

RethinkAI™: Designing the Human and AI relationship in the Future of Work

Laloti, Vali

Royal College of Art, School of Design, London, UK
vali.laloti@rca.ac.uk

The innovation landscape is drastically changing due to the adoption of Artificial Intelligence (AI), as whole industries are incorporating AI into smart products and automated processes. Designing AI for industry 4.0 requires revolutionary thinking. It requires the emergence of new design paradigms that build designers' ability to navigate the ethical and socioeconomic issues that AI brings in the future of work. This research develops RethinkAI™ a new participatory design method to address this gap. The paper focuses specifically on how to design the relationship of humans and AIs working together.

RethinkAI™ builds an interactive, social way for designers and transdisciplinary teams to explore this relationship. It creates insights on how human and AI strengths can be designed together in a co-evolving relationship in the future of work. RethinkAI™ was designed, tested and refined through a series of three workshops. The method was evaluated in a final workshop with 24 multinational industry professionals, involved in the decisions of integrating AI into their products and processes. The paper analyses the results from its application. It explores how this participatory method can be a valuable and rich medium to stimulate new thinking into the current design paradigms of humans and AI systems interaction. The analysis indicates that the use of human strengths and participatory method can help develop designers' agency in surfacing and mitigating social and ethical issues when designing AI into products and services. The paper concludes with a reflection on current insights and direction of further research.

Keywords: new design paradigms; artificial intelligence; future of work; design innovation design; participatory methods; transdisciplinary design research

1 Introduction

Designing Artificial Intelligence (AI) for industry 4.0 requires revolutionary thinking, if it is to give back human agency and ability to navigate ethical and socioeconomic issues in the future of work (Huxley, 1957) (Bostrom, 2014). AI is bringing major changes in the nature of work as a result of rapid advances in machine learning and cognitive technologies. It brings strengths that create efficiencies and help solve problems at scale, enhancing cognitive and physical capabilities of humans in the future of work (McKinsey Global Institute, 2017).

Though the benefits are clear, there are a multitude of socioeconomic implications and ethical issues in AI design which are rapidly developing the complexity and characteristics of a “wicked problem” (Emmerson & Young, 2017).

These range, from concerns over job losses (Deloitte Review, 2017), the well-being of gig economy workers and dehumanisation of the augmented worker (Bhatnagar et al., 2018) to ethical issues rising from AI autonomy (Stewart, 2018) and dystopian futures from singularity (Medeiros, 2017). Dystopian narratives about AI are compelling and the calls to regulate and create policies that safe guard societal values are justified. This happened throughout human history, as our technologies evolved and changed how we work. During the first industrial revolution, for example, regulation for child labour was put forward through 1833’s Factory Act.

1.1 Research Rational and Motivation

The goal of this research is to inform the design of work with Artificial Intelligence (AI) and to navigate forwards, beyond the polarized views and negative publicity of dystopian futures or the current positively over-hyped AI scenarios.

This paper focuses specifically on how to design the relationship of humans and AIs working together. It brings new ways on how to help designers explore this relationship, how to surface human and AI strengths and preferences on how these are combined when working together with AI. The ultimate aim is to stimulate new thinking into the current design paradigms of humans and AI systems interaction, to inform the design of human work with AI in a coevolving relationship.

As with the IoT and the need for new product development processes (Lee, Cooper, & Hands, 2018) there is a need to shift our design paradigm when designing AI in the future of work. As AI evolves, becoming more autonomous, our human-centred design paradigms also need to evolve and extend interaction and user experience design into designing the relationship with AI. This is why this paper’s design research into methods and tools to explore and design the relationship between human and AI agents at work is significant.

1.2 Research Questions and Methods

The key research questions therefore addressed in this paper and contributing towards this broader vision are:

How can a design research exploration of the relationship of human and AI at work inform the design of work with AI?

Q1: What are the current issues in the relationship between the human and AI at work?

Q2: How can we research human and AI strengths coming together at work?

Q3: How can this research exploration surface human preference in AI design?

Q4: How can this research exploration inform the design of work with AI?

The research methodology brings together research in design through literature review and use of participatory research methods in workshops with business professionals (Sanders, 2006) (Sanders, 2008). Action research is used in an iterative approach, to plan, action, observe and reflect on the research (Eikeland, 2001), for each of the questions above, with each participatory workshop informed the design of the next workshop, as shown in Figure 1.

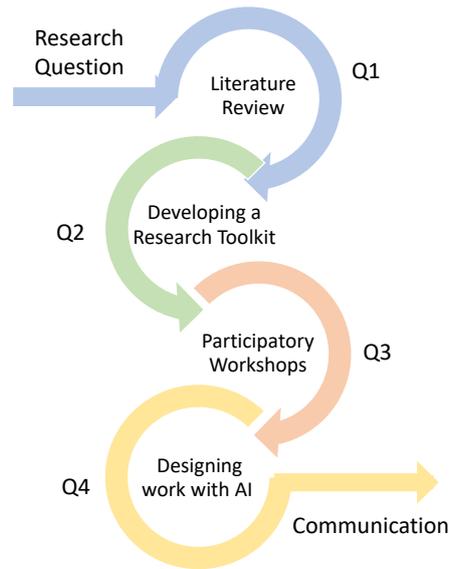


Figure 1. The Research Methodology combines research in design with research through participatory design practice, within an Action Research framework, to explore the four sub-questions, Q1, Q2, Q3 and Q4.

The participatory research kit, was developed as a glossary of human and AI strengths and designed as a game with 50 cards. It provides an interactive, social way to create insights in communication between designers and business professionals. The research frames the design of work with AI in a coevolutionary relationship of human and AI strengths working together. The participatory research approach was developed through a series of workshops with business professionals, including business coaches, MBA students and business managers from across different functions and diverse industries. The toolkit was evaluated and improved through an iterative design process. It was informed by extensive literature review on the wide range of challenges and ethical issues around AI and the future of work.

The participatory research approach is validated in a workshop with 24 Chief Finance Officers (CFOs) from a global steel corporation. CFOs are chosen because they are leading important decisions on how AI is integrated into work due to the economic drivers for AI use. It has been found that it enables the designer in communication with business decision makers, to build understanding of human strengths and challenges in our relationship with AI at work. It empowers participants to co-design AI in their work as a coevolving relationship between humans and AI. It stimulates a dialogue between designers and business professionals on how to design AI into human work.

The literature review is described in section 2 and the iterative design process and workshops are detailed in section 3 of this paper. Section 4 concludes with the future research directions.

2 AI and the future of Work

What are the current issues in the relationship between the human and AI at work?

From a technological perspective, AI is a computer program that learns by itself. There are, however, real issues with how AI learns. A key one is, that human biases are being passed on to AI algorithms, through the data sets that are used to train AIs. Google's image search engine categorising black faces as a gorilla is the most often cited example (Simonite, 2018).

Another key issue, is that of transparency of how AI algorithms are reaching their predictions and decisions, often referred to as the issue of “explainability”. This is significant in view also of AI becoming autonomous. The first fatal accidents involving autonomous cars (Stewart, 2018) have surfaced new fears around AI autonomy and the recent Cambridge Analytica scandal has brought out questions on the drivers behind AI use (Tett, 2018). These concerns are also echoed by the broader AI community. For example, Mustafa Suleyman, co-founder and Head of Applied AI at DeepMind, Google’s AI company, argues for “up-ending the processes by which technology is developed and deployed, and proposes new ways to ensure it contributes to greater economic and social justice around the world” (Suleyman, 2017).

Journalists posing as gig economy workers are gathering first hand experiences in companies such as Uber and Deliveroo of work conditions and ethical dilemmas (Bloodworth, 2018), raising awareness of the impact on human well-being (Rapacki, 2018). Smart devices, like Amazon’s wristband (Olivia Solon, 2018) that facilitate workflows, connecting AI with the human worker, also spark concerns on the de-humanization and control of the worker. These inform the debate and raise questions on the kind of work relationship between human and AI we want to create (Bostrom, 2014)(Harari, 2018).

While regulations and policies for AI are drawn up, there is also a pressing need to build ownership and the agency of the decision makers in business, to understand and negotiate the ethics and challenges when integrating AI into human work. This also requires an evolution in our design paradigms. Matt Jones, principal designer for Google AI is advocating for a human-centred approach that designs our relationship with AI, beyond the industry’s current design paradigms which stem from human-computer interaction methods and user experience toolkits (Koch, 2017) (Amico, 2017), advocating a collaborative relationship with AI (Jones, 2018) (Google, 2019).

Japan, for example, sees AI as the solution to its socioeconomic challenges resulting from its aging demographic. As Japan’s prime minister said, “Japan has no fear of AI. Machines will snatch away jobs? Such worries are not known to Japan. Japan aims to be the very first to prove that growth is possible through innovation, even when a population declines” (Kharpal, 2017). Popular culture also with comic heroes like astro-boy (Schodt, 2007) has reinforced the “AI as a friend” relationship in Japan.

The strengths that AI brings are welcomed in Japan, as a solution to its ageing population which impacts its workforce. Japan has low levels of foreign workforce hence AI and its robotic materialisations are considered a solution and not a problem. The UK is seen job losses from robotic automation which are hitting the headlines and the well-being of human workers, thus AI strengths are perceived as replacing human strengths and agency. This results in also differences in the way AI strengths are designed into work and its acceptance in the society (Ito, 2018). The next section describes the participatory methods and iterative approach used in the research toolkit design.

3 Human-AI relationship in the future of Work

To explore this relationship, a card-based toolkit and a participatory method was developed. A participatory approach was chosen as it can open up a dialogue between designers and business professionals on how to design work with AI. This is in the form of four workshops with business professionals, involving them as co-creators in the design research process.

The participatory research artefacts, were designed as thinking tools, to help designers and participants research how human and AI strengths can come together at work. These were developed as a glossary of strengths, through the first workshop with business coaches (section 3.1). The participatory research approach was tested through the second workshop with MBA students (section 3.2) and refined in the third workshop with 20 business managers from across functions and industries (section 3.3). The final workshop evaluated the method and explored how it can be used to inform the design of work with AI in a business case workshop with a group of 24 executives from a global steel industry corporation (section 3.4).

3.1 Designing a participatory research toolkit

How can we research the relationship of human and AI strengths coming together at work?

If we are to design the work relationship between human and AI, as a coevolving, collaborative relationship, a good place to start is research into how humans work together. There is a plethora of organisational research tools, to explore and improve the dynamics of how humans work together. One category of these, focuses on individuals' strengths and how these might be brought together more effectively when working in teams. Some of the most researched and well proven ones, used by global corporations and NGOs are Myers-Briggs MBTI® (Myers-Briggs & Myers, 1980) (Jung, 1971) and Belbin® (Belbin, 2010).

Through this research in related organisational tools, an initial glossary of positive traits or strengths was developed. These were translated into a glossary of strengths and designed as a kit of coloured cards, with one word on each card, describing a human strength at work. Alternatives and complementarity of this glossary of words, was tested through a one-hour workshop with six business coaches, accredited to work with MBTI® and Belbin® typologies.



Figure 2. Developing a glossary of human strengths, designed as a set of 50 cards.

For this workshop, an initial set of coloured cards were complimented by blank cards, which the participated business coaches used to add words or change existing ones. For example, “bossy” which is used in (Shapiro, 2010), was changed to “assertive” and words such as “detailed” or “rigorous” became part of the glossary, as shown on the left and right images of Figure 2. The initial colours were blue, green, red and brown. The brown colour was changed to yellow and orange hues, because brown was perceived as dull to represent people’s strengths, in comparison to the other colours (see central image in Figure 2). This resulted a kit of 50 coloured cards used as a form of participatory artefacts and a thinking tool throughout the research practice.

3.2 Testing the Design Research Kit

These participatory artefacts, were further tested in an hour-long workshop with a group of four, part-time MBA students at Westminster Business School. This group was chosen, as they are working managers and students at the same time and in the point of considering their future careers and work beyond the MBA.

After a short introduction on the future of work and how this might change because of AI, participants were invited to choose from the set of 50 cards the strengths that best describe them. The choice was restricted to 5 cards, as the intention is for participants to narrow down their selection to their most valued strengths. Participants were then asked to look at their cards and think which of these strengths AI could also have in the future. Participants dropped these strengths from their hands to a black board, as shown on lower right image in Figure 3. The black board was chosen because, according to the literature review, AI remains a black box, in terms of our understanding of how it reaches predictions and decisions.

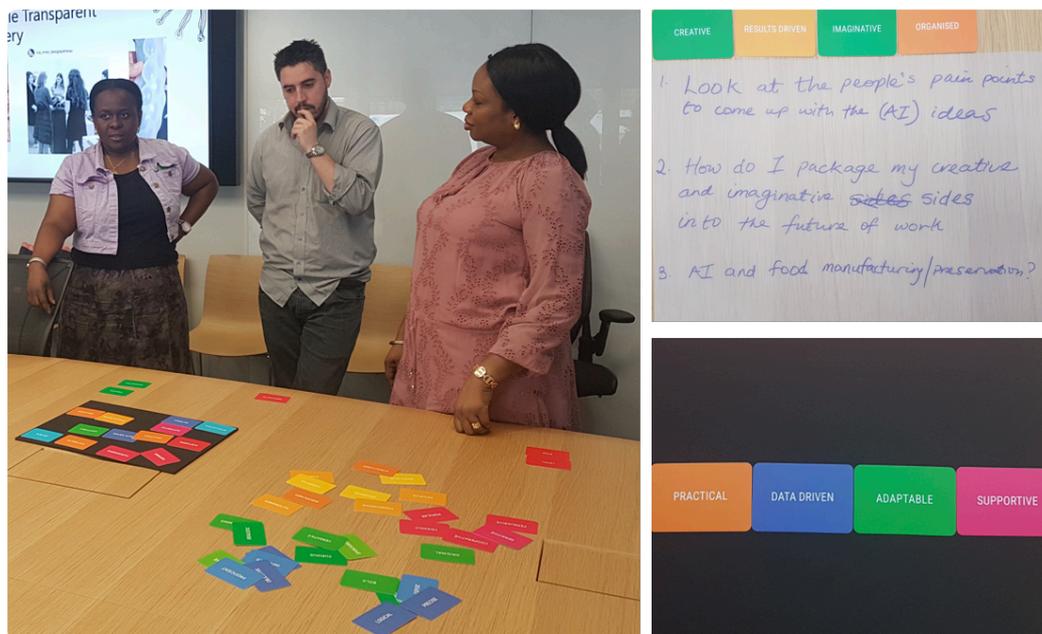


Figure 3. Participants reflect on AI strengths (left and bottom right) and their career options (upper right).

Participants discussed what it means for them and their own strengths and how they might develop these in the future. At the end participants were given time to write down how they might grow their strengths in their future careers working with AI, as shown in Figure 4. One participant, for example, decided to focus on their “kindness” and “empathy” to build their HR career on “diversity and inclusion”, while another chose to “package their creative and imaginative strengths” in their career in innovation in “food manufacturing and preservation.”

Key findings are: a) it enabled participants to surface their own strengths and which of these are most important to them, b) removing cards from their own strengths reinforced the negative, fear element about AI, c) participants needed a warm up step to familiarise themselves with the words on the cards, as presented all 50 of them was a lot of information to absorb as a first step and d) the reflective and group discussion time allowed rich insights into how participants view their future work with AI.

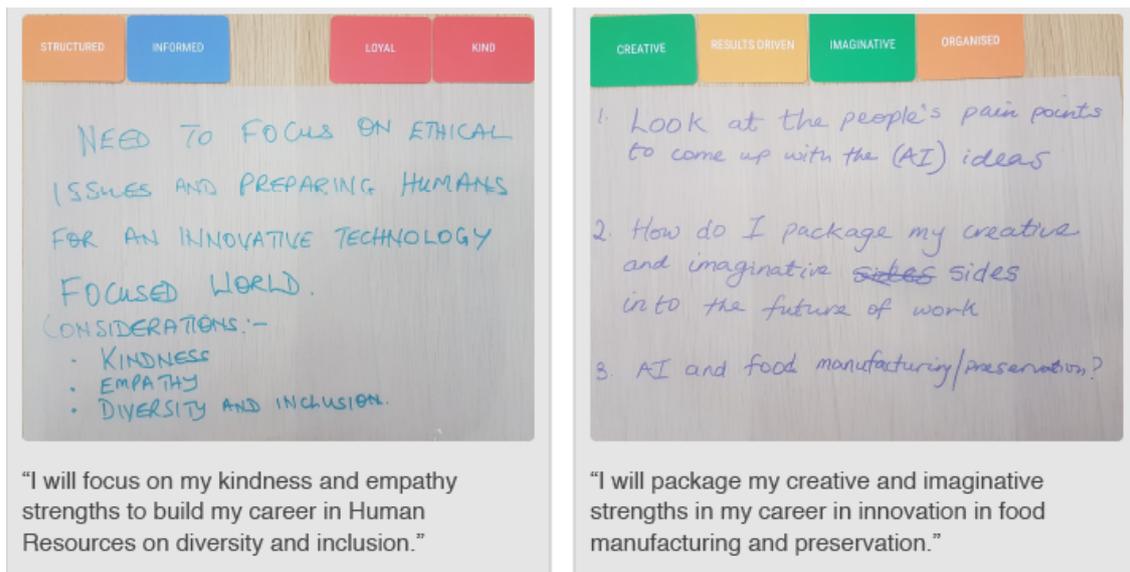


Figure 4. Participants' reflections on their strengths and how they can best use these in the future of work with AI.

3.3 Refining the design research approach

How can this research exploration surface human preference in AI design?

The glossary and participatory research methods were further refined in a third workshop of one and a half hours long, with 20 business managers attending an Agile Management Conference. Based on the key findings from the previous workshop, a warm up game was introduced to familiarise participants with the words on the cards. Participants are split into groups of 5 to 6 people, each dealt with 5 cards. Participants first exchange cards with others in their group, to end up with five cards that best represent their strengths. The remaining cards are spread on the table with the words visible to all. Participants can drop a card from their hands and pick up one from the table, ending up with five cards that represent their most valued strengths.

Participants then choose from their own and the remaining cards on the table, the strengths that they would like AI to have, to work better together with them. These are placed on a black board, as before. Therefore, participants design the AI as "a board of strengths". They are given time to reflect and negotiate their own preferences and those of others in their groups into a final design. Groups present their AI designs, with time to reflect and debate their designs with other groups (see Figure 5).

AI designs across groups were different, shown in Figure 6. For example, one team designed AI as "Loyal, Friendly, Supportive", because as the group articulated they would like to see, "AI as supportive to the user, so technology is a friendly access to services offered. AI is playing a supportive role." Another group designed their AI as "Analytical, Objective, Structured". They articulated that they prefer their AI to support them on analytical tasks and structuring their work, rather than as a friend: "we like to design AI to be friendly and supportive, but as a human if you have a machine that is supportive and empathetic, and you have a friend, where would you go? I would say, we would always go for the human."

Key findings are: a) framing this research as designing the AI they would like to work with, creates positive dynamics and energetic conversations, b) the warm up game helps participants familiarise with the glossary of strengths c) the group conversations and reflections allow participants to surface and debate their preferences and biases towards AI, building up rich insights for designers and d) empowers participants to become designers of their own AI, defining it through a set of strengths and according to what they would like AI to do for them in the future of work.



Figure 5. Participants design AI using the RethinkAI™ glossary of strengths and participatory research method.

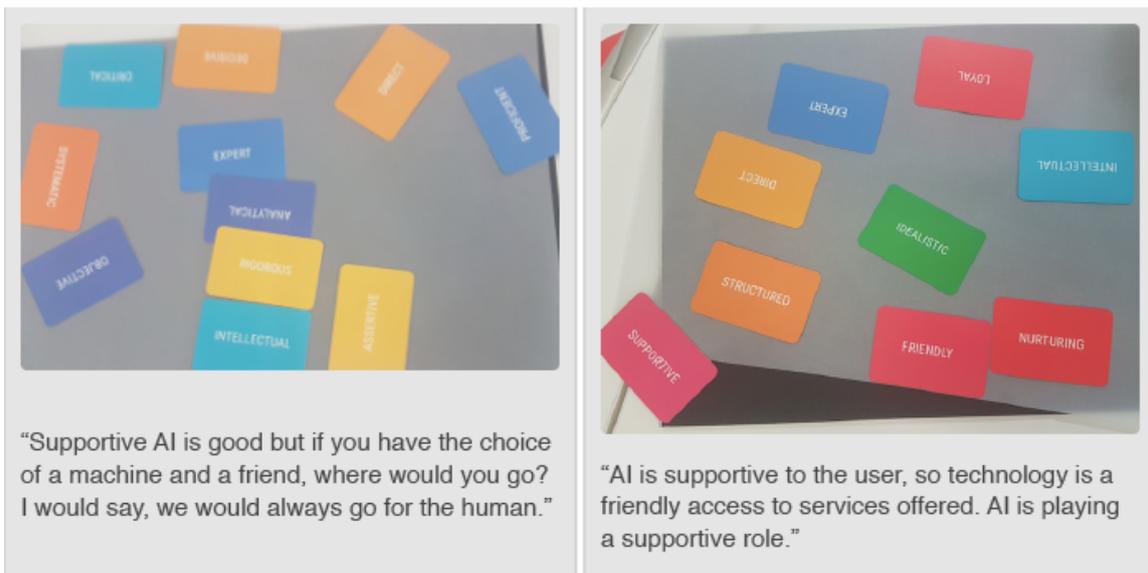


Figure 6. Reflecting on human strengths in the future of work

3.4 Validating with industry - Designing work with AI

How can this research exploration inform the design of work with AI?

The fourth workshop was a half-day workshop with 24 Chief Finance Officers (CFOs) from a global engineering corporation. The participants were CFOs of national divisions within the multinational. CFOs are chosen because they are leading important business decisions on how AI is integrated into work, due to the economic drivers for AI use in business. The aim of the fourth workshop was to evaluate the research approach in a business case and test how a participatory approach can create ownership and agency to surface and negotiate the impact of human-AI issues resulting from these decisions. It further explored, how business professionals in communication with designers can use this research approach to design their work with AI, as a coevolving relationship.

The workshop started with capturing words that the CFOs associate with AI. These were used to benchmark and compare with change of attitudes towards AI at the end of the workshop. The initial words were centred around the technical and economic benefits of introducing AI and ethical issues, such as privacy, bias and job losses. These were in line with the literature review and the emphasis of organisations in using AI to maximise efficiency and automate processes to reduce costs.

The workshop proceeded by framing the introduction of AI together with the impact on humans and workflows, as a co-evolving relationship. The octopus was used as the metaphor for designing work with AI. The octopus's nervous system is distributed, with octopus tentacles so packed with neurons that it is as if there are "thinking" independently and with their bodies (Montgomery, 2015) (Godfrey-Smith, 2017). This distributed, embodied intelligence resembles more closely the way that AIs are integrated into our workflows and business systems, connecting with other AIs and learning from humans and large data sets distributed across businesses. This reframed the design of work with AI, as designing the human and AI strengths working together in a distributed and coevolving relationship (see Figure 7).

Participants first designed their AI as a black board of strengths, as in previous workshops, and then redesigned one of their finance functions to work with AI. AI integration in organisations is currently dealt with as a technical system integration. Conversations about the impact on human strengths, human workflows and jobs are typically addressed separately and as part of business transformation processes and through Human Resources. Putting the human and AI on the same design framework, allowed the CFOs to go beyond the economics and efficiencies of AI integration. They debated what this AI integration means for the current teams and their own CFO roles.

They then designed three phases of the human-AI working relationship, based on the capabilities of the AI system in each phase: a) initial, b) learning and c) autonomous. The first phase is the introduction of AI, the second is the learning through humans and data and finally the phase of AI as more autonomous and able to make decisions. Participants used the cards to design the human and AI strengths in each phase and how these would work and evolve together as AI strengths change through these phases.

workshops with MBA students, Business Managers and Chief Finance Officers from the steel industry. Initial findings from the use of this participatory method show promising results. It is found to stimulate a dialogue between designers and business professionals on how AI and human relationship can be designed in the future of work. However, further quantitative and qualitative validation is needed, to build on current results and bring rich insights into the current design paradigms of human and AI interaction.

There are two research directions currently pursued: a) further validation with an ongoing research study within a postgraduate transdisciplinary design research educational context, complemented by quantitative data captured through a dedicated online app and b) a further investigation on how cultural differences affect AI design. The latter is currently explored through a UK/Japan collaboration, on differences in AI and human relationship in the two countries, to bring new thinking into current human and AI interaction design paradigms. Designing the Relationship of humans and AI at work remains a significant topic of research and the work presented in this paper is a timely investigation of this research topic.

5 References

- Amico, L. (2017). Designing with AI. Retrieved March 11, 2019, from <https://uniform.net/about/blog/designing-with-ai>
- Belbin, R. M. (2010). *Team roles at work*. Butterworth-Heinemann.
- Bhatnagar, S., Cotton, T., Brundage, M., Avin, S., Clark, J., Toner, H., ... Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation Authors are listed in order of contribution Design Direction*. Retrieved from https://img1.wsimg.com/blobby/go/3d82daa4-97fe-4096-9c6b-376b92c619de/downloads/1c6q2kc4v_50335.pdf
- Bloodworth, J. (2018). *Hired : six months undercover in low-wage Britain*. Atlantic Books.
- Bostrom, N. (2014). *Superintelligence : paths, dangers, strategies*. Oxford University Press.
- Deloitte Review. (2017). *Navigating the future of work*. Retrieved from <https://documents.deloitte.com/insights/DeloitteReview21>
- Eikeland, O. (2001). Action research as the hidden curriculum of the western tradition. In *Handbook of Action Research*. <https://doi.org/0432>
- Emmerson, P., & Young, R. (2017). Taking aim at 'wicked problems' A practical philosophy for educating designers in the making of wise decisions. In *International Association of Societies of Design Research Conference*. Cincinnati. Retrieved from <https://scholar.uc.edu/concern/articles/41687h449?locale=en>
- Godfrey-Smith, P. (2017). *Other minds : the octopus and the evolution of intelligent life*.
- Google. (2019). PAIR – Google AI. Retrieved March 11, 2019, from <https://ai.google/research/teams/brain/pair>
- Harari, Y. N. (2018). *21 lessons for the 21st century*. Vintage, Jonathan Cape.
- Huxley, J. (1957). *Transhumanism*. Chatto & Windus.
- Ito, J. (2018). Why Westerners Fear Robots and the Japanese Do Not | WIRED. Retrieved March 11, 2019, from https://www.wired.com/story/ideas-joi-ito-robot-overlords/?mbid=social_fb
- Jones, M. (2018). Centaurs or Butlers? Designing for Human Relationships with Non-Human Intelligences - Artificial Intelligence Videos. AI Videos. Retrieved from <https://www.artificial-intelligence.video/centaurs-or-butlers-designing-for-human-relationships-with-non-human-intelligences>
- Jung, C. G. (1971). *Psychological Types: Collected Works of C.G. Jung*. Princeton University Press. Retrieved from http://www.taylorandfrancis.com/catalogs/jung_and_analytical_psychology_2013_online_catalog/4/
- Kharpal, A. (2017). Japan has no fear of AI — it could boost growth despite population decline, Abe says. Retrieved March 11, 2019, from <https://www.cnbc.com/2017/03/19/japan-has-no-fear-of-ai--it-could-boost-growth-despite-population-decline-abe-says.html>

- Koch, J. (2017). *Design implications for Designing with a Collaborative AI. The AAAI 2017 Spring Symposium on Designing the User Experience of Machine Learning Systems Technical Report SS-17-04 Design.*
- Lee, B., Cooper, R. F. D., & Hands, D. (2018). Are Traditional NPD Processes Relevant to IoT Product and Service Development Activities ? A Critical Examination. In *Design Research Society Conference*. <https://doi.org/10.21606/dma.2018.244>
- McKinsey Global Institute. (2017). *What is the future of work?* Retrieved from <https://www.mckinsey.com/featured-insights/future-of-work/what-is-the-future-of-work>
- Medeiros, J. (2017). Exclusive Interview with Stephen Hawking. *WIRED*.
- Montgomery, S. (2015). *The soul of an octopus : a surprising exploration into the wonder of consciousness*. Retrieved from <https://www.simonandschuster.com/books/The-Soul-of-an-Octopus/Sy-Montgomery/9781451697711>
- Myers-Briggs, I., & Myers, P. B. (1980). *Gifts differing : understanding personality type*. Consulting Psychologists Press.
- Olivia Solon. (2018). Amazon patents wristband that tracks warehouse workers' movements | Technology | The Guardian. *Guardian* . Retrieved from <https://www.theguardian.com/technology/2018/jan/31/amazon-warehouse-wristband-tracking>
- PwC. (2018). *UK Economic outlook: What will be the next impact of AI and related technologies on jobs in the UK?* Retrieved from <https://www.pwc.co.uk/economic-services/ukeo/ukeo-july18-full-report.pdf>
- Rapacki, K. (2018). Precarious by Design. *Disegno*, Vol. 19.
- Sanders, L. (2006). Design Research in 2006. *Design Research Quarterly*. <https://doi.org/10.1007/s10195-015-0335-1>
- Sanders, L. (2008). ON MODELING An evolving map of design practice and design research. *Interactions*. <https://doi.org/10.1145/1409040.1409043>
- Schodt, F. L. (2007). *The Astro Boy essays : Osamu Tezuka, Mighty Atom, and the manga/anime revolution*. Stone Bridge Press.
- Shapiro, S. M. (2010). *Personality poker : the playing card tool for driving high-performance teamwork and innovation*. Portfolio Penguin.
- Simonite, T. (2018). When It Comes to Gorillas, Google Photos Remains Blind | WIRED. Retrieved February 15, 2019, from <https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/>
- Stewart, J. (2018). Tesla's Self-Driving Autopilot Was Turned On In Deadly California Crash. Retrieved February 15, 2019, from <https://www.wired.com/story/tesla-autopilot-self-driving-crash-california/>
- Suleyman, M. (2017). RSA President's Lecture: The Technologist's Dilemma - RSA. Royal Society of Art and Manufacture (RSA). Retrieved from <https://www.thersa.org/events/2017/11/rsa-presidents-lecture-the-technologists-dilemma>
- Tett, G. (2018, July 12). The Cambridge Analytica scandal echoes the financial crisis. *Financial Times*. Retrieved from <https://www.ft.com/content/b21ffb20-85b1-11e8-96dd-fa565ec55929>

About the Authors:

Lalioti, Vali: Vali has a high-profile career in academic and corporate research, with a PhD, Computer Science, MRes Design and an MBA. She worked in Europe, USA, Africa and Asia in senior posts in industry, national research institutes and as University Professor.