home, office, factory, retail, entertainment, public transport, automotive, classroom, learning) and nine vectors (players, products, academic research, investments, partnerships, associations, mergers and acquisitions, policies, events). The analysis focused on existing secondary research (e.g. publicly available data such as academic publications, press releases, reports, whitepapers, and databases) to ground protocols for subsequent study phases and help isolate key focus verticals. At the end of this first phase, smart home, autonomous vehicles and smart workspace were selected as key verticals to focus on in subsequent phases.

Survey and in-home interviews focused from a quantitative as well as qualitative perspective on two key areas: people's perceptions, attitudes, thresholds and expectations of intelligent systems; and people's perspectives toward specific scenarios of intelligent systems in home, autonomous cars, and workspace. A series of jargon-free descriptions were created and used with participants to:

- explain what intelligent systems are and what technologies they include;
- provide a series of scenarios of what such systems may enable;
- describe what smart homes, autonomous vehicles and smart workspaces are;
- offer specific examples showing what smart homes, autonomous vehicles and smart workspaces may enable.

In addition to gathering feedback on a wide range of themes, a series of metrics were collected to facilitate comparative analysis for each scenario and description:

- 1 to 5 Likert ratings to identify comfort levels or assess concepts across seven parameters (relevance, uniqueness, appeal, quality, comfort, excitement, trustworthiness);
- Word associations exercises, where participants were asked to provide feedback to concepts by selecting three items from a list of adjectives (e.g. exciting; creepy);
- Emotion association exercises, where participants provided feedback by selecting three items from a list of emotions (e.g. love/desire; worried/fearful).

While in-home interviews were conducted in the US, the survey was conducted in US, PRC (People's Republic of China) and Germany. Participants for both survey and

interviews were recruited using a screener that focused on several criteria, including: age; gender; smartphone, PC and intelligent systems ownership; and intelligent systems purchase intention (refer to Fig. 1 for sample details). The screener also focused on soft quotas, such as family composition and income, and had a natural fallout in relation to users' knowledge of intelligent systems.

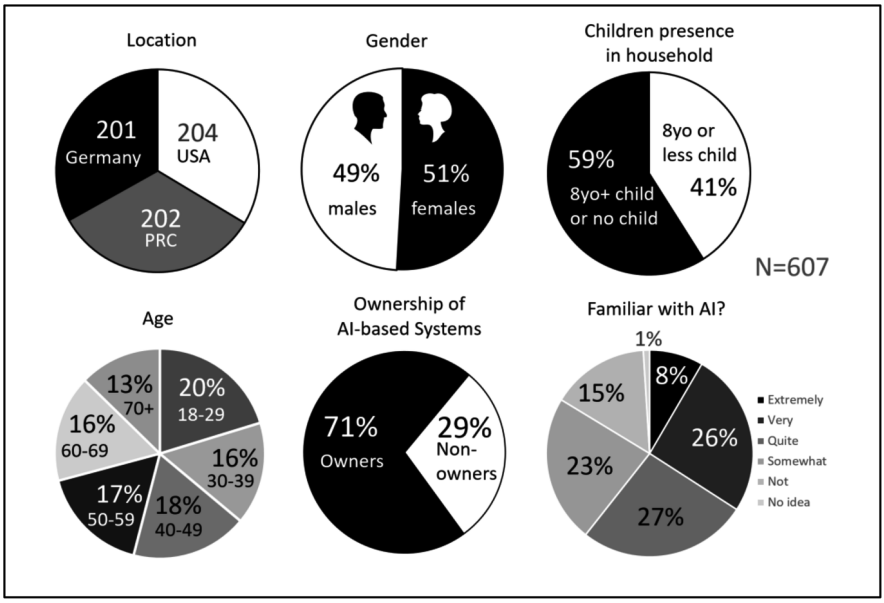

Fig. 1. Participants sample details. Source: *Loi, D. 2017*

The survey, administered to 607 participants, focused on:

- Ownership and intent to purchase intelligent systems;
- Comfort levels with embracing intelligent systems in four diverse contexts (home, car, workspace and classroom);
- Grouping intelligent systems' scenarios into one of four clusters: must have, nice to have, do not want, and not sure;
- High level feedback to smart home, autonomous cars and smart workspace;
- Comfort level with specific usages focused on smart home, autonomous vehicles and smart workspace; and
- Comparative feedback to smart home, autonomous cars and smart workspace concepts.

In-home interviews lasted a total of two hours per participant, during which observational techniques were mixed with a semi-scripted interview approach that mirrored the above-mentioned survey's flow, focus and criteria. After completing survey and in-home interviews, a subset of interviewees was invited to a participatory workshop where themes were further-explored and participants co-created a manifesto

to regulate intelligent systems futures. It should be noted that a similar workshop structure was subsequently used within a professional conference setting [17, 18].

4 Results Highlights

Given the sample size and the multiple approaches used in the study, a vast amount of data was collected and analysed. Not all data will be reviewed in this paper – while data that grounded the guidelines is discussed in the following sub-sessions, additional findings will be discussed in future publications.

4.1 What One Knows Makes a Difference

During survey and interviews, participants were asked to rate their likelihood to purchase an intelligent system for smart home, autonomous vehicle and smart workspace twice: at the start of the study and at the end, once participants had the opportunity to enrich their knowledge through provided documentation.

Overall, participants' likelihood to consider a smart home was rated higher when compared with ratings for autonomous vehicles and smart workspaces and, during interviews, participants were clearly more excited about this context of use. However, when comparing ratings collected at the start with those collected at the end of the session, data shows a drop in US and Germany ratings and an increase in PRC ones (Fig. 2). Moreover, when comparing pre- and post-ratings by gender, data shows a frequent drop in women's ratings, while male ratings stay the same or increase. A similar trend was noted during in-home interviews. This data seems to indicate that the notion that knowledge equals understanding, and that understanding may equal higher likelihood to embrace a new concept may not always apply to intelligent systems.

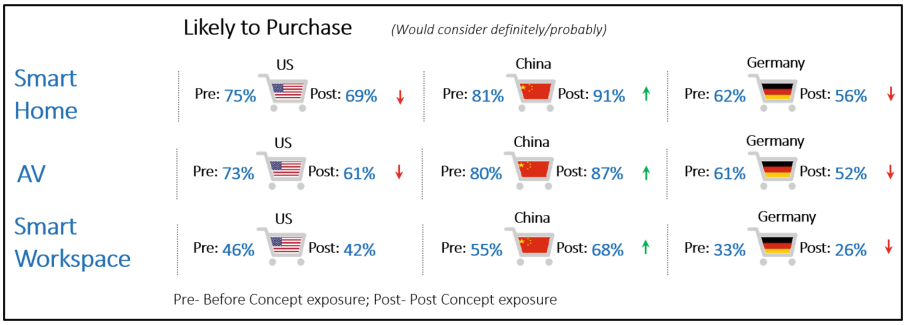

Fig. 2. Response to concepts before and after exposure to details. Source: *Loi, D. 2017*

It is proposed that amount and type of provided information play a role in people's perception and willingness to embrace intelligent systems. Culture and gender play an even more crucial role. This data highlights that *how* intelligent systems are explained

as well as demonstrated to consumers will be central to their willingness to embrace, or reject, such systems.

4.2 Once They Have One, They Want More

An interesting trend, here called the *domino effect of smart things*, emerged during in-home interviews. This will be illustrated with the story of Catherine (pseudonym), a 40 year old woman that shares a three-story detached house with her husband and 9 year old daughter (Fig. 3).

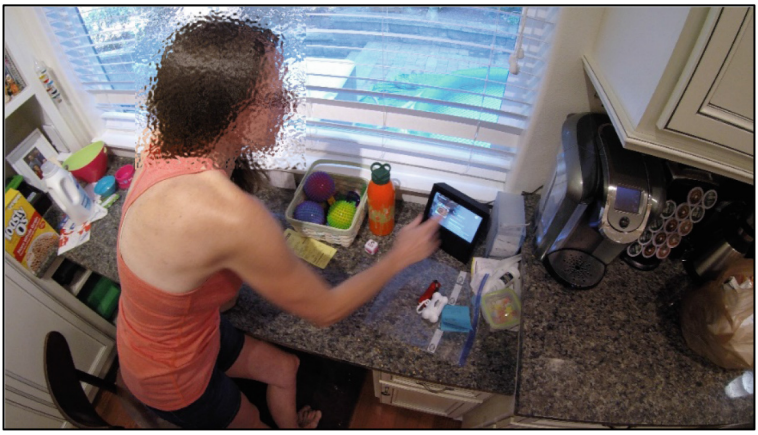

Fig. 3. Catherine discusses the benefits of her smart home systems. Source: *Loi, D. 2017*

About a year ago Catherine purchased an Amazon Echo Dot and put it in the living room. After using the device for "music and reminders", she starts enjoying more features yet realizes that the device's range is confined to one room only. She therefore purchases a second system for the TV room and, soon after that, Catherine purchases a new system for her daughter's bedroom, primarily to "listen to calming music at night". Her daughter is on the autism spectrum and Catherine shares how pleased she is with the independence these devices are providing to her daughter. A fourth Dot is soon acquired for the home office. Then, in June 2017, Amazon Echo Show start shipping – Catherine learns that she can buy two systems for a reduced cost and does not hesitate: one systems is purchased for the kitchen and the other for the main bedroom. She loves the ability to use the two new systems as a video intercom as they encourage her daughter to be more independent, while providing the ability to visually check on her as needed. Catherine proudly shows me that she can use her Dots and Shows to operate her new smart alarm system as well as the new smart light systems. She also explains that she may soon purchase a smart lock for her main door. While showing how to inter-operate her devices, she shares that she wishes they did a better job of understanding when to listen (and not listen) or when she is talking to one versus another system. "They are not perfect", she says.

Catherine's story was not unique during this study and survey data seems to indicate that this domino effect may be common. For instance, when comparing data related to ownership versus likelihood to purchase new intelligent systems, numbers show that intent is almost invariably higher than existing ownership. Not only, PRC participants (who owned the highest average of devices/person) expressed an higher intent to purchase than US and Germany counterparts. There seems to be a direct correlation between amount of owned intelligent systems and willingness to get even more. While exciting news for those that manufacture and sell such products, this trend could easily backfire if such multitudes of systems fail in satisfying people's need to have consistency and reliability in how they relate to each other.

4.3 Everyone Is Scared, yet Everyone Is Prepared to Compromise

Another clear trend identified through survey and then deepened through in-home interviews relates to participants' general fears and preoccupations with being part of an artificially intelligent world. Sonia, a 66 year old retiree that spends time between grandkids and learning about technological innovations, shared a sense of resignation and acceptance when she stated: "I am not sure I want to live in a world where everything is artificial and intelligent, even if I can see a place for these things". More combatively and critically, 33 years old small business owner Nathan shared concerns with technology with the potential to impact relationships, "affecting intimacy, creating dependency". Most interviewees referred to intelligent systems as something useful yet deeply problematic.

Many well understood the quid pro quo of this technology: to be smart, an intelligent system needs to learn and to learn, data must be fed to the system – personal data. At the same time, all participants seemed open for negotiations, prepared to accept and compromise, as long as a clear Return on Investment (ROI) is provided in return.

In some cases accessing intelligence was worth the inconvenience of everyday intrusions: "You say 'Alexa' or 'Echo' and it wants to start talking and you do not even know how or why […] It feels like somebody is in our world [laughs]. We got used to it but it is one of the things we dislike" (Catherine, 40). In other cases, people liked the convenience but wanted to ensure they could still maintain the ability to be human: "In life it's good to make mistakes so you can learn from them… here feels like you would not make mistakes anymore […] I can see the convenience but I can see that since everyone learns by doing, here I would not get a chance to" (Sheila, 69). Some were painfully aware of the fact that compromises will be needed: "I do not mind if my info is being shared but it is not good when the data can be used against you… it all goes down to what you are willing to compromise" [Stuart, 52]. Others made it clear that intelligent systems will need to provide a range of options, empowering them to choose based on personal comfort zone: "Camera is a bit too much for me […] but I know it's needed for lots of these things so I guess it depends on the privacy options you have – provided you know where the data is and what is being used for" (Amanda, 28).

As previously stated, if clear ROI is provided, people appear rather open to negotiate, accept and compromise access to their data. The key is to provide usages that have high ROI – these are discussed in the next section.

4.4 Safe, Efficient, Practical, Transparent

When survey participants that declared not having an interest in purchasing AI-based systems (N = 70) were asked to provide the rationale for their adversity (see Fig. 4), 39% of them listed security concerns, right after their top motivator: cost. At the third place they listed privacy and intrusiveness issues. The trend mirrored findings from in-home interviews: "It's always listening… How secure is it? […] what do they do to protect you?" (Jules, 39). Additionally, many participants had specific expectations: "My concern is that these things can be hacked so I expect them to be designed so they are safe" (Sonia, 66). People explained how their trust in a system is interlinked with their trust toward those providing it: "What if a business or social change occurs and the initial agreement behind the system changes? Where does (data) end? […] this is a power that can be abused and I know that people always abuse power. It's not that I do not trust technology, I do not trust people" (Nathan, 33). Moreover, participants often referred to brand trust as something that would make or break their willingness to consider an intelligent system: "It'd have to be a company I trust, that has a proven record of keeping things private, no security breaches, scams and things like that" (Amanda, 28).

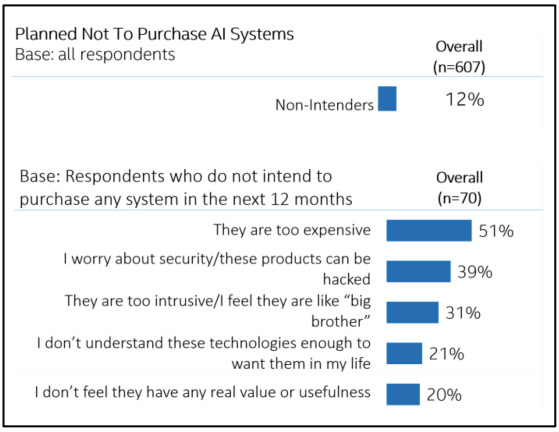

Fig. 4. Top motivators for not purchasing an AI-based system. systems. *Source*: *Loi, D. 2017*

Possibly due to these privacy and data security concerns, many expressed greater openness to and interest in intelligent systems focused on making them efficient. This is well demonstrated by survey responses to a question where participants were asked to group provided intelligent systems usages according to four clusters, namely must have, nice to have, do not want, and not sure. As illustrated in Fig. 5 (usages abbreviated) the utilitarian usage "remind me of tasks and meetings" was rated as number one "must have" overall and number one for US (62%) and Germany (66%).

Rank	Top 10 Features Shown	Overall (n=607)	US (n=204)	China (n=202)	Germany (n=201)
	Remind me of tasks/meetings	61%	1 62%	54%	1 66%
	Ask before automating things	57%	2 54%	3 55%	2 61%
	Challenge me if I am making a wrong decision	49%	37%	55%	3 55%
	Know what I'm doing to prevent making mistakes	47%	3 45%	2 58%	39%
	Provide companionship to elderly/people in need	47%	31%	1 62%	48%
	Provide personalized entertainment/news	43%	36%	54%	37%
	Offer relevant advice	43%	36%	45%	46%
	Perform manual labor	40%	33%	49%	39%
	Provide recommendations/alerts based on my activities	39%	34%	45%	38%
	Automate actions based on preferences/past behavior	32%	29%	42%	26%

Fig. 5. Top 10 Must Have usages. Includes Top 3 ranking by country. *Source: Loi, D. 2017*

Top 10 *high comfort* Smart Home usages	
Alert when house requires maintenance/upgrading	42%
Monitor activities to alert actions if someone is hurt	39%
Use audio, light switches or projections to help find items	34%
Monitor activities to alert actions to prevent property damage	32%
Trigger alerts/home automation based on behaviours/preferences	31%
Track valuable data and automatically add to health chart	26%
Keep track of physical activity to detect early sign of disease	24%
Understand who is giving commands to personalize actions	23%
Monitor/facilitate play & school activities without parental supervision	21%
Generate user profiles for those interacting with the system	19%

Fig. 6. *High Comfort* Smart Home usages (abbreviated). % indicates amount of participants that feel "very comfortable" with a Smart Home performing the activity. *Source: Loi, D. 2017*

Another indication of this efficiency trend is visible from survey feedback received in relation to Smart Home usages, where top ranking usages show a clear preference for usages focused on maintenance, prevention, and efficiency (Fig. 6, usages abbreviated).

4.5 The Secret Life of Emotions

The study also tested a number of Affective Computing [1] usages, all focused on the ability to identify the emotional state of a person or group of people to activate a series of context-appropriate actions (e.g. personalized recommendations or interventions).

When faced with usages that relate to such an intimate topic, people often paused and their responses included a deep sense of skepticism, aversion, curiosity, and distrust.

Some participants felt intrigued yet did not trust the ability of an affective system to be reliable and smart enough: "Thought provoking […] In theory is great but it'd need to be super sophisticated. Not sure it can be THAT sophisticated [emphasis added to mirror participant's vocalization]" (Jim 65) and "the issue is not with discomfort with the action but doubts that it can do it properly and reliably (Sheila, 69).

Others felt such usages would be intrusive: "Having a device monitoring my mood is too personal […] This is beyond what a machine should be doing. Keeping track of things is ok, but emotional state? I do not see this as a positive thing at all" (Esther, 34).

Some participants did not oppose affective systems but opposed to the idea of a system with a conversational human agency: "I do not want the car to check up on me but I do not think this [idea] is a bad thing" (Catherine, 40), while only a minority of interviewees expressed excitement about the notion of human systems: "The more personal technology gets… that'd be great. Not only smarter but actually human-like" (Amanda, 28).

It should be finally added that during in-home interviews people often offered ideas on how an affective system may specifically benefit them in given situations, demonstrating the contextual nature of their willingness to embrace them.

4.6 Smart Versus Intelligently Independent Systems

Feedback demonstrated the clear line that people draw between smart versus independent systems. Many interviews clarified that an intelligent system to them means convenience and that, although open to some serendipity, they need control, predictability and consistency. This need to be and feel in control is exemplified by data related to the usage "ask before automating things", which ranked second, with 57% of users (aggregate numbers) selecting it as a must have feature (Fig. 5). People's discomfort toward independent systems (and their need to keep control) was often interlinked with discomfort toward technology with its own personality and perspectives, as such traits were often seen as yet another way for a system to become independent, overstepping beyond *acceptable smartness*.

It should be noted that while personality and perspectives were generally poorly received, participants saw specific contexts where they would be not only acceptable, but desirable. For instance, the "provide companionship to elderly or people in need" usage was very well received, ranking in the top five must have features for USA and Germany and first for PRC (Fig. 5). This indicates that in specific application-contexts (such as companionship) traits such as personality and the ability of having a perspective, acting with some degree of independence, are acceptable if not desirable. It appears that the ROI of companionship usages is high, since people showed openness to compromise on system traits that would be otherwise considered undesirable.

5 Ten Design Guidelines for Intelligent Systems Futures

The previous section focused on insights gathered over the course of the study at the center of this paper. This section showcases 10 design guidelines that were directly informed by those insights.

5.1 ONE: Take a Firm, Unambiguous Ethic Stand – Be a Trusted Brand

Intelligent systems have been and will continue to be exposed to high scrutiny, and rightly so. In fact, scrutiny will increase, thanks to increased mainstream awareness, standards, and governmental mandates. Intelligent System designers and developers must not only promote ethical practices, they must design them: be firm in their ethical stand, ensuring that such a stand is present throughout their design and development process. They must be a trusted brand and design trusted systems – this includes a responsibility to speak up and not commit to (nor enable) new ideas, designs and developments if they appear to break one's ethic stands.

5.2 TWO: Adopt the *Minimize Intrusion Mantra* and a Less-Is-More Approach

In case of doubt, intelligent system designers and developers should use minimalism as a compass. This means ensuring that intelligent systems strictly collect the minimal data (type and amount) that is required for successfully achieving a requested transaction. During the study here reported, it was clear that the more one becomes familiar with an intelligent system, the more one trusts and feels comfortable in using it. However, familiarity requires not only time, it requires careful design and consideration.

A way to use a minimalist approach is to set a system's default settings at a basic level – basic functions, mirrored by basic amount of data collection, use, storage and exposure. An intelligent system should be capable of dynamically changing its settings, based on direct users' requests or feedback loops embedded in the system. Such a system should be conceived as an organism that adapts to the user and should never collect more data than what is required to satisfactorily complete a task, unless specified otherwise by its users.

5.3 THREE: Design Socially Trusted and Trustworthy Platforms

This guideline incorporates a number of sub-guidelines, all centred on ensuring intelligent systems are designed to be socially trusted and trustworthy. These include:

- Intelligent systems must fail safe;
- Privacy and hacking concerns must addressed upfront – for instance, by offering data protection services and warranties as part of the product;
- Check and balances mechanism must be embedded into the fabric of a system;
- By default, all data should be encrypted;

- Data types should be separated – only the user's system should have the ability to assemble them into a cohesive picture;
- Similarly to online money transaction models, *how* an intelligent system does something should be separated from *what* is being done and *where* it is done;
- Intelligent systems' motivations and actions must be transparent;
- Users must have the ability to provide feedback to an intelligent system and the system must take all feedback into account as well as explain *how* and *when* provided feedback will be executed; and
- An intelligent system must explain *where* data is stored, *where* it may go, *who* can access it (and *why*), and whether it will stay somewhere permanently – in accessible, transparent ways.

5.4 FOUR: Do not Make Systems Human, but Capable of Helping Humans

During the study at the center of this paper, interviewees clearly articulated what type of relationship they wish to have with intelligent systems: one where the system is in a subordinate role, never a peer. People expressed a need to be in charge and, in respecting such a need, the system should be designed to ensure that there is no ambiguity on who is in control. Unless otherwise specified and authorized by the user, an intelligent system should always ask before acting, with the exception of emergencies, where additional behavioural rules will be needed and agreed on. Since human-like attributes are typically associated with an unwelcome level of independence, it is recommended to:

- Design *helper systems,* with clear power boundaries;
- Avoid designing systems that behave (or are perceived) as assuming or arrogant – this is a particularly important point for affective systems;
- If emotion recognition is an available capability, tackle emotions by context and embed in the system ways to educate people about them. However, never assume on behalf of a user and always leave full control on actions and behaviors to end users;
- Do not underestimate people's scepticism on affective usages' reliability; and
- Consider using emotion understanding to help people help and connect with other people.

5.5 FIVE: Prioritize Usages that Matter – Helper Usages

The fact that a technology *could do something* does not imply that it should. Designers and developers should be mindful and reflective of this precept: in the case of intelligent systems this is an extremely important point to consider. Pushing usages with low (or no perceived) ROI or usages that may be (or be perceived as) ethically questionable or low in purpose will have long term repercussions on the product's success, users' willingness to embrace it, and potentially society.

People during this study clearly expressed what usages have high ROI: utilitarian usages that make them feel efficient yet in charge; usages that make them save money,

energy and time; usages that reduce their preoccupations and remove frustrations; and usages that allow them to remove boundaries and focus on what really matters. When designing intelligent systems, it is recommended that everyday chores, efficiency, and helper usages that keep a clear hierarchical distinction between helper and master are tackled first. Systems should then be capable of identifying what type of advanced usages may be pertinent and of interest to end users and should be designed to educate as well as ramp up users to such advanced opportunities.

5.6 SIX: Design Systems with Consistent Behaviors, Yet Design for Serendipity

People expressed a duality throughout the study: on one side, they asked for consistency and reliability, on the other they did not want to feel predictable and asked for technology that can enrich their understandings and even surprise them (within comfort zone). It is recommended to design systems that are enriching and predictable yet designers and developers should embrace the challenge of designing for contextual, personalized serendipity.

5.7 SEVEN: Make People Feel Unique and Empower Their Unique Goals

Part of human nature is the need to feel special, acknowledged as an individual. During many interviews people described how they want to feel throughout their technological interactions: unique. Many for instance resented the notion of a system that is so smart that can predict their behavior spotlessly, as feeling predictable makes them feel boring and less unique. Additionally, most expressed a need for technology capable of empowering them so they can achieve their unique goals – especially those goals that would be out of their reach otherwise.

Intelligent Systems should make people feel connected, wanted, and acknowledged – a system should make users feel that they are cared for and it should have ability to help users care for others and their surroundings. Companionship, social connectedness, and mediated social interactions all offer great design and development opportunities for addressing such a human need and for enriching people's everyday lives, especially the lives of those that may be in higher need for assistance, support, and nurturing (for instance, senior citizens).

5.8 EIGHT: Create Multiple and Diverse Educating Tools

One size fits all approaches are rarely satisfactory – in the case of intelligent systems design, they would heavily compromise how people understand, perceive, relate to, and embrace such systems. During the study many asked for and expressed a strong sentiment that they have the right to get a clear idea of what a system is, what it does, where data goes as well as who does what and when.

Designers and developers should focus on empowering people so they can make informed choices on whether and how to incorporate intelligent systems in their lives. This requires not only an appreciation that people have diverse baseline understandings

on these matters: it requires the development of multiple, diverse, contextual ways (content, methods, tools) to educate people on what they are in for and how to choose what is best for them, their surroundings and communities. These educating tool must avoid cryptic, tech-centric and confusing lingo – people need to understand these systems, not be confused or feel betrayed by it. While this guideline should apply to any product, in the case of complex systems powered by AI, where trust is a massive sticky point, it becomes fundamental.

5.9 NINE: Design On-Boarding Mechanisms that Grow and Evolve

Borrowed from human resources [19], onboarding is a term that user experience designers adopted to describe the process used to ramp up users, making them familiar with a new site, app, or service and increasing the likelihood that they will continue using such a site, app, or service. When well implemented, onboarding assists users in learning how to use an application incrementally, avoiding cognitive overload. In the context of intelligent systems, it is suggested to use this technique not only to ramp up users as they start engaging with a system, but also to:

- Adjust the system's behavior as it increases its knowledge of the user and as a user's understanding of, and trust toward, the system changes. For instance, a system may ask the user: *I noticed you tend to do X every day at the same time, do you wish me to do Y instead of Z in the future?*; and
- Contextually explain to the user what will be compromised or what could be gained if a setting was changed or activated in relation to a system's recommendation. For instance, the system may say: *Given you seem to like X, I think you may also enjoy Y. If you wish to try Y out, note that I will have to gather Z data under W circumstances.*

Intelligent systems should be able to evolve as their relationship with their users evolve, they should have ability to grow up and grow old with their owners, and should transparently empower people, equipping them to choose what is best for them and to change their choices over time.

5.10 TEN: Create Families of Products

During interviews with owners of several smart devices, a need for smart inter-operability across devices often emerged as a key theme. Additionally, during the research participants showed a general tendency to use (or want to use) technology outside its initially designed purposes – tendency that reached new levels when they interacted with their multiple intelligent systems: they expected such systems to be smarter, able to dynamically adjust to their everyday practices, and capable of perfectly *collaborating* and understanding each other, regardless of who manufactured them and regardless of their original purpose. In light of such behaviors, designers and developers should adopt a family-of-products design mindset, carefully designing for users' expectations, potential cross-devices usages and likely mis-usages. Each intelligent system should be designed as a node in a complex, dynamic network of systems.

6 Conclusions and Implications

The study reported in this paper aimed at identifying design guidelines that are inspired from, supported by and grounded in everyday people's perspective, attitudes, thresholds and expectations toward intelligent systems. Thanks to a multi-pronged approach that included qualitative and quantitative tools, a number of key insights were discussed. First, the paper discussed how people's knowledge of intelligent systems impacts their understandings of (and willingness to embrace) such systems. After an overview of the *domino effect of smart things*, the paper articulated that while people have great concerns, they are prepared to flex their comfort zones if there is an evident ROI. Then, it was demonstrated that people want to maintain control over intelligent systems and that they have a preference for efficiency, *helper* usages. Finally, insights on how people view Affective Computing [1] were offered, alongside a discussion showing that while people are open to smart things, they are less enthusiastic toward intelligent independent ones. These insights were then used to articulate ten design guidelines, namely:

- Take a firm, unambiguous ethic stand – be a trusted brand
- Adopt the *minimize intrusion mantra* and a less-is-more approach
- Design socially trusted & trustworthy platforms
- Do not make systems human, but capable of helping humans
- Prioritize usages that matter – *helper* usages
- Design systems with consistent behaviors, yet design for serendipity
- Make people feel unique and empower their unique goals
- Create multiple and diverse educating tools
- Design on-boarding mechanisms that grow and evolve
- Create families of products

The study here discussed was not intended to produce the ultimate design guidelines – rather, it was conducted to identify practical people-centric recommendations that will hopefully spark a healthy debate on the processes used to develop intelligent systems and the agency that designers and developers have and should have in such processes. Within such a debate, a number of questions remain in need for deepening and practical development, including:

- What ethical considerations should designers and developers prioritize?
- What level of autonomy and agency should intelligent systems have?
- Should autonomy and agency change contextually or by context of use? How?
- What level of transparency should be provided to end users? How?
- How should an intelligent system relate to, converse and engage with users?
- What specific design attributes may enable systems that are effective and accurate yet unobtrusive, respectful, intuitive and transparent intelligent?
- Can a human-centric approach to intelligent systems be effective while enabling sustainable business models and technological progress?
- What social and behavioral contracts should underpin people's interactions with intelligent systems?

Designers and developers have the moral and ethical responsibility to engage with how intelligent systems futures are being and will be shaped. A future enriched and enabled by intelligent yet trustworthy, ethical systems requires careful implementation of guidelines that govern the actions of those in charge of deciding what to design, how, why and what data to feed into a given system. Designers and developers are called on to be challenged by and contribute to the complex yet exciting task of shaping the present and future of intelligent systems.

References

1. Picard, R.W.: Affective Computing. MIT Press, Cambridge (1997)
2. Global Artificial Intelligence study: sizing the prize. PwC. https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html. Accessed 11 Oct 2017
3. The age of AI surveillance is here, Quartz. https://qz.com/1060606/the-age-of-ai-surveillance-is-here/. Accessed 30 Sep 2017
4. Are we about to witness the most unequal societies in history? The Guardian. https://www.theguardian.com/inequality/2017/may/24/are-we-about-to-witness-the-most-unequal-societies-in-history-yuval-noah-harari. Accessed 23 June 2017
5. Bostrom, N., Yudkowsky, E.: The ethics of artificial intelligence. In: The Cambridge Handbook of Artificial Intelligence, pp. 316–334. Cambridge University Press (2011)
6. Google's New AI is better at creating AI than the company's engineers, futurism. https://futurism.com/googles-new-ai-is-better-at-creating-ai-than-the-companys-engineers/. Accessed 23 June 2016
7. Trolls turned Tay, Microsoft's fun millennial AI bot, into a genocidal maniac, The Washington Post. https://www.washingtonpost.com/news/the-intersect/wp/2016/03/24/the-internet-turned-tay-microsofts-fun-millennial-ai-bot-into-a-genocidal-maniac/?utm_term=.388462f65470. Accessed 23 June 2016
8. Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. The Verge. https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist. Accessed 24 Mar 2016
9. Wang, Y., Kosinski, M.: Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. J. Pers. Soc. Psychol. **114**, 246 (2018)
10. GLAAD and HRC call on Stanford University & responsible media to debunk dangerous & flawed report claiming to identify LGBTQ people through facial recognition technology. https://www.glaad.org/blog/glaad-and-hrc-call-stanford-university-responsible-media-debunk-dangerous-flawed-report. Accessed 7 Nov 2017
11. Study claiming AI can detect sexual orientation cleared for publication. KQED. https://ww2.kqed.org/futureofyou/2017/09/13/can-facial-recognition-detect-sexual-orientation-controversial-stanford-study-now-under-ethical-review/. Accessed 7 Nov 2017
12. AI research is in desperate need of an ethical watchdog, Wired. https://www.wired.com/story/ai-research-is-in-desperate-need-of-an-ethical-watchdog/. Accessed 14 Jan 2017
13. Gunkel, D.J.: The Machine Question: Critical Perspectives on AI, Robots, and Ethics. MIT Press, Cambridge (2012)
14. Bostrom, N.: Superintelligence: Paths, Dangers, Strategies. Oxford University Press, Dangers (2014)
15. Barocas, S., Boyd, D.: Engaging the ethics of data science in practice. Commun. ACM **60** (11), 23–25 (2017)

16. Ackerman, M.S.: The intellectual challenge of CSCW: the gap between social requirements and technical feasibility. Hum. Comput. Interact. **15**(2–3), 179–220 (2000)
17. Loi, D., Raffa, G., Esme, A.A.: Design for affective intelligence. In: HCII 2017, San Antonio (2017)
18. Loi, D., Lodato, T., Wolf, C.T., Arar, R., Blomberg, J.: PD manifesto for AI futures. In: PDC 2018. Hasselt & Genk, Belgium (2018)
19. Bauer, T.N., Erdogan, B.: Organizational socialization: the effective onboarding of new employees. In: Zedeck, S. (ed.) APA Handbook of Industrial and Organizational Psychology, vol. 3, pp. 51–64 (2011)