AuctentionAR - Auctioning Off Visual Attention in Mixed Reality

Adrian Pandjaitan
adrian.pandjaitan@student.unisg.ch
University of St. Gallen
St.Gallen, Switzerland

Jannis Strecker
jannisrene.strecker@unisg.ch
University of St. Gallen
St.Gallen, Switzerland

Kenan Bektaş
kenan.bektas@unisg.ch
University of St. Gallen
St.Gallen, Switzerland

Simon Mayer
simon.mayer@unisg.ch
University of St. Gallen
St.Gallen, Switzerland

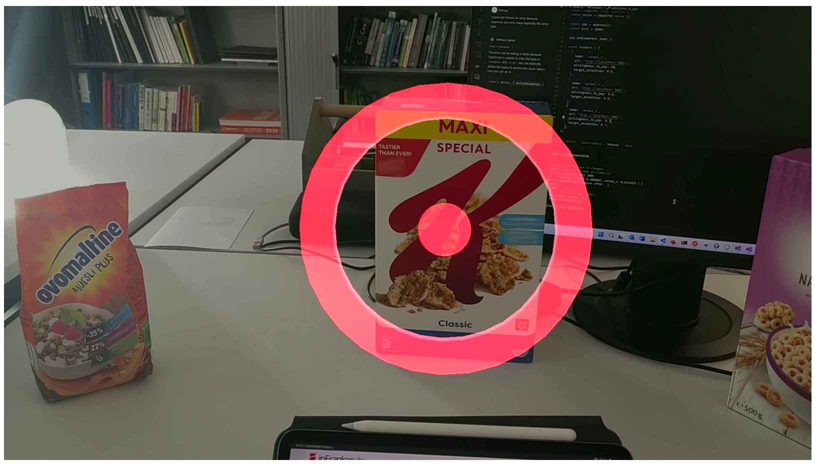

Figure 1: A visual attention marker of the AuctentionAR system highlights a product which has won the auction for our attention as viewed through a HoloLens 2. The red circle attempts to draw our visual attention to the cereal box.

ABSTRACT

Mixed Reality technologies are increasingly interwoven with our everyday lives. A variety of powerful Head Mounted Displays have recently entered consumer electronics markets, and more are under development, opening new dimensions for spatial computing. This development will likely not stop at the advertising industry either, as first forays into this area have already been made. We present AuctentionAR which allows users to sell off their visual attention to interested parties. It consists of a HoloLens 2, a remote server executing the auctioning logic, the YOLOv7 model for image recognition of products which may induce an advertising intent, and several bidders interested in advertising their products. As this system comes with substantial privacy implications, we discuss what needs to be considered in future implementation so as to make this system a basis for a privacy preserving MR advertising future.

CHI EA '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0331-7/24/05
https://doi.org/10.1145/3613905.3650941

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; **Ubiquitous and mobile computing systems and tools**; • **Computing methodologies** → **Computer vision**.

KEYWORDS

visual attention, notifications, auctioning, object detection

ACM Reference Format:

Adrian Pandjaitan, Jannis Strecker, Kenan Bektaş, and Simon Mayer. 2024. AuctentionAR - Auctioning Off Visual Attention in Mixed Reality. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3613905.3650941

1 INTRODUCTION

Recent developments in the domain of Mixed Reality (MR) have pushed this field again towards the center of public attention, not least due to the hype around the release of Apple's Vision Pro Head Mounted Display (HMD) which is said to be compatible to most third-party apps which are already existing for iOS devices. It is thought to make the development of such apps more attractive to external vendors as they can be used by a potentially large market. But not only in terms of interoperability has there been increased progress—also regarding models of technology acceptance, new generations of consumers exhibit high values concerning MR technologies [19, 24]. One of these models is the well-proven Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) [26]. As an extension of the original UTAUT model which itself was based on a synthesis of previous technology acceptance research [25, 26], it defines multiple dimensions which influence the intention behavior to use technology.

When examining current MR systems, we can argue that they score high in most of the model's dimensions: For instance, with new HMD models which have very user-friendly setup processes (e.g. apps on the Apple Vision Pro can easily be installed from its AppStore[1]) effort expectancy was lowered. Even though compared to other smart devices they are expensive, compared to the price of former HMDs such as Microsoft's HoloLenses, the utility per dollar (a combination between performance expectancy and price value in UTAUT2) which is offered by the technology can be regarded as much higher due to the above-mentioned factors. The Social influence dimension will also likely ameliorate as the form factor of these consumer HMDs gets ever closer to the one of everyday sunglasses (as showcased multiple times at the Consumer Electronics Show 2024 already[2]). Others might then not be as skeptical of an individual wearing such a device [25]. In conclusion, it can be said that many dimensions of acceptance of MR technology are expected to improve and consequently its diffusion will likely increase. Hence, this is why our research took interest in it.

1.1 Advertising in MR Settings

Having established this potential trajectory of wearable MR technology, it is not difficult to imagine that attached industries will emerge and grow in size (see [17]). In this paper, the focus lies on the advertising industry since it is easily imaginable that similar to the way the modern Web works, first- and third-party MR applications will try to find different ways of funding their endeavors and create "Hidden Revenue" business models [8]. The possibilities for engaging the end user are extended drastically in MR settings (see [23, 28]). This promises to make standing out easier as the creative freedom of advertisement creators is increased by allowing them to make use of the physical context an end user of the system is located in. As an example for this, the Hyper-Reality concept by Matsuda[3] could be taken, where interactions with physical objects are overlaid with custom tailored gamified interactive advertisements or a mobile application by the furniture store IKEA allowing users to place items from the catalogue virtually in their own homes before purchase [23]. On the other hand, also increased competition between advertisers can be expected. In contrast to the Web, where scrolling space is virtually unlimited, in an MR setting the physical space limits the positioning, size and amount of advertisements which can be displayed. Lastly, the topic of measuring advertisement impact is a recurring theme already in the established offline and online advertising industry, with traditional performance indicators such as Impressions or Click Through Rates being compared to newer ones like calculated attention from mouse movements [2].

[1]https://developer.apple.com/news/?id=ssmnoze2
[2]https://www.forbes.com/sites/charliefink/2024/01/14/all-the-xr-at-ces-2024/
[3]http://hyper-reality.co/

While it can be said that the advertisement industry can expect a range of new opportunities and challenges by adopting MR technology and the use of HMDs, the same goes for the end user in this regard. Similarly again to the Web of today, with the controversy around tracking cookies, MR advertising technology which might gather additional information such as gaze data about a much bigger part of a user's life can be viewed as even more intrusive. Privacy leakage through the misuse of sensitive spatial data has already been described in the context of mobile MR and MR HMDs [5]. Furthermore, it is not unlikely that in a scenario where MR advertisements have to heavily compete for the users' attention in limited space, they need to try to go "all-in" and possibly limit the users' agency by overwhelming, distracting and manipulating them regularly, employing dark patterns [7]. Examples for this are given for instance in [16], where the authors argue that large platforms might use the collected data to perform predatory pricing or even engage in "Reality Censorship". This feeling is apparently shared by people among the general population: Having seen the Hyper-Reality concept, one commenter of the video remarked that "[t]his Hyper-Reality need [sic] an Ad-Block app urgent [sic]"[4]; there is also research on physical world ad-blockers (see [21]).

1.2 Aligning the goals of advertisers and end users

It is visible that the goals of advertisers and end users in traditional advertising and MR settings alike diverge to a large extent. As was established above, due to the opportunities it opens up to advertisers, it is very unlikely, however, that MR will be ad-free forever and some advancements in this domain have already been made (see e.g. [6, 29]). While a framework for identifying manipulative MR advertising techniques has already been devised [15], rather than focusing on the problems, the objective of this paper is to find a way to tackle at least part of this problem. Hence we employ the economic principle of supply and demand to align the goals of the two parties to some extent through auctions. For this reason, after giving some context and drawing parallels to Web advertising, the prototype of an attention auctioning system for advertisers, called AuctentionAR is devised and its functionality as well as privacy implications and further research directions are explained below.

2 CONTEXT AND BACKGROUND

In the following, we explain the underlying assumptions and inspirations for our design and prototype. Furthermore, peculiarities of advertisement performance measurements are addressed as MR allows for new ways of implementing it.

2.1 Advertising in the modern Web as inspiration for MR

Many popular offerings in the Web, such as news platforms, search engines, video streaming services and social networks make use of advertising as their primary source of revenue. While there exist different methods of shaping such advertisement contracts between the platforms and the advertisers, it often boils down to auctioning systems, where advertisers bid on slots to have their tailored ads

[4]https://www.youtube.com/watch?v=YJg02ivYzSs

displayed to a user which belongs to a certain interest category [18]. These categories are created based on past user behavior to optimally match the advertisers with the end users and display relevant information to them. Success as the subject of the auctions is normally tied to Impressions or Click Through Rates [18] but other performance indicators such as attention start to gain traction, at least in the scientific discourse [2]. The benefit of auctioning systems such as generalized second price (GSP) lies in the property that it generally finds the advertiser's optimal willingness to pay [2, 18].

Even though the user rarely receives (a share of) the profits in such scenarios, building a business model around it seems to be worth pursuing as it has the possibility to give back agency to the user. For instance, just like with some real auctions, users could be allowed to set their own thresholds or reservation prices, which need to be surpassed to be shown an advertisement. This would provide them with the agency to themselves decide the trade-off of how much advertisements in return for a monetary or different remuneration they want to see.

Summarized, advertising in the modern Web has already very much matured and best practices have been devised to a certain extent. Thus, for the system in this paper, we will follow a similar approach and embed an auctioning mechanism for advertisements into it.

2.2 Predicting Attention in an MR Setting

As has been shown by [2], using attention as a predictor of advertisement impact and performance works just as well as using traditional indicators. In contrast to the method of mouse pointer tracking, predicting visual attention in MR settings with HMDs can be done by recording eye gaze tracking data [4]. This data can then be used to generate a highly precise understanding about which parts of a scene a user is paying attention to in particular. For this reason it might be of high relevance for advertisers and [3, 12] argue that visual attention can be used to infer a person's degree of interest as well as purchasing and consumption behavior. Our prototype will therefore make use of such a method, not least to account for the benefits MR advertising can have over traditional advertising.

2.3 Implications for MR Advertising Systems

We have established that MR technologies open up possibly very promising new kinds of user engagement for advertisers. User visual attention can easily be measured via the eye trackers which are a common sight in HMDs [15] as they are often needed to calibrate the displays correctly. Their prevalence removes the need for complex attention calculation models as it is the case in Web based advertising systems but needs special care when it comes to privacy issues as the collected data can reveal sensitive personal information easily [5, 12]. We thus strive for a user controlled system, where people are given back the agency about what they want to be shown and can define themselves what they want to receive in return.

3 THE AUCTENTIONAR PROTOTYPE

Drawing from the context above, we have created a system which implements these properties in a proof of concept. It consists of two main parts: First, we utilize a Microsoft HoloLens 2 (HL2) as an MR HMD which executes an application created with the Unity Game Engine to display the holographic user interface and record the user's gaze. This HL2 is connected to a NodeJS server running on a remote computer which provides attention auctioning as the main functionality of our system.

3.1 AuctentionAR Initialization

As illustrated in Fig. 2, our system works as follows: When started, the HL2 (1) connects to our back-end server (2) and starts its own HTTP listener. The server accesses the HL2's forward-facing camera video stream to detect objects (3) that are currently in the field of view of the user. For this task, the YOLOv7 [27] model (4) is used to determine the presence and, if applicable, the position of the objects it was trained on in a coordinate system relative to the camera's field of view. For our training data, we used different types and brands of cereal bags and boxes. The full dataset alongside the model weights can be found on GitHub.[5]. Our custom YOLOv7 model was trained on an NVIDIA Tesla T4 GPU and includes 35 classes of different grocery products. Its mean Average Precision score is 99.4% at an intersection over union threshold of 0.5 and 97.6% over the range [0.5, 0.95]. As soon as the object detection phase has concluded, the back end sends the detected objects' bounding box coordinates to the HL2. The HL2 next transforms the received screen coordinates into world coordinates by first compensating for the distance between the camera and the user's eyes and then doing a raycast from the average of the bounding box coordinates against the physical surface mesh which is automatically generated by the HL2. The 3D coordinates of the point of intersection are then saved and remain stable in the physical world. This means that the user can move their head and walk around and the HL2 itself compensates as much as possible for it. However, after some time a certain drift of the bounding box from the initially calculated point can occur, which would require a re-calibration of these coordinates. At the intersection of the ray and the surface mesh, a Unity GameObject with a collider which we call *gaze sensor* is then placed and linked with the specific type of cereal. The gaze sensors register when they are hit by the user's gaze ray which is calculated by doing a raycast from the 2D gaze coordinates provided by the eye tracker.

3.2 AuctentionAR Operation

During operation, our prototype's HL2 application starts collecting gaze information; we use this data as a proxy for attention. The application calculates a basic attention score by calculating the fraction of time that the user gazes at each gaze sensor over intervals of 2 seconds. After each interval, a message with all attention scores of the different objects is sent to the remote server alongside the request to start an attention auction (5). Agents (6), as defined by [14], can now bid on receiving the attention of the user. This is done via an auction where the registered agents can signal their willingness to pay for a specific object to be gazed at. Since a shift in the user's attention is likely correlated with increased mental

[5]https://github.com/Interactions-HSG/Datasets-And-Weights-For-Yolo

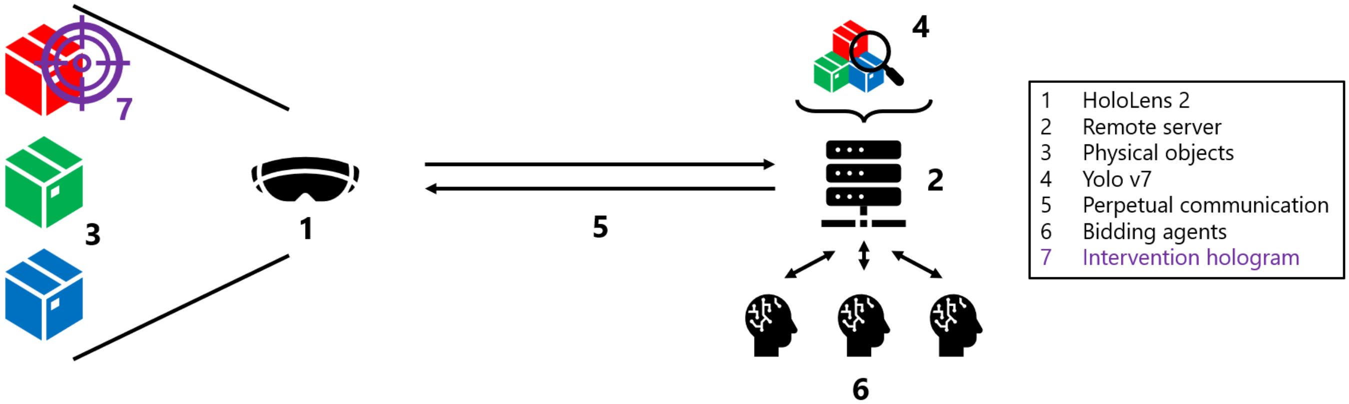

Figure 2: The Structure of the AuctentionAR System

load due to the disturbance, we divide the offered bid by $(1 - AttentionScore_{Objecti})$ for each object, which yields the price per unit of attention *shift*. The winner of the auction is then selected as the bidder with the highest willingness to pay for the attention shift. After the auction has been won, the server sends another HTTP request to the HL2's endpoint to notify it of the attention shifts; The HL2 then places the intervention (7)—a marker consisting of two concentric circles, forming a target—at the previously calculated coordinates of the designated object and starts an emphasizing animation to smoothly pulse the marker, oscillating it between 1-1.15 times its size for half a second. After the intervention has been concluded—regardless of whether the object has actually gotten the desired attention in our prototype—the marker vanishes and the main cycle starts again and repeats until the user closes the application.

4 DISCUSSION

Having created and preliminary presented our prototype to fellow researchers, we received feedback that for some people it looked rather dystopian. They had the feeling that companies having access to their gaze data would allow them to intrude into their privacy even more than they already were. Furthermore, in previous research it has been argued that far reaching information can be deducted by analyzing eye gaze data (e.g., gender, age, sexual preferences) which might lead to further privacy concerns [12]. Thus, in this section we discuss the implications for personal privacy of our system, provide suggestions of how the potential negative effects can be alleviated and show where our system or a technical system in general might not be the only solution.

4.1 Implications for Privacy

To discuss the implications on personal privacy of our prototype, we need to view the two main components of our system separately. First, there is the video stream and object recognition part. Here, the privacy concerns largely depend on the implementation and especially on where the data is processed in detail. For our prototype, we transmitted the video stream from the HL2's camera directly to a server under our control, to process the information and to do the classification without the need to use third party hardware or software which we could not independently check (i.e., cloud services). However, given the open concept of the system, there is virtually no limitation to where the processing and classifications can take place. Moving forward, we see that this could thus sway into two directions. Either the information processing is further outsourced to third party providers where users are at high risk of losing control over their own data or, on the other hand, the information processing is done directly on the device, eliminating the need for data transmission and a dedicated server application. At least with the HL2 as HMD, 3D bounding box alignment of objects has already been done on device [10], but actual object recognition, e.g., to determine which brand our cereal boxes belong to, remains difficult. This is due to the limited power mobile devices and especially HMDs such as the HL2 still possess, making it hard to run the required convolutional neural networks on them [11, 13]. Since an on-device object recognition approach might, however, be better suited for casual users of such HMDs who become an ever more important target group of HMD manufacturers, optimizing local object recognition models might be necessary to make possible increased levels of privacy. The average consumer most likely cannot afford to allot much time or budget to set up their own processing servers, for instance inside the walls of their own home. Moreover, manufacturers who are launching ever more powerful HMDs also need to open up their systems enough to make them compatible with such on-device applications.

The second part of our system, the auctioning, is different from the first part in the way that it necessitates communication and to some extent also the sharing of private information with third parties as they are part of a transaction of attention against a currency and want to measure their impact, as discussed above. However, also here there are strategies to minimize the data transmitted while still achieving acceptable results for all involved parties. For example, the auctioning process can be highly decentralized, creating a "marketplace" for each user and hence removing the need for a central all-knowing entity. This could, for instance, be achieved by employing systems such as SOLID pods[6], where users have full

[6] https://solidproject.org/

control over their own private and possibly local data containers (cf. [3, 9]). There they could then set their preferences, e.g., regarding who has access to their data (in our case: who can bid) themselves easily.

4.2 Limitations of Our Prototype System

Apart from the missing performance measurement for advertisers (this is something we have planned for future iterations of our system, where we will evaluate different options such as [1] or [20]), only one type of intervention, the red circle, has been examined. What is more, our system does also not exhibit all of the privacy characteristics highlighted in the previous section. In its current version it should thus only be used as a basis for future systems which have implemented more rigid privacy controls. Furthermore, while we believe that there would be a high acceptance of our system in a privacy preserving finalized form by end users, we believe that the same might not be true for advertisers. Even though, in general they would take direct damage from such a solution (i.e., bidding generally leads also to an optimal outcome for them), it is easy to see that they might not want to give up power over the user data they could collect easily. Thus, aside from our system as a technical solution, regulatory measures should be investigated as MR advertising technology progresses. This could for example mean that the HMDs of gate keeper manufacturers need to be open enough regarding their software (such as operating systems) and/or hardware to make it possible to independently assess their adherence to privacy legislation. First steps in this direction are already taken in some regions, for example with recently enacted Digital Markets Act in the European Union[7].

5 FUTURE RESEARCH DIRECTIONS

When testing the system in our research group, it became apparent that the system works reasonably well in extracting contextual information from the physical surroundings of the user. The auctioning system as well as the intervention placement also led to the reported change in attention. However, it was noticed that the interventions, especially when placed in intervals of two seconds, are too distracting and would need to be fine-tuned further. Moreover, the attention score calculation based on gaze time and using it as a proxy for actual attention could be improved by including additional variables such as the activity a user is currently pursuing. What is more, while the image recognition and placement of the gaze sensors were done on the remote external server, latency could probably be improved, by executing parts of the process such as bounding box detection directly on the HL2 (see [10]).

In future research, aside from a user study to find out system usability and actual acceptance, variations in the type of interventions could also be examined (see [22]). Finally, even though advertising is a large area of application for this technology, we believe that it is beneficial to look into other fields (outside of advertising) where objects (or their owners/producers/associates) might have a certain willingness to compete for user attention. This could for instance be devices in a smart home system where the goal of the user is to save energy: To prevent the user from getting overwhelmed with many energy-saving hints from all devices at once, our prototype could be employed together with a fixed amount of a fictional currency per device. In an auction, the devices could then bid as much of their currency as they believe their energy savings are worth. The proceedings could then be distributed to the user in the sense of gamifying their energy saving efforts or to the other devices. Another possible application scenario might be online or offline grocery shopping, where third parties (e.g., health insure companies) might have a willingness to pay for shifting a shopper's attention to healthier products. It should be noted, however, that restricting the the allocation of attention using an auctioning system might not be a sensible solution for high-risk applications. In an airplane cockpit, for example, it would not make sense to have different instruments bid for the pilot's attention. If there is an emergency, every useful piece of information needs to be relayed. Applications which could be reasonably devised, thus, exist rather in areas like the above-mentioned energy saving smart homes and similar low-risk domains.

6 CONCLUSION

Advertising in MR environments is hardly stoppable with the increased acceptance and distribution of capable HMDs. We have developed a prototype of a system which uses auctioning mechanisms similar to the Web to sell attention to the highest bidder. Our system alone does not guarantee privacy but needs to be taken as a basis to be combined with other privacy preserving systems such as SOLID pods. Thus our contribution is concerned mostly with finding a way that does not require an all-knowing central authority (as auctions can run local and be decentralized), illustrate privacy concerns about such systems and indicate possible future research directions which can build upon our system to ensure MR privacy.

REFERENCES

[1] Antti Ajanki, Mark Billinghurst, Hannes Gamper, Toni Järvenpää, Melih Kandemir, Samuel Kaski, Markus Koskela, Mikko Kurimo, Jorma Laaksonen, Kai Puolamäki, Teemu Ruokolainen, and Timo Tossavainen. 2011. An augmented reality interface to contextual information. *Virtual Reality* 15 (6 2011), 161–173. Issue 2-3. https://doi.org/10.1007/S10055-010-0183-5/FIGURES/5
[2] Ioannis Arapakis, Antonio Penta, Hideo Joho, and Luis A. Leiva. 2020. A Price-per-attention Auction Scheme Using Mouse Cursor Information. *ACM Transactions on Information Systems (TOIS)* 38 (1 2020), 13. Issue 2. https://doi.org/10.1145/3374210
[3] Kenan Bektaş, Jannis Strecker, Simon Mayer, and Kimberly Garcia. 2024. Gaze-enabled activity recognition for augmented reality feedback. *Computers & Graphics* (2024), 103909. https://doi.org/10.1016/j.cag.2024.103909
[4] Eunji Chong, Nataniel Ruiz, Yongxin Wang, Yun Zhang, Agata Rozga, and James M. Rehg. 2018. Connecting Gaze, Scene, and Attention: Generalized Attention Estimation via Joint Modeling of Gaze and Scene Saliency. , 383-398 pages.
[5] Jaybie A. de Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. 2019. A First Look into Privacy Leakage in 3D Mixed Reality Data. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11735 LNCS (2019), 149–169. https://doi.org/10.1007/978-3-030-29959-0_8/FIGURES/11
[6] Ko de Ruyter, Jonas Heller, Tim Hilken, Matthew Chylinski, Debbie I. Keeling, and Dominik Mahr. 2020. Seeing with the Customer's Eye: Exploring the Challenges and Opportunities of AR Advertising. *Journal of Advertising* 49 (3 2020), 109–124. Issue 2. https://doi.org/10.1080/00913367.2020.1740123
[7] Chloe Eghtebas, Gudrun Klinker, Susanne Boll, and Marion Koelle. 2023. Co-Speculating on Dark Scenarios and Unintended Consequences of a Ubiquitous(ly) Augmented Reality. (7 2023), 2392–2407. https://doi.org/10.1145/3563657.3596073
[8] Oliver Gassmann, Karolin Frankenberger, and Michaela Choudury. 2020. *The Business Model Navigator.* PEARSON EDUCATION LIMITED.
[9] Jan Grau, Simon Mayer, Jannis Strecker, Kimberly Garcia, and Kenan Bektaş. 2024. Gaze-based Opportunistic Privacy-preserving Human-Agent Collaboration.

[7] https://digital-markets-act.ec.europa.eu/about-dma_en

In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3613905.3651066

[10] Yongjie Guan, Xueyu Hou, Nan Wu, Bo Han, and Tao Han. 2022. DeepMix: Mobility-aware, Lightweight, and Hybrid 3D Object Detection for Headsets. *MobiSys 2022 - Proceedings of the 2022 20th Annual International Conference on Mobile Systems, Applications and Services* (6 2022), 28–41. https://doi.org/10.1145/3498361.3538945

[11] Adam Ibrahim, John B Lanier, Brandon Huynh, John O'Donovan, and Tobias Höllerer. 2017. Real-time Object Recognition on the Microsoft HoloLens. *technical report, UC Santa Barbara* (2017).

[12] Jacob Leon Kröger, Otto Hans Martin Lutz, and Florian Müller. 2020. What does your gaze reveal about you? on the privacy implications of eye tracking. *IFIP Advances in Information and Communication Technology* 576 LNCS (2020), 226–241. https://doi.org/10.1007/978-3-030-42504-3_15/FIGURES/1

[13] Luyang Liu, Hongyu Li, and Marco Gruteser. 2019. Edge assisted real-time object detection for mobile augmented reality. *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM* 19 (8 2019). https://doi.org/10.1145/3300061.3300116

[14] Pattie Maes. 1995. Artificial life meets entertainment. *Commun. ACM* 38 (11 1995), 108–114. Issue 11. https://doi.org/10.1145/219717.219808

[15] Abraham Mhaidli and Florian Schaub. 2021. Identifying manipulative advertising techniques in xr through scenario construction. *Conference on Human Factors in Computing Systems - Proceedings* (5 2021). https://doi.org/10.1145/3411764.3445253

[16] Joseph O'Hagan, Jan Gugenheimer, Jolie Bonner, Florian Mathis, and Mark McGill. 2023. Augmenting People, Places & Media: The Societal Harms Posed by Everyday Augmented Reality, and the Case for Perceptual Human Rights. *ACM International Conference Proceeding Series* (12 2023), 219–229. https://doi.org/10.1145/3626705.3627782

[17] Jason Orlosky, Misha Sra, Kenan Bektaş, Huaishu Peng, Jeeeun Kim, Nataliya Kos'myna, Tobias Höllerer, Anthony Steed, Kiyoshi Kiyokawa, and Kaan Akşit. 2021. Telelife: The Future of Remote Living. *Frontiers in Virtual Reality* 2 (11 2021), 763340. https://doi.org/10.3389/FRVIR.2021.763340/BIBTEX

[18] Michael Ostrovsky, Michael Schwarz, and Benjamin G. Edelman. 2005. Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords. https://papers.ssrn.com/abstract=847037

[19] José M. Ponzoa, Andrés Gómez, Silvia Villaverde, and Vicente Díaz. 2021. Technologically empowered? perception and acceptance of AR glasses and 3D printers in new generations of consumers. *Technological Forecasting and Social Change* 173 (12 2021), 121166. https://doi.org/10.1016/J.TECHFORE.2021.121166

[20] Pernilla Qvarfordt and Shumin Zhai. 2005. Conversing with the user based on eye-gaze patterns. *Conference on Human Factors in Computing Systems - Proceedings* (2005), 221–230. https://doi.org/10.1145/1054972.1055004

[21] Philipp A. Rauschnabel. 2021. Augmented reality is eating the real-world! The substitution of physical products by holograms. *International Journal of Information Management* 57 (4 2021), 102279. https://doi.org/10.1016/J.IJINFOMGT.2020.102279

[22] Sara Romano, Enricoandrea Laviola, Michele Gattullo, Michele Fiorentino, and Antonio Emmanuele Uva. 2023. More Arrows in the Quiver: Investigating the Use of Auxiliary Models to Localize In-View Components with Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 29 (11 2023), 4483–4493. Issue 11. https://doi.org/10.1109/TVCG.2023.3320229

[23] Joachim Scholz and Andrew N. Smith. 2016. Augmented reality: Designing immersive experiences that maximize consumer engagement. *Business Horizons* 59 (3 2016), 149–161. Issue 2. https://doi.org/10.1016/J.BUSHOR.2015.10.003

[24] Kuttimani Tamilmani, Nripendra P. Rana, and Yogesh K. Dwivedi. 2021. Consumer Acceptance and Use of Information Technology: A Meta-Analytic Evaluation of UTAUT2. *Information Systems Frontiers* 23 (8 2021), 987–1005. Issue 4. https://doi.org/10.1007/S10796-020-10007-6/FIGURES/4

[25] Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. 2003. User acceptance of information technology: Toward a unified view. *MIS Quarterly: Management Information Systems* 27 (2003), 425–478. Issue 3. https://doi.org/10.2307/30036540

[26] Viswanath Venkatesh, James Y.L. Thong, and Xin Xu. 2012. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly: Management Information Systems* 36 (2012), 157–178. Issue 1. https://doi.org/10.2307/41410412

[27] Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. 2022. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. (7 2022), 7464–7475. https://doi.org/10.1109/cvpr52729.2023.00721

[28] Michel Wedel, Enrique Bigné, and Jie Zhang. 2020. Virtual and augmented reality: Advancing research in consumer marketing. *International Journal of Research in Marketing* 37 (9 2020), 443–465. Issue 3. https://doi.org/10.1016/J.IJRESMAR.2020.04.004

[29] Yu You, Alain Boyer, Tero Jokela, and Petri Piippo. 2017. SelfieWall: A Mixed Reality Advertising Platform. *Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017* (10 2017), 240–244. https://doi.org/10.1109/ISMAR-ADJUNCT.2017.77