A Cognitive Immersive Room for Intelligence Analyst Scenarios (CIRIAS)

Shannon Briggs¹, Bryan Burns², Samuel Chabot¹, Erfan Al-Hossami³, Collen Roller², Abraham Sanders⁴, Tomek Strzalkowski⁴, and Jonas Braasch¹

¹EMPAC Research, Rensselaer Polytechnic Institute, Troy, NY 12180 USA

²Air Force Research Laboratory, Rome, NY, 13441 USA

³Department of Computer Science, University of North Carolina at Charlotte, Charlotte, NC 28223 USA

⁴Department of Computer Science, Rensselaer Polytechnic Institute, Troy NY 12180 USA

ABSTRACT

This paper describes the CIRIAS system, a cognitive, immersive suite of technologies that are designed to enable sensemaking for intelligence analysis. We describe the current technology array in the immersive room, past use cases of the CIRIAS, and further describe our current use case of the immersive environment. We are currently enabling a use case for Air Force intelligence analysts, by integrating a dataset concerning NYC cab traffic analysis and anomalous behaviors. For future work, we will study the effect of the immersive room on sensemaking metrics, such as information accuracy, retention, absorption, and collaborative behavior.

Keywords: Human systems integration, Immersive technology, Sensemaking, Virtual environments

INTRODUCTION

Intelligence can be understood as the timely delivery of actionable information. Our Cognitive Immersive Room for Intelligence Analysis Scenarios (CIRIAS) supports foraging and processing information during time-critical scenarios. Intelligence has an ambiguous meaning and could either refer to the ability to learn and reason well using a logical approach or to use a standard procedure to gather and process public and secret information about an adverse entity to forecast threats and opportunities. While the latter definition of intelligence roots in military operations, similar methods have been successfully applied in the civil domain, for example, forensic sciences and corporate business decisions. In this paper, we describe the use of cognitive immersive environments for collaborative decision-making using the general procedures of intelligence analysis, especially the concept of the foraging loop by Pirolli and Card (Pirolli and Card, 2005). We focus on a traffic-pattern analysis with NYC cabs to explain the benefit of virtual environments for the efficient and time-constrained decision-making process. By leveraging the cognitive immersive technology, we will transfer some of the granular search and sort activities to the system, reducing the cognitive load experienced by users during intelligence tasks. The progressive dialog system paired with our map views allows users to plan points of interest across travel itineraries and allows users to plan routes during challenging traffic. Our brainstorming tool (Briggs *et al.*, 2020) supports text source discovery, allowing users to build a knowledge base, and marshal evidence.

This approach aids analysis in reducing time and effort; timely analysis is typically critical in reconnaissance and other intelligence analyst tasks. During collection and analysis, information has to be pulled from various sources and shared among an expert team. CIRIAS possesses matured technologies to source information through personal interfaces such as computer terminals, handheld devices, and dialog systems while also allowing interfacing between groups of people.

The latter is important within the shared context between analysts to allow sharing the most relevant information while deferring other information. To bridge this technology gap, we propose a Situations Room environment that enables small teams to pursue intelligence analysis tasks together. In this cognitive immersive environment, each member can gather information individually while also exchanging and displaying relevant data on large, immersive displays. The use case we are developing is based on NYC cab traffic pattern data (He *et al.*, 2022); we anticipate this use case will be generalizable to our expert user domain, Air Force analysts¹

CIRIAS: CONCEPT AND IMPLEMENTATION

Immersive Room

The theoretical underpinning of the CIRIAS system is derived from Pirolli and Card's notational sensemaking model (Pirolli and Card, 2005). In this model, Pirolli and Card describe a 16-step recursive sensemaking process that analysts undergo during the analysis process. These steps are incorporated into two major loops, the information foraging loop and the sensemaking loop. We integrated this model to inform the system of the users' cognitive schema, supporting users during different activities that occur during discrete steps of the sensemaking and information foraging processes.

We also are interested in examining domain experts' sensemaking activities in the cognitive immersive room. We have designed and implemented domain users' cognitive tools, known as structured task analyses, within the cognitive, immersive technologies. The CIRIAS system can also be presented on a variety of displays; users can interact with the system through the global view, as can be seen on Fig. 1, which is presented on the 14' 360 degree screen. Large displays, such as the screen in the cognitive immersive room, increase sensemaking by enabling information foraging and schema generation through physical navigation, as opposed to forcing users to disrupt their

¹Co-author Bryan Burns is a Technical Sergeant (TSgt) and works as an intelligence analyst for aerial operations at the Air Force Base in Rome, NY. He is the domain expert on this project and is interested in new technologies to improve the workflow for intelligence analysts to make the process faster and more reliable. He ensures that the civilian use cases presented in this paper are relevant for the useful design of CIRIAS for the daily work of intelligence analysts.



Figure 1: CIRIAS overhead view and ground floor view with users.

tasks to switch applications and rely on memory, thereby offloading cognitive effort from the user (Endert *et al.*, 2014). A personal view is also enabled, which can be accessed via any web-enabled personal device, such as tablet, smartphone, or laptop. The CIRIAS system is currently composed of a data visualization tool, a map tool, and a brainstorming tool, which is informed by a structured analytic technique (Beebe and Pherson, 2014). Interaction with the system is accomplished with gesture and verbal commands, where users' body frames are registered via Kinect systems, and lapel microphones that allow users to issue commands. Gestural recognition is enabled via Microsoft Kinects, and verbal interaction through the Watson Assistant AI, as well as a conversational progressive dialog system. We leverage the array of sensors and displays to aid users in sensemaking. The progressive dialog system acts as a recommender system, lessens the amount of time spent in information foraging, and allows users to reach decisions more efficiently and more quickly.

Prior use cases in the sensemaking room involve the Shahikot Valley, which describes an intelligence failure that could be corrected through use of the CIRIAS, as well as the Jonathan Luna use case (Beebe and Pherson, 2014). This latter use case was enabled for a user study conducted in the cognitive immersive room. This user study examined the effect the digital brainstorming tool had on sensemaking behaviors. We discovered that the digital brainstorming, and their note and category contents were more closely representative of the sample text than the analog brainstorming tool (Briggs *et al.*, 2020, 2021, 2022).

Technical Background and Implementation

The goal of this research is to develop a modular system to support intelligence analysts. Intelligence analysis is typically a multi-expert process where different domain experts work together. AI-based dialog technologies, however, do not differentiate between individual users but rather treat the group on a holistic level where it does not matter who makes what input or contribution. Solutions exist in systems where analysts are networked together on the internet through individual terminals (e.g., the COLLANE project (Strzalkowski *et al.*, 2009)), but matters are much more difficult when the domain experts are physically present in the same room, which is often the case in operation centers. In this paper, we present a solution based on a panoramic screen system that consists of a spherical microphone system and multi-camera tracking to localize and receive input from individual team members. The Collaborative-Research Augmented Immersive Virtual Environment Laboratory (CRAIVE-Lab), shown in Fig. 2 addresses the need for a specialized virtual-reality system for the study and enabling of communication-driven tasks with groups of users immersed in a high-fidelity multimodal environment located in the same physical space. The CRAIVE-Lab was erected with the support of the National Science Foundation (Sharma, Braasch and Radke, 2017). The underlying technical infrastructure, shown in Fig. 2, is a collaborative, immersive system that focuses on intelligence analysts' cognitive processes to facilitate decision-making during analysis. We define collaborative, immersive cognitive systems as environments where: (i) groups of humans can communicate naturally with each other without obstruction from wearable devices (collaborative), (ii) groups of people are embedded in a panoramic display with surround sound capabilities (immersive), and (iii) users can draw from intelligent computing capabilities (cognitive systems).

Our front-projection video design, produced with eight projectors, differs substantially from the traditional Cave Automatic Virtual Environment (CAVE) back-projection design(Cruz-Neira, Sandin and DeFanti, 1993; DeFanti *et al.*, 2009; Wefers *et al.*, 2015).

Unlike the latter or head-mounted VR goggles, front-projection video systems do not lead to unnecessary fatigue. In contrast, it has been shown for VR goggles that even when watching a 3D movie, about 55% of the viewers complain about cybersickness side effects (Solimini, 2013).

For the acoustic domain, a 134-loudspeaker-channel system has been designed and installed for Wave Field Synthesis (WFS) with the support of Higher-Order-Ambisonic (HoA) sound projection to render inhomogeneous acoustic fields (Braasch, Peters and Valente, 2008; Chabot and Braasch, 2017; Agrawal and Braasch, 2018) (Fig. 2, left box in the second row from the top).

Addressing the dependencies of the different system components is not unique to our system approach. For example, changes in one presentation

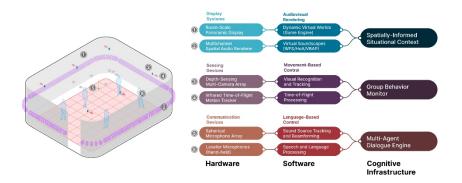


Figure 2: Domains of human systems integration. (Adapted from U.S Air Force, 2005.)

modality also affect the other modalities. This was solved by using A microperforated screen is used to preserve acoustics, and allow other acoustical environments to be created virtually through the loudspeaker system. An intelligent lighting system illuminates the participants without degrading the project video significantly, drawing users into the virtual world. The optimal lighting scenario will be context-dependent, in our technical solution is achieved by a DMX lighting system.

The audio-visual system can be connected through a dynamic world model, such as one provided in a game engine. This way, congruent audio-visual scenes can be created. In addition, the video display system can be used as a giant computer desktop with a resolution of 11636×1200 square pixels. The goal here is twofold: (1) to create immersive audio-visual imagery presenting physical spaces of interest to solve the intelligence analysis problem and (2) to display symbolic data, often in dense spatial form, to augment (1). Commercial software is used to remove the projection distortion and provide a seamless overlap between the images rendered by the individual projectors (Chabot and Braasch, 2017).

An intelligent position-tracking system estimates current user locations and head orientations as well as positioning data for other objects. For the tracking system, a hybrid visual/acoustic sensor system is being used to emulate the humans' ability to extract robust information by relying simultaneously on different modalities.

The tracking system uses a 16-channel spherical microphone array in conjunction with a sparse iterative beamforming algorithm to enable low-latency estimation of acoustical sound sources at a low computational cost. Our Sparse Iterative Search (SIS) method starts out by analyzing the spherical area with equal-area grid size. The algorithm computes the received energy in each area and contracts the analyzed area and grid size based on the received energy. Using an iterative approach, the algorithm zones into the energyemitting sound sources with high accuracy due to the small grid size. The algorithm can track up to four simultaneous static or moving sound sources.

Further, a network of six video cameras has been installed in the ceiling of the CRAIVE-Lab to provide a seamless top-view picture of the floor area (Sharma, Braasch and Radke, 2017). A time-of-flight (TOF) sensor array using Microsoft Kinects has been installed in the same ceiling grid (Divekar *et al.*, 2019). A sensor-fusion system combines the audio input from the spherical microphone and the visual/graphical input from the camera and TOF arrays. For this purpose, the fusion system provides a coordinate transformation for the microphone array from polar into Cartesian coordinates to match the output of the camera/TOF systems. The fusion system also weighs the input according to saliency in case the tracking cues conflict with each other. Using a multimodal tracking system also utilizes the circumstance that all cues are not available all the time, for example, when participants are quiet or when the DMX lighting is turned off while a movie is being projected on the screen.

CIRIAS also makes use of panoramic imagery for the immersion of participants in various realities. Crowd-sourced repositories such as Google Street View imagery are harnessed to provide content rapidly. This dynamic immersion has been utilized for a variety of applications, from decision-making to education. Students of language are able to travel to foreign countries to experience language and culture elements (Chabot *et al.*, 2020); students of architecture can explore potential project sites at human-scale (Elder, 2017); and students of acoustics can pair auralizations of performance spaces with their respective visuals (Chabot and Braasch, 2022).

To facilitate rapid collaboration, a Pin-Up platform has been developed which provides users of the environments with a browser-based, deviceagnostic interface for contributing content to the immersive screens. Users are presented with a familiar web page for uploading content which posts directly to the screen. This content covers a wide variety of commonly used media types including images, videos, audio, PDFs, websites, Street View panorama, and YouTube videos. Once contributed to the screen, these media can be positioned and presented on, and even maximized to encompass the entirety of the display. This application platform has been utilized in courses of acoustics, architecture, physics, and others. For decision-makers, this platform enables multiple collaborators to make use of a single canvas simultaneously.

The ATICA system is a dialogue system created to assist users to complete domain-specific tasks. The initial system incorporates the framework of the Schema Guided Dialogue (Rastogi *et al.*, 2020), and uses a unified multi-task neural language model (Raffel *et al.*, 2019) for its Dialogue State Tracking (DST) subsystem (Balaraman, Sheikhalishahi and Magnini, 2021) and for generating responses. ATICA tracks the progress of the conversation using a progression function (Sanders, 2022), and evaluates the state of the conversation to task completion. Currently, ATICA is being used to help Air Force analysts through the information foraging phase. A prior recommender system, COLLANE system, enabled collaborative work and information sharing during intelligence analysis tasks.

Current Use Case

There is broad potential for the CIRIAS system for the Air Force. Air Force intelligence analysts' concerns center around resource identification and allocation, collaborative sensemaking during joint tasks, and analysis cohesion and production. We hypothesize that the CIRIAS system can positively impact intelligence analysts during information foraging and sensemaking stages. The use case that we will implement to investigate this hypothesis is based on traffic pattern analysis – a task with some resemblance to Air Force efforts. Air Force analysts regularly work in air operations centers, where large groups of analysts work on air traffic activity, working with other airmen that collect and disseminate intelligence, and work on problem solving (Hurter *et al.*, 2012).

Analysts often face discrepancies across the analysis process, and the results of a finished intelligence product can mask the dysfunctional parts of the workflows. Finished intelligence products are a variety of information products, ranging from text, image, and data files. Our goal for the CIRIAS system is to provide a cognitively immersive and intuitive, approachable space where analysts can interact with information through the sensemaking processes to create these intelligence products. As most analysts' processes are done with standard programs, part of the challenge is enabling programs and processes they are familiar with, but providing a space that will not require them to break their workflow to reach a finished product. Analysts' approach to intelligence requests typically require a solid basis in the information foraging phase in order to produce specific and sophisticated insights in the sensemaking phase, and eventually a finished intelligence product. Analysts who are more experienced can often rely on their knowledge base to reduce time in the information foraging phase, but the challenges still arise in marshaling pertinent and important information. CIRIAS' technology, including the ATICA dialog system, are designed to enable analysts to easily access necessary information. Specific tools we have enabled in the immersive room include the brainstorming tool, map based data tools, and a dialog agent.

Since analysts often work collaboratively in simultaneous lines in the same intelligence problem, scheduling and resource challenges make it difficult to easily and quickly share information. The CIRIAS system is designed to enable collaborative work, and has been tested with small groups of users (Briggs *et al.*, 2020, 2021, 2022). The Pin-Up tool (Chabot and Braasch, 2022) is designed to enable collaborative sharing of information, including text, image, and video.

The NYC cab dataset (He *et al.*, 2022) addresses similar concerns as analysts would deal with standard air tasking orders, and would allow us to determine an operational baseline to compare against CIRIAS system usage. Tasks we envision analysts to undertake that could generalize to the Air Force domain are convoy detection, meeting detection, and suspicious activity indicators. Specific metrics we will measure for this user study are information accuracy, absorption, and comprehension.

The NYC traffic cab (He et al., 2022) use case will provide several different angles to assess analysts' use of information. The traffic patterns use case is similar to the Air Force analysts' common workflow. However, the challenge in generalizing the traffic patterns for the NYC cabs to analysts' work is the nature of the data. Traffic patterns include a lot of data points where it is easy to find regular patterns, but difficult to track anomalies and individual data points. The authors describe their use case of ridership prediction and anomaly detection (He et al., 2022). However, for air traffic, there are relatively fewer data points, where it can be more difficult to find regular patterns, but easier to track and spot individual data points. However, we anticipate that anomaly detection can be generalized to analysis tasks in the anticipated user domain with pattern of life analysis. Comparing common routes with anomalous routes is a familiar task for analysts that will be enabled in the CIRIAS map and street view visualizations. Another problem faced by analysts is the granularity and reliability of information, as well as competing sources of information, which can problematize and confuse analysis. We anticipate our technology (CIRIAS) can be used to help analysts determine the most relevant information, allowing them to marshal and present evidence in a collaborative space. The ATICA dialog agent will be used to help link analysts to evidence. We intend to integrate this use case into the cognitive immersive room, and develop a user study in order to determine the effects of the cognitive immersive room on points like accuracy, information absorption and retention, and ease of collaboration.

CONCLUSION AND FUTURE WORK

In this paper, the concept and first prototype of the Cognitive Immersive Room for Intelligence Analyst Scenarios (CIRIAS) has been described. The next steps of this ongoing research are twofold. Firstly, a user study is currently under preparation. In the first study, the NYC traffic scenario will be used to test the effectiveness of the CIRIAS environment as a recommender system for expert groups. The goal of this study is the effectiveness of the information gathering, selection, and evaluation process. While important, the user friendliness of the system is secondary to the results it produces. The second goal is the modularization of the system so scalable solutions can be provided to many operation-room scenarios from small temporary installations in tents to large facilities in permanent bases. Single or multiple flat-panel screens can be used to reduce the complexity and footprint of the video projection-based system. An audio system of four loudspeakers and a spherical microphone with four capsules is sufficient to provide a rudimentary version of the immersive sound presentation and tracking system. The effectiveness of differently scaled solutions will be addressed in future user studies.

ACKNOWLEDGMENT

This paper is based upon work supported in parts by the United States Air Force under Contract No. FA8750-21-C-0075, the National Science Foundation (#1909229), the IBM Corporation under the Artificial Intelligence Research Collaboration Agreement No. W1771793 between IBM and Rensselaer. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of USAF, NSF, IBM or the U.S. Government.

REFERENCES

- Agrawal, S. and Braasch, J. (2018) 'Impression of Spatially Distributed Reverberation in Multichannel Audio Reproduction', in. *Audio Engineering Society Convention 145*, Audio Engineering Society. Available at: https://www.aes.org/elib/browse.cfm?elib=19802 (Accessed: 30 January 2023).
- Balaraman, V., Sheikhalishahi, S. and Magnini, B. (2021) 'Recent Neural Methods on Dialogue State Tracking for Task-Oriented Dialogue Systems: A Survey', in *Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue. SIGDIAL 2021*, Singapore and Online: Association for Computational Linguistics, pp. 239–251. Available at: https://aclanthology.org/2021.sigdial-1.25 (Accessed: 27 April 2022).
- Beebe, S. M. and Pherson, R. H. (2014) Cases in intelligence analysis: structured analytic techniques in action. CQ Press.

- Braasch, J., Peters, N. and Valente, D. L. (2008) 'A Loudspeaker-Based Projection Technique for Spatial Music Applications Using Virtual Microphone Control', *Computer Music Journal*, 32(3), pp. 55–71. Available at: https://doi.org/10.1162/ comj.2008.32.3.55.
- Briggs, S. *et al.* (2020) 'Brainstorming for Sensemaking in a Multimodal, Multiuser Cognitive Environment', in, pp. 66–83. Available at: https://doi.org/10.1007/978-3-030-60114-0_4.
- Briggs, S. et al. (2021) 'Thematic Units Comparisons Between Analog and Digital Brainstorming', in S. Yamamoto and H. Mori (eds.) Human Interface and the Management of Information. Information Presentation and Visualization. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 257–267. Available at: https://doi.org/10.1007/978-3-030-78321-1_20.
- Briggs, S. et al. (2022) 'Category Creation Between Digital and Analog Sensemaking Tools in a Cognitive Immersive Environment', in S. Yamamoto and H. Mori (eds) Human Interface and the Management of Information: Visual and Information Design. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 451–460. Available at: https://doi.org/10.1007/978-3-031-06424-1_33.
- Chabot, S. et al. (2020) 'A Collaborative, Immersive Language Learning Environment Using Augmented Panoramic Imagery', in 2020 6th International Conference of the Immersive Learning Research Network (iLRN). 2020 6th International Conference of the Immersive Learning Research Network (iLRN), pp. 225–229. Available at: https://doi.org/10.23919/iLRN47897.2020.9155140.
- Chabot, S. and Braasch, J. (2017) 'An Immersive Virtual Environment for Congruent Audio-Visual Spatialized Data Sonifications'. Available at: https://doi.org/10. 21785/icad2017.072.
- Chabot, S. and Braasch, J. (2022) 'Walkable auralizations for experiential learning in an immersive classroom', *The Journal of the Acoustical Society of America*, 152(2), pp. 899–910. Available at: https://doi.org/10.1121/10.0012985.
- Cruz-Neira, C., Sandin, D. J. and DeFanti, T. A. (1993) 'Surround-screen projectionbased virtual reality: the design and implementation of the CAVE', in *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pp. 135–142.
- DeFanti, T. A. *et al.* (2009) 'The StarCAVE, a third-generation CAVE and virtual reality OptIPortal', *Future Generation Computer Systems*, 25(2), pp. 169–178. Available at: https://doi.org/10.1016/j.future.2008.07.015.
- Divekar, R. R. et al. (2019) 'CIRA: An Architecture for Building Configurable Immersive Smart-Rooms', in K. Arai, S. Kapoor, and R. Bhatia (eds) Intelligent Systems and Applications. Cham: Springer International Publishing, pp. 76–95. Available at: https://link.springer.com/10.1007/978-3-030-01057-7_7 (Accessed: 1 November 2019).
- Elder, R. L. (2017) Influence of Immersive Human Scale Architectural Representation on Design Judgment. Rensselaer Polytechnic Institute.
- Endert, A. *et al.* (2014) 'The human is the loop: new directions for visual analytics', *Journal of Intelligent Information Systems*, 43(3), pp. 411–435. Available at: http s://doi.org/10.1007/s10844-014-0304-9.
- He, M. *et al.* (2022) 'Ridership prediction and anomaly detection in transportation hubs: an application to New York City', *The European Physical Journal Special Topics*, 231(9), pp. 1655–1671. Available at: https://doi.org/10.1140/epjs/s 11734-022-00551-4.

- Hurter, C. *et al.* (2012) 'Air traffic analysis'. Available at: https://publica.fraunhofer .de/handle/publica/248231 (Accessed: 30 January 2023).
- Pirolli, P. and Card, S. (2005) 'Sensemaking processes of intelligence analysts and possible leverage points as identified through cognitive task analysis', in *Proceedings* of the 2005 International Conference on Intelligence Analysis, McLean, Virginia. Available at: https://phibetaiota.net/wp-content/uploads/2014/12/Sensemak ing-Process-Pirolli-and-Card.pdf.
- Raffel, C. *et al.* (2019) 'Exploring the limits of transfer learning with a unified textto-text transformer', *arXiv preprint arXiv:1910.10683* [Preprint].
- Rastogi, A. *et al.* (2020) 'Towards Scalable Multi-Domain Conversational Agents: The Schema-Guided Dialogue Dataset', *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05), pp. 8689–8696. Available at: https://doi.org/10. 1609/aaai.v34i05.6394.
- Sanders, A. (2022) 'Towards a Progression-Aware Autonomous Dialogue Agent', in Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.
- Sharma, G., Braasch, J. and Radke, R. J. (2017) 'Interactions in a Human-Scale Immersive Environment: the CRAIVE-Lab', Cross-Surface 2016, in conjunction with the ACM International Conference on Interactive Surfaces and Spaces [Preprint]. Available at: https://par.nsf.gov/biblio/10026264-interactions-human-sca le-immersive-environment-craive-lab (Accessed: 30 January 2023).
- Solimini, A. G. (2013) 'Are There Side Effects to Watching 3D Movies? A Prospective Crossover Observational Study on Visually Induced Motion Sickness', *PLOS ONE*, 8(2), p. e56160. Available at: https://doi.org/10.1371/journal.pone .0056160.
- Strzalkowski, T. et al. (2009) 'COLLANE: An Experiment in Computer-Mediated Tacit Collaboration', in M. Marciniak and A. Mykowiecka (eds) Aspects of Natural Language Processing: Essays Dedicated to Leonard Bolc on the Occasion of His 75th Birthday. Berlin, Heidelberg: Springer (Lecture Notes in Computer Science), pp. 411–448. Available at: https://doi.org/10.1007/978-3-642-04735-0_17.
- Wefers, F. *et al.* (2015) 'Audiotechnik des aixcave virtual reality-systems', *Fortschritte der Akustik–DAGA*, pp. 467–470.