MASTER-XR: Mixed Reality Ecosystem for Teaching Robotics in Manufacturing

– *preprint* –

Michael Barz[1,2], Panagiotis Karagiannis[3], Johan Kildal[4], Andoni Rivera-Pinto[4,]Judit Ruiz de Munain[5], Jesús Rosel[5], Maria Madarieta[6], Konstantina Salagianni[7], Panagiotis Aivaliotis[7], Sotiris Makris[3], and Daniel Sonntag[1,2]

[1] Interactive Machine Learning, German Research Center for Artificial Intelligence (DFKI), 66123 Saarbrücken, Germany
`michael.barz@dfki.de`, `daniel.sonntag@dfki.de`
[2] Applied Artificial Intelligence, University of Oldenburg, 26111 Oldenburg, Germany
[3] Laboratory for Manufacturing Systems and Automation, Department of Mechanical Engineering and Aeronautics, University of Patras
`karagiannis@lms.mech.upatras.gr`, `makris@lms.mech.upatras.gr`
[4] Fundación Tekniker
`johan.kildal@tekniker.es`, `andoni.rivera@tekniker.es`
[5] Mondragon Lingua-Alecop S.Coop
`juditr@alecop.es`, `jrosel@alecop.es`
[6] Virtualware 2007 S.A., Usausuaga 7, 48970 Basauri (Bizkaia), Spain
`mmadarieta@virtualwareco.com`
[7] Teaching Factory Competence Center
`salagianni@teachingfactory-cc.eu`, `aivaliotis@teachingfactory-cc.eu`

Abstract. Many industries are transitioning to Industry 4.0 production models by adopting robots in their manufacturing processes. In parallel, Extended Reality (XR) technologies have reached sufficient maturity to enter the industrial applications domain, with early success cases often related to training workers, remote assistance, access to contextual information, and interaction with digital twins. In the future, robots will be increasingly enhanced with XR applications, which requires that industrial workers understand both technologies and use and control hybrid solutions confidently. Specific education and training programs will be essential to this transition, especially for vocational school students and professionals in upskilling. They must learn how to program robots and establish a safe and productive human-robot collaboration. The new EU-funded project MASTER will improve the XR ecosystem for teaching and training robotics in manufacturing by providing an open XR platform that integrates key functionalities like creating safe robotic environments, programming flexible robotic applications, and integrating advanced interaction mechanisms based on eye tracking. It will also provide high-quality training materials for robotics. We report on the project plan, our objectives, and milestones.

Keywords: Industry 4.0 · Extended Reality (XR) · Robotics · Worker Training · Manufacturing · Human-Robot Collaboration · Eye Tracking.

1 Introduction

Many industries adopt robots and Extended Reality (XR) technologies to realize Industry 4.0 production models. This requires industrial workers to understand both technologies, particularly in flexible and collaborative manufacturing settings that require frequent programming of robots and safe shared workspaces. For instance, high awareness of robot actions and intentions can help mitigate stress and potential dangers. Consequently, vocational school students and professionals must learn how to program robots and collaborate with them safely. XR technologies can support the learning process [49] and may become a key technology of future manufacturing processes. According to data published in 2021, XR will create 1.2-2.4 million new jobs in the EU by 2025, and European XR markets are expected to grow between € 35 billion and € 65 billion by 2025 [50]. Specific education and training programs will be essential for successfully transitioning to XR-based robotic workspaces like the example illustrated in figure 1. The presented EU project, MASTER, aims to provide an open XR platform for worker training in robotics and XR-enhanced human-robot collaboration. The platform will facilitate the creation of XR-based training scenarios and materials and provides key functionalities for creating safe robotic environments, programming flexible robotic applications, and integrating advanced interaction techniques based on eye tracking. In this report, we outline the objectives of our project, describe the planned XR platform and its major features, provide insight into didactic considerations for educating workers, and provide an outlook on how we plan to evaluate the educational platform in realistic settings through two open calls.

Fig. 1. Exemplary shared workspace in robotic-assisted manufacturing based on XR.

1.1 Technological Objectives

The main objective of the MASTER project is to boost the XR ecosystem for teaching and training of robotics in manufacturing. To achieve this, an open XR platform will be provided, which shall enable creators to deliver rich training content on robotics and integrate key functionalities for creating safe robotic environments, programming flexible robotic applications, and integrating advanced interaction mechanisms. Four technical objectives have been identified, which act as the main pillars of the project:

1. **Provide an open platform for the creation and management of XR content in robotics by non-experts.** We focus on the creation of an XR platform that allows non-expert programmers to create XR learning content and teaching materials on robotics. The targeted stakeholders are robot experts who know about robotic processes and related components. The platform shall enable the transition of this knowledge into XR content, exercises, and learning sessions (see section 2).
2. **Create virtual environment to program and interact with robots in safety configurations.** We address safety aspects in manufacturing, focusing on means to increase the safety and safety awareness of operators in human-robot collaboration, also concerning ergonomics (see section 2.1).
3. **Bridge the gap between the use of virtual and real robots for training and real operation.** We aim to integrate a programming toolbox that allows a user to program a robot easily. A special focus will be on Programming-by-Demonstration (PbD), as well as machine vision assisted techniques, such as Look-and-Move (L&M). These techniques exploit the benefits of the dual real/virtual nature of robots, the parts that they manipulate, and the environment, all of which the user can inspect and operate in conjunction through mixed reality (MR) (see section 2.2).
4. **Improve the interaction with XR systems using multimodal interaction methods.** We integrate gaze-based interaction techniques in 3D virtual and augmented reality settings tailored to educational use cases. We aim to integrate active gaze-based interaction for the manipulation of, e.g., robots, their digital twins, or digital menus. In addition, we plan to integrate methods for real-time analysis of the human gaze behavior for assessing the learning progress of users (see section 2.3).

1.2 Industrial and Educational Objectives

Besides our technical objectives, we identified one industrial and one educational objective that will guide our development and project activities:

1. **Industrial Objective: Boost companies, especially SMEs and start-ups, to deploy XR technology.** SMEs and start-ups rely on external financial and technical support to reach the market since only 26.6% are able to sustain themselves with their own funds. Thus, the MASTER project

focuses on providing financial, technical, and business support for companies in two open calls that will be launched during the project to develop and test new XR technologies for human-robot collaboration (see section 4).

2. **Educational Objective: Create robotics educational material based on state-of-the-art XR technologies.** Educational material is one of the most important resources in any learning process, both in educational and professional fields. The MASTER project aims to deliver state-of-the-art material created using the XR platform. We aim to develop teaching material for XR in multiple formats and adapted to learners in relevant target sectors. The XR contents will address the three core technologies presented in sections 2.1, 2.2, and 2.3, and topics of third-party industry partners (see section 3). The material will be validated by experts in robotics and education and by third parties selected in the second open call (see section 4).

2 XR Platform

A key outcome of our project will be an XR platform with three main modules that refer to our technological objectives. The implementation requires consolidating, maturing, and unifying existing technologies from all partners into a single framework. We will extend the VIROO® platform [1], a virtual reality (VR) platform with a focus on industrial applications, for that purpose. It includes tools and services for easy creation, management, and deployment of immersive content for single-user and remote collaborative multi-user settings. For instance, VIROO® *Studio* enables low code VR content creation based on the Unity 3D engine [2], also empowering users without programming experience to create VR content. In addition, it maintains simplicity in certification processes required for industrial applications. VR contents are hosted in a cloud to enable quick and easy deployment and are managed using the *Portal* service. The three modules to be developed in the MASTER project cover the key technologies of the involved research partners. The modules will focus on creating safe robotic environments, programming flexible robotic applications, and integrating gaze-based interaction techniques using eye tracking sensors integrated into modern head-worn XR devices.

2.1 Safe Robotic Environments

In the MASTER project, as described in the technical objectives, one focus is on safety aspects in manufacturing. In general, safety systems, like cameras, curtains, safety mats, etc. are installed in the production lines, to protect the human operators from hazards such as getting hit by a robot or getting trapped by an unexpected robot movement. Nevertheless, such safety sensors do not include MR technologies, but these technologies can be used to provide real-time information on the status of the resources and their operations, to help humans be proactive in their actions. Thus, integrating MR in industrial settings enables the creation of intuitive and effective interaction between robots and humans to increase the operator's safety awareness and support easy robot programming. This can be achieved through the creation of flexible frameworks, able to simulate real-life assembly operations through immersive VR [15]. Combining VR and wearable robotics with exoskeleton technology can be a promising solution that could be used in multiple industrial areas and tasks, leveraging their benefits for the operators [23]. Beyond traditional VR systems that use wearable or portable devices, a 3-wall CAVE [3] consisting of three big projectors and four infrared cameras can be used to track multiple users virtually performing assembly tasks, aiming to visualize key performance indicators (KPIs) in order to support production engineers in decision-making [37]. Using augmented reality (AR) can enable real-time interaction between operators and the workplace. For instance, insights of experienced operators can be integrated into common execution flows through audio commands or manual robot guidance [20], also using smartwatch applications [51]. In addition, the provision of safety aspects, such as notifications or warnings, is a crucial feature for enabling safe human-robot collaboration [29]. AR can be used for visualizing the robot envelope of a static robot inside the operator's field of view using a depth sensor [21].

Although a lot of research has been done on how to exploit AR and VR in human-robot collaboration environments, there is a lack of integrated solutions that facilitate a seamless interaction of robots with human operators in the manufacturing environment. The lack of interfaces to support human operators in cases of unexpected events related to the robot is an important limitation in the existing literature that will be addressed in our project.

2.2 Programming Flexible Robotic Applications

The demand for a high degree of personalization in production is a feature of the Industry 4.0 production paradigm. Satisfying these demands requires that industry operators become agile in generating robot programs without writing code. Various advanced techniques have been proposed in the literature, referred to as **robot teaching** [54]. In the robot teaching paradigm, rather than explicitly writing code, the operator interacts with the robot through natural (i.e., human and unlearned) communication channels to transfer knowledge to the robot. While doing so, the code, corresponding to the skills taught to the robot, is generated automatically behind the scenes. **Kinaesthetic programming-by-demonstration (KPbD)** is a key set of techniques developed to achieve this robot teaching goal. We bring to the MASTER project expertise in developing and studying such techniques, including contributions to the body of relevant prior work. In particular, we have contributed by developing multimodal and multi-channel interactions that merge, in the same collaborative application scenario, a real robot and its holographic digital twin (DT) [45,46].

In the MASTER project, we will advance the codeless programming of robot actions and routines through natural multimodal interaction with its holographic twin via manual guidance, voice, and gestures. Interactions with the robot to teach new skills will be assisted by a **virtual agent** designed and implemented in the project. Foundational operations for robot routine programming will be first supported as building blocks for the programming of collaborative production processes. These basic operations will support the programming of trajectories for a robot to approach its end effector to a region of interest. Then, support for the precise positioning of the end effector of a robot on selected points of interest is discovered in the region of interest that the robot previously approached. Precise positioning will be achieved with the help of machine vision algorithms that the operator will invoke through interaction with the virtual agent.

To teach approach trajectories, the KPbD technique will be used through interaction with the hologram of the robot. This will involve teaching robot trajectories, actions, and decision-making criteria, by moving the holographic digital twin robot by hand, either in a VR scene with respect to a virtual version of the production environment or in an MR environment within a real production context, which may include the physical version of the robot in addition to its holographic DT. In addition to our references cited above, KPbD in XR is present in the literature as visual MR implementations of the KPbD technique. Using a HoloLens device, pick-and-place action programming is described in [42,38]. To improve the perception of physicality (absent in the handling of a hologram), the authors provided in a previous project tactile feedback using ultrasound actuators, resulting in a visuotactile interaction that improved the user experience (UX) of handling a hologram [45]. In VR, a description of a controller-free implementation of similar pick-and-place interactions exists [32]. To implement precise positioning capabilities for the end effector, machine vision-assisted techniques will be used, such as a **look and move (L&M) technique**. With an onboard camera calibrated with its end effector, the robot visually inspects the region of interest from close-up positions. The user will engage with the virtual agent to perform detailed visual analysis using machine vision applications that can identify features of interest in the image captured by the robot. When programming approach and precise-positioning trajectories, XR technologies will allow the user to preview simulated visualizations of the executions of the trajectories. This will help the user decide about the proposed trajectories' correctness before adding them to the program. This can all be done through interaction with the holographic version of the robot. The operator will use these tools by interacting with the virtual agent through an interface in the XR device, which will also support voice interaction based on semantic technologies, for which we also bring in expertise (e.g., [33]). This agent will monitor the progress of the approach-and-position sequence and provide guidance and step-by-step instructions to help the user interpret intermediate results, decide on the next steps, and complete the programming task.

This module of the MASTER project will be developed for and validated with two highly relevant real-world production scenarios: a machine tending scenario and a line following scenario. The **machine tending scenario** will involve the robot picking unsorted parts and feeding them one by one in an orderly way to a specific point of the production process. For that, the robot must first perform an approach to the region of interest in which the parts to be picked are located. Then, using machine vision analysis, the robot must identify features of the parts that lead it to identify a grasping point. After that, the robot will grasp a part (completing the L&M routine) and feed it to the process. With this module, the operator will use KPbD and L&M functionalities and, by interacting with the interface of the virtual agent, generate a behavior tree as the executable program for the robot to perform the machine tending task. The **line following scenario** will reflect tasks in which the robot must follow lines with high precision, such as when applying welding along a line or when fixing a crack in the material. For that, using the KPbD functionality of the module, the robot will first approach the region containing the line's starting point that will be followed. Then, the robot will find the starting point of the line that it will follow. This will be done with machine vision image analysis features of the L&M functionality of the module, invoked through interaction with the virtual agent. This interaction will be used to define the criteria for the robot to follow the line while acting on it (welding, applying a chemical treatment, etc.). Criteria to determine that the end of the line has been reached will be determined in a similar way. The whole sequence of actions and all decision criteria will be added to the executable behavior tree.

2.3 Interaction based on Eye Tracking

In the MASTER project, we implement gaze-based interaction methods for XR and investigate their impact on educational XR interfaces for training workers in robotics. Our goals include reducing extraneous cognitive load [34, p. 39] induced by the XR-based education system to enable better learning and to integrate a reliable method for, e.g., observing a user's learning progress. We bring in our expertise in cyber-physical manufacturing systems [47], human-in-the-loop systems [7], and, particularly, active and passive gaze-based interfaces. In active gaze-based interfaces, users explicitly move their eyes to interact with a digital system, for instance, for object selection or manipulation. Passive gaze-based interfaces implicitly observe users' eye movements to adapt the interaction to their current state, such as cognitive load. In addition, we aim to develop novel interactive machine learning (IML) methods that facilitate the continuous improvement of user models or interaction methods in the MASTER project.

Eye trackers have evolved from pure diagnostic sensors to input devices for real-time interactive systems [16]. This was partially driven by advances in eye tracking hardware over the last decade concerning the devices' affordability, performance, and form factor. However, the characteristics of the human visual system, including the eyes, are essential to this development. Directing the gaze to a specific spot is very efficient, and the gaze direction implicitly reveals where a human allocates their limited attentional resources [22, p. 26]. This visual attention can be attracted by generally appealing, bottom-up factors and goal- or task-oriented, top-down factors [13]. Modern eye trackers can capture and stream a user's gaze directions in real-time, allowing developers to use the human gaze as an input to interactive systems. The gaze signal can be captured in stationary settings using remote eye trackers and in mobile settings using head-mounted eye trackers. Remote eye trackers have a fixed position in the world, usually attached to a screen, and allow little to no movement by a user. Head-mounted eye trackers are worn by the user like glasses and allow more freedom of movement. Eye trackers in MR glasses are head-mounted eye trackers. Still, interaction settings can be stationary if users do not move. From stationary to mobile interaction settings, gaze can be employed as an active input modality or as a passive sensor-based cue, which spans the design space for gaze-informed multimodal interaction introduced by [43, pp. 382-384]. Gaze-informed multimodal interaction directly relates to the theory of multimodal-multisensor interfaces, i.e., interfaces that "combine one or more user input modalities with sensor information" [39, p. 3].

Active Gaze-based Interfaces When the human gaze signal is used as an active input mode, i.e., in active gaze-based interfaces, a user's intentional eye movements directly impact the interaction. Examples include gaze-based selection [48] and text input [19]. We plan to integrate an active gaze-based interaction method that supports the educational use case. This can be achieved by transferring an existing interaction method to an MR setting or by optimizing an existing MR interaction technique based on eye tracking for the educational use case at hand. We aim to investigate the impact of using eye tracking versus no eye tracking in educational settings. We assume that using eye gaze for interacting with MR interfaces can reduce the required effort for the interaction and leaves more capacities for learning. The related cognitive load theory states that reducing extraneous cognitive load can lead to better learning outcomes [53,34]. Recent examples of active gaze-based interfaces in MR include the selection of virtual objects [31], the manipulation of virtual objects, i.e., scaling, rotating, and translating them [41], the interaction with virtual menus [44], and gaze-based text entry systems [30].

Passive Gaze-based Interfaces When the human gaze signal is used as a passive input mode, i.e., in passive gaze-based interfaces, naturally occurring eye movement behavior is monitored and can indirectly impact the interaction. Typically, the gaze signal is used to infer information about the user, such as the ongoing activity [28] or the induced cognitive load of an activity [24], to tailor the interaction to the user and their current state. This includes two stages, the detection of certain states and a suitable adaptation to them. Good results have been observed in estimating participants' expertise based on their eye movements [14]. Other researchers investigated whether the change in knowledge can be observed from eye movements [11] and whether user confusion can be detected [27]. In prior work, we investigated whether eye movements and fixated visual stimuli can be used to predict the target of a visual search [9] or the relevance of a text snippet read by a user [6]. Also, we implemented and evaluated methods for detecting what objects have been fixated based on head-mounted eye trackers [8]. First attempts have been made to adapt user interfaces based on a user's preferences determined through eye tracking. One example concerned the presentation of details for flat advertisements based on the perceived relevance per user [18]. Such monitoring methods can complement multimodal learning analytics tools, which aim at analyzing students' learning behavior and progress [12]. We aim to identify a user characteristic that can be reliably tracked by observing a user's eye movements in the context of teaching robotics via XR headsets. However, our experiences show modeling user states is connected to certain risks. Typically, it is hard to identify generic patterns that generalize across different users, either because we cannot collect sufficient data through user studies or because eye movement behavior is connected to individual experiences and varies significantly between users. Our goal in the MASTER project is to measure a user characteristic relevant to the educational use case with high validity and reliability. We can imagine showcasing our final method in an adaptive educational user interface that adapts to the state of this characteristic or in a dashboard for teachers that visualizes the characteristic for multiple learners.

Interactive Machine Learning (IML) In interactive machine learning systems, the user engages in a two-way interaction with a machine learning system. The goal is to improve the model based on user feedback [17]. Such human-in-the-loop machine learning systems are very relevant when transferring non-perfect machine learning systems to real use cases in the industry. Typically, the datasets for training are not large enough to reach a sufficient level of accuracy in predictions, or the task is very sensitive to wrong classifications and requires human quality assurance. We aim to apply the principle of interactive machine learning to machine learning models in the context of gaze-based interaction. For instance, passive gaze-based interfaces rely on models that take human eye movements as input to predict a certain user characteristic or to understand which elements of a virtual or real scene are viewed by a user in which order. Such models typically suffer from a low number of training samples because collecting and annotating data is expensive and time-consuming. We envision a system for interactive data exploration or interactive modeling of user characteristics based on recorded or real-time eye movement data. One concrete example is the semi-automatic annotation of mobile eye tracking data based on computer vision models [8,40,52,25]. Such models must be adjusted for every new context. IML can enable education experts with no prior machine learning experience to apply machine learning technology [5,26].

3 Didactic Approach

Educational technology is "the study and ethical practice of facilitating learning and improving performance by creating, using, and managing appropriate technological processes and resources" [10, p. 1]. The use of XR technology as an effective educational technology in diverse educational and training settings has been the subject of extensive research for many years. Alnagrat et al. [4] presented a related survey that highlights the potential benefits of XR for training and education, such as greater user engagement, lower (operational) costs, better knowledge acquisition and retention, training as often as needed, and training in safe environments. Recent advances and increased affordability of XR technologies have led to a proliferation of related educational resources. The major challenges remain in corporate training, technical and professional higher education, and, generally, where complex technical skills or high-stake learning goals are required. An example is teaching robotics in manufacturing, assisted by XR. Our vision in the MASTER project is that industrial workers, vocational and higher education students of robotics, robotics teachers and trainers, and other stakeholders in real-world education and training environments will be immersed in high-quality teaching and learning experiences based on XR technology.

To this end, the XR platform will extend the existing VIROO® platform with additional functionalities arising from the technological areas of the research partners. These include the creation of safe robotic environments, the programming of flexible robotic applications, and the integration of gaze-based interaction techniques. These technologies must be merged into an effective educational technology with the learning content from specific educational use cases. This shall promote the learning outcomes formulated in the targeted robotics domains in our open call projects. The educational system will be configured functionally in compliance with the principles of effective teaching for complex learning in technical contexts, as these have been operationalized in instructional models like the ones introduced in [36,35]. The didactic approach in the XR platform is to have problem classes and associated problem sequences related to robotics study targets, such as collaborative robot programming. This content and supporting information shall enable progressive and self-directed learning of students. It will be presented with pertinent immersive and realistic characteristics and required interactions. The system will be equipped with functionalities to manage the teaching and learning processes and facilitate the work of teachers and administrators, including capabilities for content management, student access management, and tracking students' activity and learning progression. In addition, the XR platform will also allow teachers, trainers, or instructional designers to plan new activity scenarios, teaching and learning sequences, and XR content. The XR platform and educational content will be conceived considering audiences' specific expectations, requirements, and feedback. They will be formatively evaluated during the second open call, as described next.

4 Evaluation via Open Calls

The MASTER project aims to develop and deliver a high-quality platform and training content to effectively support the project objectives. This platform will be supported by two open calls. The first one targets technical SMEs to integrate their XR developments into the platform, enhancing its capabilities. The second one shall attract educational institutes that can prepare training material using the above tools and test the proposed functionalities. A validation methodology will be proposed and iteratively improved during the project to achieve this. Through this experimental validation process, the project will focus on evaluating user acceptance, usability, and trainees' perceptions of safety. It shall provide the necessary feedback throughout the design and development phases of the XR platform and training material, ensuring continuous improvement and refinement. The first open call addresses technical challenges, supporting and funding satellite projects focusing on the technologies being developed. This way, a knowledge-sharing channel will be created while raising awareness among various companies and organizations regarding the project's advancements and potential. The technologies developed by the core partners of the MASTER project and those provided by the open call participants will be evaluated via experiments designed and implemented during the project's lifecycle. The second open call addresses educational challenges. The developed technologies and the created training content will be evaluated through empirical user studies. This approach allows an inclusive evaluation of the platform and the training content, considering the various perspectives and needs of our broad target audience. Findings from these studies will be used to improve the final platform and didactic materials through an iterative design process.

5 Conclusion

In this paper, we presented the EU-funded project MASTER. We summarized the aim of the project, which is to advance the application of innovative XR technologies in robotics education for manufacturing companies. Six objectives were defined, namely, four technical, one educational, and one industrial. The technical objectives include the development of an open XR platform that is extended by key functionalities for creating safe robotic environments, programming flexible robotic applications, and integrating advanced interaction techniques based on eye tracking. We also described the planned technological developments in detail. The educational objective addresses the creation of related learning material. In this context, we detailed our didactic considerations and the challenges connected to educating workers. The industrial objective addresses the support and the transfer of our technologies to SMEs and start-up companies. Further, we described the evaluation methodology. We planned two open calls: the first enables technology transfer and evaluation in realistic settings; the second focuses on the assessment of the educational material that will be created by open call beneficiaries. News and recent developments concerning the MASTER project will be published on our project website https://www.master-xr.eu.

Acknowledgements This work is co-funded by the European Union (EU) under grant number 101093079 (project MASTER, https://www.master-xr.eu). Thanks to Abdulrahman Mohamed Selim and Nathalie Barz for the very useful feedback on the initial draft and to the European Science Communication Institute (ESCI) gGmbH for the permission to use the image.

References

1. Enterprise VR made easy with VIROO® (2023), https://www.virtualwareco.com/viroo/
2. Unity Real-Time Development Platform | 3D, 2D, VR & AR Engine (2023), https://unity.com
3. Virtual Reality CAVE Solutions (2023), https://steantycip.com/vr-cave/
4. Alnagrat, A., Che Ismail, R., Syed Idrus, S.Z., Abdulhafith Alfaqi, R.M.: A Review of Extended Reality (XR) Technologies in the Future of Human Education: Current Trend and Future Opportunity. Journal of Human Centered Technology **1**(2), 81–96 (Aug 2022). https://doi.org/10.11113/humentech.v1n2.27
5. Barz, M., Bhatti, O.S., Alam, H.M.T., Nguyen, D.M.H., Sonntag, D.: Interactive Fixation-to-AOI Mapping for Mobile Eye Tracking Data Based on Few-Shot Image Classification. In: Companion Proceedings of the 28th International Conference on Intelligent User Interfaces. pp. 175–178. IUI '23 Companion, Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3581754.3584179, https://doi.org/10.1145/3581754.3584179, event-place: Sydney, NSW, Australia
6. Barz, M., Bhatti, O.S., Sonntag, D.: Implicit Estimation of Paragraph Relevance From Eye Movements. Frontiers in Computer Science **3** (2022). https://doi.org/10.3389/fcomp.2021.808507, https://www.frontiersin.org/article/10.3389/fcomp.2021.808507
7. Barz, M., Poller, P., Schneider, M., Zillner, S., Sonntag, D.: Human-in-the-Loop Control Processes in Gas Turbine Maintenance. In: Mařík, V., Wahlster, W., Strasser, T., Kadera, P. (eds.) Industrial Applications of Holonic and Multi-Agent Systems. pp. 255–268. Springer International Publishing, Cham (2017)
8. Barz, M., Sonntag, D.: Automatic Visual Attention Detection for Mobile Eye Tracking Using Pre-Trained Computer Vision Models and Human Gaze. Sensors **21**(12), 4143 (Jan 2021). https://doi.org/10.3390/s21124143, https://www.mdpi.com/1424-8220/21/12/4143, number: 12 Publisher: Multidisciplinary Digital Publishing Institute
9. Barz, M., Stauden, S., Sonntag, D.: Visual Search Target Inference in Natural Interaction Settings with Machine Learning. In: Bulling, A., Huckauf, A., Jain, E., Radach, R., Weiskopf, D. (eds.) ACM Symposium on Eye Tracking Research and Applications. pp. 1–8. ETRA '20 Full Papers, Association for Computing Machinery (2020). https://doi.org/10.1145/3379155.3391314, https://doi.org/10.1145/3379155.3391314
10. Betrus, A.K., Branch, R.M., Doughty, P., Molenda, M., Pearson, R., Persichitte, K.A., Rezabek, L.L., Robinson, R.S., Stoddard, C., Januszewski, A.: Definition. In: Januszewski, A., Molenda, M. (eds.) Educational Technology: A Definition with Commentary, p. 384. Routledge, 2nd edn. (2008), https://www.routledge.com/Educational-Technology-A-Definition-with-Commentary/Januszewski-Molenda/p/book/9780805858617
11. Bhattacharya, N., Gwizdka, J.: Relating Eye-Tracking Measures with Changes in Knowledge on Search Tasks. In: Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications. ETRA '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3204493.3204579, https://doi.org/10.1145/3204493.3204579, event-place: Warsaw, Poland
12. Blikstein, P.: Multimodal Learning Analytics. In: Proceedings of the Third International Conference on Learning Analytics and Knowledge. pp. 102–106. LAK '13, Association for Computing Machinery, New York, NY,

USA (2013). https://doi.org/10.1145/2460296.2460316, https://doi.org/10.1145/2460296.2460316, event-place: Leuven, Belgium
13. Borji, A., Itti, L.: State-of-the-Art in Visual Attention Modeling. IEEE Transactions on Pattern Analysis and Machine Intelligence **35**(1), 185–207 (2013). https://doi.org/10.1109/TPAMI.2012.89
14. Castner, N.J.: Gaze and visual scanpath features for data-driven expertise recognition in medical image inspection. Dissertation, Universität Tübingen (Dec 2020). https://doi.org/10.15496/publikation-51677, https://publikationen.uni-tuebingen.de/xmlui/handle/10900/110301, accepted: 2020-12-07T07:53:09Z ISBN: 9781742181882
15. Dimitropoulos, N., Togias, T., Michalos, G., Makris, S.: Framework enabling the design of Virtual Environments used for simulation of assembly operations. Procedia Manufacturing **51**, 571–576 (2020). https://doi.org/https://doi.org/10.1016/j.promfg.2020.10.080, https://www.sciencedirect.com/science/article/pii/S2351978920319375
16. Duchowski, A.T.: Gaze-based interaction: A 30 year retrospective. Computers & Graphics **73**, 59–69 (2018). https://doi.org/https://doi.org/10.1016/j.cag.2018.04.002, https://www.sciencedirect.com/science/article/pii/S0097849318300487
17. Dudley, J.J., Kristensson, P.O.: A Review of User Interface Design for Interactive Machine Learning. ACM Transactions on Interactive Intelligent Systems **8**(2) (Jun 2018). https://doi.org/10.1145/3185517, https://doi.org/10.1145/3185517, place: New York, NY, USA Publisher: Association for Computing Machinery
18. Feit, A.M., Vordemann, L., Park, S., Berube, C., Hilliges, O.: Detecting Relevance during Decision-Making from Eye Movements for UI Adaptation. In: ACM Symposium on Eye Tracking Research and Applications. ETRA '20 Full Papers, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3379155.3391321, https://doi.org/10.1145/3379155.3391321, event-place: Stuttgart, Germany
19. Feng, W., Zou, J., Kurauchi, A., Morimoto, C.H., Betke, M.: HGaze Typing: Head-Gesture Assisted Gaze Typing. In: ACM Symposium on Eye Tracking Research and Applications. ETRA '21 Full Papers, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3448017.3457379, https://doi.org/10.1145/3448017.3457379, event-place: Virtual Event, Germany
20. Gkournelos, C., Karagiannis, P., Kousi, N., Michalos, G., Koukas, S., Makris, S.: Application of Wearable Devices for Supporting Operators in Human-Robot Cooperative Assembly Tasks. Procedia CIRP **76**, 177–182 (2018). https://doi.org/https://doi.org/10.1016/j.procir.2018.01.019, https://www.sciencedirect.com/science/article/pii/S2212827118300325
21. Hietanen, A., Pieters, R., Lanz, M., Latokartano, J., Kämäräinen, J.K.: AR-based interaction for human-robot collaborative manufacturing. Robotics and Computer-Integrated Manufacturing **63**, 101891 (2020). https://doi.org/https://doi.org/10.1016/j.rcim.2019.101891, https://www.sciencedirect.com/science/article/pii/S0736584519307355
22. Holmqvist, K., Andersson, R.: Eye tracking: A comprehensive guide to methods, paradigms and measures. Lund Eye-Tracking Research Institute, Lund, Sweden (2011)
23. Karvouniari, A., Michalos, G., Dimitropoulos, N., Makris, S.: An approach for exoskeleton integration in manufacturing lines using Virtual Reality techniques. Procedia CIRP **78**, 103–108 (2018).

https://doi.org/https://doi.org/10.1016/j.procir.2018.08.315, https://www.sciencedirect.com/science/article/pii/S2212827118312289

24. Klingner, J.M.: Measuring Cognitive Load During Visual Tasks by Combining Pupillometry and Eye Tracking. Ph.D., Stanford University, Department of Computer Science (May 2010), http://purl.stanford.edu/mv271zd7591
25. Kumari, N., Ruf, V., Mukhametov, S., Schmidt, A., Kuhn, J., Küchemann, S.: Mobile Eye-Tracking Data Analysis Using Object Detection via YOLO v4. Sensors **21**(22) (2021). https://doi.org/10.3390/s21227668, https://www.mdpi.com/1424-8220/21/22/7668
26. Kurzhals, K.: Image-Based Projection Labeling for Mobile Eye Tracking. In: ACM Symposium on Eye Tracking Research and Applications. Association for Computing Machinery, New York, NY, USA (2021), https://doi.org/10.1145/3448017.3457382
27. Lallé, S., Toker, D., Conati, C.: Gaze-Driven Adaptive Interventions for Magazine-Style Narrative Visualizations. IEEE Transactions on Visualization and Computer Graphics **27**(6), 2941–2952 (2021). https://doi.org/10.1109/TVCG.2019.2958540
28. Li, Y., Liu, M., Rehg, J.M.: In the Eye of Beholder: Joint Learning of Gaze and Actions in First Person Video. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) Computer Vision – ECCV 2018. pp. 639–655. Springer International Publishing, Cham (2018)
29. Lotsaris, K., Fousekis, N., Koukas, S., Aivaliotis, S., Kousi, N., Michalos, G., Makris, S.: Augmented Reality (AR) based framework for supporting human workers in flexible manufacturing. Procedia CIRP **96**, 301–306 (2021). https://doi.org/https://doi.org/10.1016/j.procir.2021.01.091, https://www.sciencedirect.com/science/article/pii/S2212827121001190
30. Lystbæk, M.N., Pfeuffer, K., Grønbæk, J.E.S., Gellersen, H.: Exploring Gaze for Assisting Freehand Selection-Based Text Entry in AR. Proceedings of the ACM on Human-Computer Interaction **6**(ETRA) (May 2022). https://doi.org/10.1145/3530882, https://doi.org/10.1145/3530882, place: New York, NY, USA Publisher: Association for Computing Machinery
31. Lystbæk, M.N., Rosenberg, P., Pfeuffer, K., Grønbæk, J.E., Gellersen, H.: Gaze-Hand Alignment: Combining Eye Gaze and Mid-Air Pointing for Interacting with Menus in Augmented Reality. Proceedings of the ACM on Human-Computer Interaction **6**(ETRA) (May 2022). https://doi.org/10.1145/3530886, https://doi.org/10.1145/3530886, place: New York, NY, USA Publisher: Association for Computing Machinery
32. Masurovsky, A., Chojecki, P., Runde, D., Lafci, M., Przewozny, D., Gaebler, M.: Controller-Free Hand Tracking for Grab-and-Place Tasks in Immersive Virtual Reality: Design Elements and Their Empirical Study. Multimodal Technologies and Interaction **4**(4) (2020). https://doi.org/10.3390/mti4040091, https://www.mdpi.com/2414-4088/4/4/91
33. Maurtua, I., Fernández, I., Tellaeche, A., Kildal, J., Susperregi, L., Ibarguren, A., Sierra, B.: Natural multimodal communication for human–robot collaboration. International Journal of Advanced Robotic Systems **14**(4), 1729881417716043 (Jul 2017). https://doi.org/10.1177/1729881417716043, https://doi.org/10.1177/1729881417716043, publisher: SAGE Publications
34. Mayer, R.E. (ed.): The Cambridge Handbook of Multimedia Learning. Cambridge Handbooks in Psychology, Cambridge University Press, 2 edn. (2014). https://doi.org/10.1017/CBO9781139547369

35. van Merrienboer, J.: Training complex cognitive skills: A four-component instructional design model for technical training. Educational Technology Publications, Englewood Cliffs, NJ (1997)
36. van Merriënboer, J.J.G., Kirschner, P.A.: Ten Steps to Complex Learning: A Systematic Approach to Four-Component Instructional Design. Routledge, New York, 2nd edn. (Oct 2012). https://doi.org/10.4324/9780203096864, https://doi.org/10.4324/9780203096864
37. Michalos, G., Karvouniari, A., Dimitropoulos, N., Togias, T., Makris, S.: Workplace analysis and design using virtual reality techniques. CIRP Annals **67**(1), 141–144 (2018). https://doi.org/https://doi.org/10.1016/j.cirp.2018.04.120, https://www.sciencedirect.com/science/article/pii/S0007850618301446
38. Ostanin, M., Mikhel, S., Evlampiev, A., Skvortsova, V., Klimchik, A.: Human-robot interaction for robotic manipulator programming in Mixed Reality. In: 2020 IEEE International Conference on Robotics and Automation (ICRA). pp. 2805–2811 (2020). https://doi.org/10.1109/ICRA40945.2020.9196965
39. Oviatt, S., Schuller, B., Cohen, P.R., Sonntag, D., Potamianos, G., Krüger, A.: Introduction: Scope, Trends, and Paradigm Shift in the Field of Computer Interfaces. In: The Handbook of Multimodal-Multisensor Interfaces, pp. 1–15. Association for Computing Machinery and Morgan & Claypool, New York, NY, USA (2017). https://doi.org/10.1145/3015783.3015785, https://doi.org/10.1145/3015783.3015785
40. Panetta, K., Wan, Q., Kaszowska, A., Taylor, H.A., Agaian, S.: Software Architecture for Automating Cognitive Science Eye-Tracking Data Analysis and Object Annotation. IEEE Transactions on Human-Machine Systems **49**(3), 268–277 (2019). https://doi.org/10.1109/THMS.2019.2892919, https://ieeexplore.ieee.org/document/8633339/
41. Pfeuffer, K., Mayer, B., Mardanbegi, D., Gellersen, H.: Gaze + Pinch Interaction in Virtual Reality. In: Proceedings of the 5th Symposium on Spatial User Interaction. pp. 99–108. SUI '17, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3131277.3132180, https://doi.org/10.1145/3131277.3132180, event-place: Brighton, United Kingdom
42. Quintero, C.P., Li, S., Pan, M.K., Chan, W.P., Machiel Van der Loos, H., Croft, E.: Robot Programming Through Augmented Trajectories in Augmented Reality. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 1838–1844 (2018). https://doi.org/10.1109/IROS.2018.8593700
43. Qvarfordt, P.: Gaze-Informed Multimodal Interaction. In: The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1, pp. 365–402. Association for Computing Machinery and Morgan & Claypool (2017), https://doi.org/10.1145/3015783.3015794
44. Reiter, K., Pfeuffer, K., Esteves, A., Mittermeier, T., Alt, F.: Look & Turn: One-Handed and Expressive Menu Interaction by Gaze and Arm Turns in VR. In: 2022 Symposium on Eye Tracking Research and Applications. ETRA '22, Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3517031.3529233, https://doi.org/10.1145/3517031.3529233, event-place: Seattle, WA, USA
45. Rivera-Pinto, A., Kildal, J., Lazkano, E.: Multimodal Mixed Reality Impact on a Hand Guiding Task with a Holographic Cobot. Multimodal Technologies and Interaction **4**(4) (2020). https://doi.org/10.3390/mti4040078

46. Rivera-Pinto, A., Kildal, J., Lazkano, E.: Toward Programming a Collaborative Robot by Interacting with Its Digital Twin in a Mixed Reality Environment. International Journal of Human–Computer Interaction pp. 1–13 (Jun 2023). https://doi.org/10.1080/10447318.2023.2221599, https://doi.org/10.1080/10447318.2023.2221599, publisher: Taylor & Francis
47. Sonntag, D., Zillner, S., van der Smagt, P., Lörincz, A.: Overview of the CPS for Smart Factories Project: Deep Learning, Knowledge Acquisition, Anomaly Detection and Intelligent User Interfaces. In: Jeschke, S., Brecher, C., Song, H., Rawat, D.B. (eds.) Industrial Internet of Things: Cybermanufacturing Systems, pp. 487–504. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-42559-7_19, https://doi.org/10.1007/978-3-319-42559-7_19
48. Stellmach, S., Dachselt, R.: Look & Touch: Gaze-Supported Target Acquisition. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2981–2990. Association for Computing Machinery, New York, NY, USA (2012), https://doi.org/10.1145/2207676.2208709
49. Thees, M., Altmeyer, K., Kapp, S., Rexigel, E., Beil, F., Klein, P., Malone, S., Brünken, R., Kuhn, J.: Augmented Reality for Presenting Real-Time Data During Students' Laboratory Work: Comparing a Head-Mounted Display With a Separate Display. Frontiers in Psychology **13** (2022), https://www.frontiersin.org/articles/10.3389/fpsyg.2022.804742
50. Vigkos, A., Pauer, A., Bevacqua, D., Turturro, L., Kulesza, M.: XR and its potential for Europe. Tech. rep., ECORYS, Brussels (Apr 2021), https://xreuropepotential.com/assets/pdf/ecorys-xr-2021-report.pdf
51. Villani, V., Sabattini, L., Battilani, N., Fantuzzi, C.: Smartwatch-Enhanced Interaction with an Advanced Troubleshooting System for Industrial Machines. IFAC-PapersOnLine **49**(19), 277–282 (2016). https://doi.org/https://doi.org/10.1016/j.ifacol.2016.10.547, https://www.sciencedirect.com/science/article/pii/S2405896316321371
52. Wolf, J., Hess, S., Bachmann, D., Lohmeyer, Q., Meboldt, M.: Automating areas of interest analysis in mobile eye tracking experiments based on machine learning. Journal of Eye Movement Research **11**(6) (Dec 2018). https://doi.org/10.16910/jemr.11.6.6, https://bop.unibe.ch/JEMR/article/view/4415, section: Articles
53. Zhou, J., Yu, K., Chen, F., Wang, Y., Arshad, S.Z.: Multimodal Behavioral and Physiological Signals as Indicators of Cognitive Load. In: The Handbook of Multimodal-Multisensor Interfaces: Signal Processing, Architectures, and Detection of Emotion and Cognition - Volume 2, pp. 287–329. Association for Computing Machinery and Morgan & Claypool (2018), https://doi.org/10.1145/3107990.3108002
54. Zhou, Z., Xiong, R., Wang, Y., Zhang, J.: Advanced Robot Programming: a Review. Current Robotics Reports **1**(4), 251–258 (Dec 2020). https://doi.org/10.1007/s43154-020-00023-4, https://doi.org/10.1007/s43154-020-00023-4