Ethical Considerations of Augmented Reality in High-Tech Manufacturing

Sander van der Hoek,[1] Marlies van Steenbergen,[2] Pascal Ravesteijn[3]

[1] ASML, Eindhoven, Netherland
sander.van.der.hoek@asml.com
[2] HU University of Applied Sciences Utrecht, Research Group Digital Ethics, Utrecht, Netherland
marlies.vansteenbergen@hu.nl
[3] HU University of Applied Sciences Utrecht, Research Group Process Innovation & Information Systems, Utrecht, Netherland
pascal.ravesteijn@hu.nl

The use of Augmented Reality (AR) in industry is growing rapidly, driven by benefits such as efficiency gains and ability to overcome physical boundaries. Existing studies stress the need to take stakeholder values into account in the design process. In this study the impact of AR on stakeholders' values is investigated by conducting focus groups and interviews, using value sensitive design as a framework. Significant impacts were found on the values of safety, accuracy, privacy, helpfulness and autonomy. Twenty practical design choices to mitigate potential negative impact emerged from the study.

Keywords: augmented reality, mixed reality, OHMD, human values, value sensitive design, industry 4.0

University of Maribor Press

DOI https://doi.org/10.18690/um.fov.6.2023.22
ISBN 978-961-286-804-8

1 Introduction

Driven by the travel restrictions caused by the Covid-19 outbreak, industrial companies that need to perform complex maintenance tasks sought out possibilities for overcoming the issue of not being able to have their engineers on site. These companies began to quickly adopt augmented reality technology that enabled senior engineers to instruct less experienced staff on site[1].The vast number of professional and academic publications that mention Covid-19 and Augmented Reality (AR) keywords indicates that this not only happened in the industrial sector but also in other sectors such as healthcare. The use of AR might provide an effective solution for overcoming travel restrictions (Lamberti et al., 2014), but an important question is whether companies are ready to use these technologies in an effective and sustainable way and whether they understand the possible implications.

Challenges and success factors for implementing AR are of technological, organizational and/or environmental type (Egger & Masood, 2020). Masood & Egger (2019) show that though academic research focuses more on the *technological* aspects, professionals are more concerned with the *organizational* consequences. All but one of the success factors professionals labeled as most relevant were user-centric: user acceptance, visibility of information, ergonomics, and usability of the user interface. Other research shows that, if human values are not incorporated effectively enough, companies may not be able to leverage all of AR's potential (Burleigh et al., 2020; Hofmann et al., 2017; Rousi, 2016).

This study addresses ethical aspects of using AR technology in an Industry 4.0 context. The main research question for this study is: How can the application of AR technology in the high-tech semiconductor industry take human values into account to positively influence the impact on its stakeholders? Examples of stakeholders are the users of AR glasses, but also their managers, group leads and customers. These stakeholders all have different stakes in regard to the use of AR technology.

The key concepts investigated in this study are stakeholders and their values. These concepts are studied within the context of specific AR use cases in the semi-

[1] See for instance https://www.smartindustry.com/tools-of-transformation/augmented-reality/article/11414929/how-covid-19-boosted-augmented-reality-and-virtual-reality-in-manufacturing).

conductor industry, to stimulate discussion of values in more detail. AR technology might impact the personal values of stakeholders in a positive or negative way. Based on insights into these impacts on stakeholder values, designs can be adapted to foster the positive impact and limit the negative impact, resulting in a design that is in balance with the values of the different stakeholders.

In section 2 we discuss the theoretical background to our study, followed by a description of our research method in section 3. Section 4 presents the results, which are discussed in section 5. The paper ends with conclusions in section 6.

2 Theoretical background

Within the high-tech industry the concepts of industry 4.0 and 5.0 are often mentioned. Industry 4.0 represents the 4th industrial revolution with high levels of mechanization, automation, digitalization and miniaturization (Lasi et al., 2014). While in the smart manufacturing paradigm of Industry 4.0 the role of human workers has become increasingly smaller, Industry 5.0 is bringing human workers back into the picture (Longo et al., 2020). The 5th Industrial Revolution combines human and machine, creating synergy between the two (Nahavandi, 2019). It is expected that the use of AR technology will significantly increase with Industry 5.0 concepts as new technical improvements are introduced (Fernández del Amo et al., 2018). Technical obstacles seem to become less inhibitory in large-scale usage and AR appears to be effective in both transferring knowledge *to* users and discovering and capturing knowledge *from* users (Fernández del Amo et al., 2018). Future factories will combine different technologies such as Big Data, AI and AR (Shi et al., 2020), enabling individualized human-machine interaction (Xu, 2021).

Though Industry 5.0 is purported to be human-centric instead of technology-driven (Xu, 2021), Longo et al. (2020) raise ethical concerns as the new technologies of the factories of the future are expected to have great impact on human values. They argue that more attention must be paid to the means by which technologies in Industry 5.0 systems can be designed for human values rather than treating these values as a side issue. This is the more relevant when the boundaries between human and technological capabilities are blurred, as is the case with AR technology. Regarding AR outside the context of Industry 4.0, Pase (2012) argues that AR is a persuasive technology, having the potential to intentionally influence or modify behaviors, values or attitudes, and as such raises ethical concerns. The main impacted values he discusses are manipulation of people, privacy and physical safety. Slater et al. (2020) explains the concept of physical vs. psychological realism. They discuss ethical issues that may arise when it becomes more difficult to differentiate the real from the virtual world. A third example of research regarding ethical considerations in AR is by Hofmann et al. (2017), who found a variety of ethical issues related to privacy, safety, justice, change in human agency, accountability, responsibility, social interaction, power and ideology.

3 Research Method

This study uses the Value Sensitive Design (VSD) approach. VSD is a theoretically based approach that takes human values into account in technical design. In VSD human value is defined as "What is important to people in their lives, with a focus on ethics and morality" (Friedman & Hendry, 2019, p. 24). VSD distinguishes between three types of investigations into stakeholder values: conceptual (general knowledge from literature or experts), empirical (eliciting actual values and norms from stakeholders) and technical (translating values into technical affordances). These investigations are intended to be iterative, feeding each other and allowing the designer to continuously modify the design (Friedman & Hendry, 2019).

In this study, first, values that might be impacted by technology design were distilled from existing academic literature (conceptual investigation). Next, these values were validated by a focus group (Bell et al., 2019) existing of four AR specialists who were involved with AR initiatives in the same high tech industry company: 1) a business leader 2) a business process specialist and expert in AR 3) a business solutions architect and 4) a systems architect. The focus group also identified two relevant scenarios of AR technology use in high tech industry as well as the stakeholders that might be impacted in these use cases.

The scenarios and value list were input for semi-structured interviews with representatives from the stakeholder groups that were identified (empirical investigation). These stakeholders were part of the same company but worked mainly at customer sites (other companies) or were associated with them. In total, 13 employees with different roles (engineers, project leads, site managers and team leads) and work locations in Northern America, Europe and Asia were interviewed.

During the interviews the participants were asked how they are involved in and affected by this technology, what values they think are most important in this context, how these values are impacted and how they think the impact can be influenced. The two scenarios from the focus group were discussed. The first consisted of Guided Work Instructions (GWI), where an engineer wears AR glasses to view step-by-step work instructions for a specific task in the form of text, holograms, audio and video. The second scenario was GWI+ where GWI is extended with more interaction between engineer and technology, because AI can "observe" the environment and work of the engineer and interact more with the engineer based on these observations. Interviewees were also asked to rank human values in order of importance to them in the context of the two given scenarios. The data gathering phase of this study took place in the summer and fall of 2021.

A thematic analysis was performed across the interviews to identify themes and search for patterns involving repetitions, similarities and differences. This was accomplished by using the code groups: harms, benefits, constraints, do's, don'ts, enablers, values, value tensions and mitigating measures. During the coding, at first the inductive method was used to code all content, while the grouping of codes used a more deductive approach (Linneberg, 2019). Relations between items and code groups were identified. Benefits, harms, do's and don'ts were linked to values in cases where there was a clear and strong relationship. Two kinds of relationships were used: relationships that indicate a positive contribution to the value and relationships that indicate a negative impact on the value. Creating this structure allowed further analyses of the raw interview data and helped find patterns and draw conclusions.

Subsequently, a formula was used to calculate the overall value ranking based on all interviews. One point was given to a value if it was selected in third place, 2 points if it was rated as second, and 3 points were given to values chosen as most important. These scores were added up, leading to a ranking with overall scores per value. In addition to the scores, two other variables are used to compare the outcome of the value ranking: groundedness and density. Groundedness refers to the number of quotations associated with a code, and density refers to the number of codes linked with a value.

After the interviews a second focus group session was conducted with the aim of validating and complementing the mitigation measures that emerged from the interviews (technological investigation).

4 Results

In the interviews the following values emerged in descending order of priority: *safety, accuracy, privacy, helpfulness, autonomy*, care, comfort, self-control, trust, independence, power, efficiency, security, accountability, pride, social interaction, cleanliness, professionalism, health, knowledge, quality, accessibility, change in human agency, reliability, effectiveness, ergonomics, responsibility as displayed in figure 1 (note that the scales of the four dimensions scenario 1, scenario 2, groundedness and density are different; dimensions can thus be compared between values but not within one and the same value).

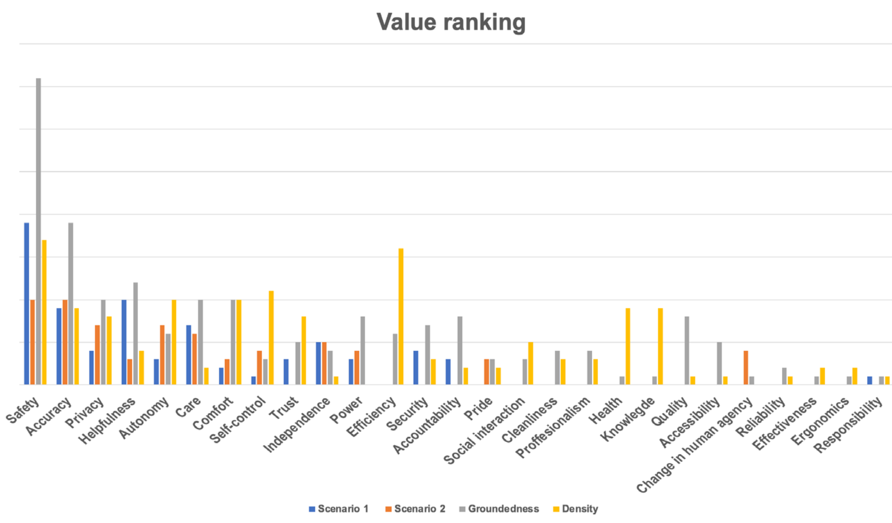

Figure 1: Value Ranking

The five values with highest priority (in italic above) are discussed in the following paragraphs, as well as the mitigating measures (numbered below) that were suggested by the participants to prevent potential negative impact on these values.

Many times, the phrase "safety first" was mentioned. Stakeholders identified several issues regarding *safety*. The wearer of an AR device is more vulnerable to falling, hitting something with their head or arms or when controlling a heavy moving object. Persons who work with lasers were cautious about open-beam hazards since a safety glass was not integrated in the AR headset. Safety sensors on the device that warn about potential safety hazards were seen as a significant improvement (1). Incorporation of safety glasses that protect the eyes from laser beams (2) was also suggested as a mitigating measure. The focus group indicated that safety regulations (3) and work instructions accompanied by safety tips (4) were the most feasible and effective measures to improve safety. Improvements in the viewing angle (5) were also mentioned. A narrator function could help prevent text objects from appearing in the field of sight (6). Advanced object recognition and safety hazard detection (7) are further possible future solutions.

Accuracy was the second most important value mentioned. The use of AR technology potentially enables the participants to do their jobs more accurately with fewer errors, and requiring less interpretation compared to the existing method of working. Accuracy can be considered an industry value in the semiconductor industry. Performing their jobs more accurately was important to the participants, but the accuracy of the AR technology itself is also important. Several measures for improvement were given such as: more accurate work instructions (8), and higher resolutions and less screen jitter (9). More accurate control of the device (10) was also often mentioned.

Privacy became an important value in the second scenario that contained more intelligent interaction between user and technology, which requires an AI function that observes the environment as well as the person. In this scenario participants became more sensitive about their privacy, and they began to raise concerns about being evaluated using AR and AI technology. Some mentioned that people behave differently when someone is wearing an AR device because others might think they are being recorded. The most significant measures to ensure privacy were to adhere to privacy laws (11), anonymize personal data (12), prevent storage of personal information (13) and not share the data with management (14). Privacy-ensuring measures such as blurring faces (15) and transparent privacy controls such as a blinking light when the device is recording (16) were also mentioned by participants.

Helpfulness was rated as the fourth important value for the stakeholders; the interviewees stated that if AR was not helpful to them, they were not going to use it. AR technology is considered a potentially helpful technology, but if it is not helpful on the spot, there is no time to make it so. Others mentioned that they need more help in using the tool to make it more helpful to them. The focus group suggested getting context-aware help in the form of projected suggestions (17). Additionally, reducing the administrative burden for engineers was proposed through intelligently assessing the components used during the assembly or maintenance task (18) and making the license portal accessible via the headset (19).

It became clear during the interviews that *autonomy* is a two-sided topic; autonomy is influenced in both positive and negative ways according to the interviewed stakeholders. They indicated a positive influence because an engineer can do more work autonomously as the information to do the task is at hand in a clear and easy to understand way. On the negative side stakeholders indicated that AR can also take away autonomy. Work instructions can become more prescriptive and rigid, leaving less room for the creativity of the engineer. The engineer becomes more dependent on the technology itself to do the job. Stakeholders expressed that if the technology, specifically in the form of work instructions, is too strict, workers would start feeling like robots or puppets, making them feel less appreciated as professionals. There appears to be a tension between working as an independent professional, and ensuring efficiency and quality by providing uniform detailed instruction; we will return to this topic later in this section. The focus group suggested giving the engineers more control and providing them with the space to make decisions as part of the work instructions by applying a state-based design principle (20).

Besides the individual values, the VSD concepts of value dams (features that some stakeholders strongly object to), value flows (features that many stakeholders support) and value tensions (values contradicting each other) were identified. The most significant value dam mentioned during the interviews involved security, more specifically the lack of intellectual property (IP) protection in the current design. Value dams are explained as design options to which even a small percentage of stakeholders strongly object (Friedman & Hendry, 2019). For companies in the semiconductor industry, protecting their IP is a major concern. IP is often considered one of the most valuable assets of a company working in this domain. Companies in the semiconductor industry often do not allow devices from suppliers that they do not control in their cleanroom as these can potentially be used to observe and steal IP-sensitive information. Thus, the technology is not easily accessible to the end user because lengthy approval processes and security controls build a metaphorical dam that hinders fostering other values such as autonomy, efficiency and helpfulness. A solution to overcome this value dam mentioned by the focus group was the capability to blur specific objects based on object recognition or virtual curtains placed on specific coordinates.

A value tension was identified between efficiency and autonomy or independence. Efficiency can be reached in the area of doing work faster by reducing execution, learning and travelling time, making fewer mistakes and having the right information at hand, leading to higher productivity. This higher productivity can be attained by building AR technology that provides more consistent work instructions, understands the context and how the activities are performed by the user, and can even correct users. Potentially, this could lead to more prescriptive and rigid ways of instructing users. Increasing autonomy means providing individuals with more freedom to determine their working procedures and task scheduling (Niessen & Volmer, 2010). This tension shows the importance of finding a proper balance between AR technology and its capabilities to instruct engineers and the level of freedom given to the engineers. Especially as autonomy is a critical aspect of both justice and wellbeing (Calvo et al., 2020). Lu et al. (2022) describes this phenomenon as empathic machines or empathic robot control where empathy skills generate human-robot shared actions based on understanding of the human state and situational circumstances.

5 Discussion

Comparing the values mentioned by the participants with existing research about Industry 4.0/5.0 and AR values, it seems that the values mentioned are more related to the AR technology itself than to the context of the semi-conductor industry. However, context does influence some of the values in terms of prioritization. Safety, privacy, comfort and accountability seem to be tightly related to the technology itself. Accuracy, care, helpfulness, and autonomy seem to have a stronger relation to the semiconductor industry.

As mentioned in the result section safety stood out as the most significant value. This relates to research by Lu et al. (2022) which positions Safety as a ground level value in their Industrial Human Needs Pyramid (Lu, 2022). The first aspect that must be ensured in a manufacturing environment is safety which can also be concluded from our study in the area of AR technology usages. Once safety is ensured other values become more important.

During the interviews two scenarios (GWI and GWI+) were described to the interviewees. In the second scenario other values emerged from the interviews. The values autonomy and privacy were considered more important in Scenario 2. The stakeholders worried that if artificially intelligent technology becomes too prescriptive, they may lose their autonomy. AR technology opens new opportunities to measure the performance of employees by using cameras and sensors. Performance monitoring of employees has been a topic of study for many years. Hawk (1994) investigated the effects of computerized performance monitoring, concluding that it can result in health problems, stress and the perception of unfairness. He urges organizations to be selective in which tasks are to be measured by computerized monitors. In 2017 an empirical study on the effects of employee surveillance showed that it leads to reduced sense of privacy and self-esteem and increased uncertainty, employee vulnerability and changes in behavior (Indiparambil, 2017). There seems to be a very thin line between using observation technology and AI for the good of the employees and using it in such a way that individuals change their behaviors for the bad, resulting in undesired behaviors and lack of intentional benefits.

Although the value of trust was not selected in the top five values during the interviews, it was mentioned quite a few times that positive outcomes on certain values such as helpfulness, privacy, autonomy, and security will build trust in AR technology, which leads to more usage and broader adoption. Since experience with AR is limited, and the technology is immature in many ways, building trust in the technology is important (Mcknight et al., 2011). Humanity's trust in technology has always been somewhat problematic, however (Taddeo, 2010). Considering the overall importance of trust in technology, the growing maturity level of AR, the potential negative outcomes and value dams and the high level of human-computer interaction (HCI) that this technology brings (Dünser et al., 2007), trust should play an important role in the further design and development of this technology.

6 Conclusion

The stakeholders of AR can be impacted by this technology in a positive but also in a negative way. This study revealed that the values stakeholders are most concerned with are safety, accuracy, privacy, helpfulness and autonomy. Twenty appropriate measures that can be embedded in the design to reduce potential negative impact on these values, were formulated during the value discussions. In addition to these five values, the importance of building trust in AR technology frequently emerged. The most significant value tension that this study disclosed was between autonomy and efficiency. Important design questions need to be answered for AR applications that provide work instructions. How much freedom will an engineer have in using GWI and will they be able to develop a personal way of working? Will GWI take existing experience into account? Will it adapt the instructions based on the level of details an engineer needs? How will mistakes or faulty executions of the work instructions by the engineer be reported?

A limitation to this study is that is performed in the context of only one company in the semiconductor industry. However, although the persons interviewed were employees of this company, they often executed their work as engineer at other major semiconductor companies. Therefore, they were able to some extent to share perspectives from those companies.

Situational factors such as local culture might influence stakeholder values and how they perceive the impact that AR has on those values. Situational factors are taken into account in this study by considering multiple stakeholders and stakeholders from different regions in the world. Interviews were held with stakeholders from three different regions (US, EU, APAC) to represent the whole population of the stakeholder group. Since the interviewees from the US, Taiwan, Japan, Korea and the Netherlands represented the global footprint of the semiconductor industry (Semiconductor Industry Association, 2020) this research provides reliable results on a global level by ruling out local influences.

The use of AR technology in the semiconductor industry is a novelty. Some interviewees had more extensive experience with AR, but overall, the experience was limited. As a minimum requirement the stakeholders selected for the interviews needed to have experimental experience with AR. Ideally, when using VSD, all participants can experience the technology extensively, including new design improvements. That was not the case in this research. Envisioned scenarios were presented that required some imagination and interpretation by the participants. We recommend to continue using VSD to validate future development.

Valuable future research could be performed in developing working instructions for AR devices. There is still a great deal to explore in developing personalized, effective work instructions. Many work instructions consist of step-by-step guides that do not take the existing skills, learning curves and personal preferences of the users into account. Since AR technology will significantly drive human computer interaction, and users need to work more and more in symbiosis with the technology, this area will be an important area for future research.

We argue that the further development of AR technology needs to profoundly take human values into account. AR technology will further mature over time, experience new iterative design cycles, and new applications will be combined with other technologies like AI. Embedding a value-sensitive design approach may prevent harm and balance value tensions in future designs.

Acknowledgements

We would like to thank all participants in the focus groups and interviews for their time and valuable insights.

References

Bell, E., Bryman, A., & Harley, B. (2019). Business research methods (Fifth edition ed.). Oxford University Press.

Burleigh, C., Kroposki, M., Magabo, M., & Bailey, L. (2020). Ethical Considerations in Designing Virtual and Augmented Reality Products--Virtual and Augmented Reality Design with Students in Mind: Designers' Perceptions. Journal of Educational Technology Systems, 49(2), 219-238. https://hu.on.worldcat.org/oclc/8686785393

Calvo, R. A., Peters, D., Vold, K., & Ryan, R. M. (2020). Supporting human autonomy in AI systems: A framework for ethical enquiry. Ethics of digital well-being: A multidisciplinary approach, 31-54.

Dünser, A., Grasset, R., Seichter, H., & Billinghurst, M. (2007). Applying HCI principles to AR systems design.

Egger, J., & Masood, T. (2020). Augmented reality in support of intelligent manufacturing - A systematic literature review. Computers & Industrial Engineering, 14010.1016/j.cie.2019.106195

Fernández del Amo, I., Erkoyuncu, J. A., Roy, R., Palmarini, R., & Onoufriou, D. (2018). A systematic review of Augmented Reality content-related techniques for knowledge transfer in maintenance applications. Computers in Industry, 103, 47-71. 10.1016/j.compind.2018.08.007

Friedman, B., & Hendry, D. G. (2019). value sensitive design. MIT Press. https://doi.org/10.7551/mitpress/7585.001.0001

Hawk, S. R. (1994). The effects of computerized performance monitoring: An ethical perspective. Journal of Business Ethics, 13(12), 949-957.

Hofmann, B., Haustein, D., & Landeweerd, L. (2017). Smart-Glasses: Exposing and Elucidating the Ethical Issues. Science and Engineering Ethics, 23(3), 701-721. 10.1007/s11948-016-9792-z

Indiparambil, J. J. (2017). An empirical study on the detrimental effects of employee surveillance in India. International Journal of Research in Computer Application & Management, 7(12), 48-51.

Lamberti, F., Manuri, F., Sanna, A., Paravati, G., Pezzolla, P., & Montuschi, P. (2014). Challenges, opportunities, and future trends of emerging techniques for augmented reality-based maintenance. IEEE Transactions on Emerging Topics in Computing, 2(4), 411-421.

Lasi, H. D., Fettke, P. P. D., Kemper, H. P. D., Feld, T. D., & Hoffmann, M. D. (2014). Industrie 4.0. Wirtschaftsinformatik, 56(4), 261-264. 10.1007/s11576-014-0424-4

Linneberg, M. S., & Korsgaard, S. (2019). Coding qualitative data: A synthesis guiding the novice. Qualitative research journal, 19(3), 259-270.

Longo, F., Padovano, A., & Umbrello, S. (2020). Value-Oriented and Ethical Technology Engineering in Industry 5.0: A Human-Centric Perspective for the Design of the Factory of the Future. Applied Sciences, 10(12), 4182. 10.3390/app10124182

Lu, Y., Zheng, H., Chand, S., Xia, W., Liu, Z., Xu, X., Wang, L., Qin, Z. & Bao, J. (2022). Outlook on human-centric manufacturing towards Industry 5.0. Journal of Manufacturing Systems, 62, 612-627.

Masood, T., & Egger, J. (2019). Augmented reality in support of Industry 4.0—Implementation challenges and success factors. Robotics and Computer Integrated Manufacturing, 58, 181-195. 10.1016/j.rcim.2019.02.003

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. ACM Transactions on Management Information Systems (TMIS), 2(2), 1-25.

Nahavandi S. (2019). Industry 5.0 - A Human-Centric Solution. Sustainability, 11(16)

Niessen, C., & Volmer, J. (2010). Adaptation to increased work autonomy: The role of task reflection. European Journal of Work and Organizational Psychology, 19(4), 442-460.

Pase, S. (2012). Ethical considerations in augmented reality applications. Paper presented at the Proceedings of the International Conference on E-Learning, E-Business, Enterprise Information Systems, and E-Government (EEE), 1.

Rousi, R. (2016). Using human-values as a guide for understanding worthy design directions in augmented reality. AcademicMindtrek 2016 - Proceedings of the 20th International Academic Mindtrek Conference, , 243-252. 10.1145/2994310.2994322

Semiconductor Industry Association. (2020). 2020 STATE OF THE U.S. SEMICONDUCTOR INDUSTRY. https://www.semiconductors.org/wp-content/uploads/2020/06/2020-SIA-State-of-the-Industry-Report.pdf

Shi, Z., Xie, Y., Xue, W., Chen, Y., Fu, L., & Xu, X. (2020). Smart factory in Industry 4.0. Systems Research and Behavioral Science, 37(4), 607-617. 10.1002/sres.2704

Slater, M. (2020). The Ethics of Realism in Virtual and Augmented Reality. Frontiers in Virtual Reality, 1

Taddeo, M. (2010). Trust in Technology: A Distinctive and a Problematic Relation. Knowledge, Technology Policy : Knowledge, Techology Policy (Untill 2011), 23(3-4), 283-286.

Xu, X. (2021). Industry 4.0 and Industry 5.0 Inception, conception and perception. Journal of Manufacturing Systems, 61, 530-535.