From Pilot to Infrastructure: Embedding XR as a Training Layer for Emergency Service Organisations

Leon Pietschmann*, Fritz Pickhardt

Federal Agency for Technical Relief (THW), Bonn/Düsseldorf, Germany

Abstract

Emergency service organisations must balance realism, safety, and scalability under tight logistical constraints, both in training and in the preparation of operational support. Extended Reality (XR) promises repeatable practice without hazard exposure, yet many deployments remain short-lived pilots that fail to survive organisational friction. In this paper, we share lessons learned from building and embedding XR within a national emergency organisation across multiple domains: procedural training for emergency scaffolding (EGS), situational judgement for flood defence (dyke inspection), hazardous-material response preparation (CBRN), and section leader training. Drawing on longitudinal deployment experience and measurable outcomes from a real certification context, including a 21% improvement in pass rates in EGS training, we argue for a shift in emphasis: from "XR as a tool" to XR as an organisational layer that can be adopted, maintained, and trusted alongside hands-on practice. We distil five practical lessons for moving beyond pilots, including co-creation as legitimisation, positioning XR as an optional add-on rather than a replacement, and designing for logistics, variability, and evidence-based buy-in. While our cases focus on pre-operational training, we argue these deployments may serve as a realistic entry point toward durable operational augmentation in challenging environments. We conclude with a forward-looking agenda focused on institutional embedding, cost-benefit clarity, and cross-domain transfer in high-stakes ecosystems.

Keywords

Emergency Services, Case Study, VR, XR, Immersive, Training, Support, Effectiveness, Flood Protection, HCI

1. Why XR pilots struggle in emergency services

Emergency service organisations frequently operate in decentralised, high-stakes settings where training is both essential and difficult to scale. Realistic exercises demand space, equipment, instructors, and time; they are often constrained by costs, weather, availability, and safety risk. At the same time, training accidents are non-trivial, and hazardous domains (e.g.Flash Floods, CBRN) further raise the cost and risk of rehearsal. Data from Germany's Federal Agency for Technical Relief (THW) indicates that 65% of reported accidents involving harm to ES workers occur during live operational training [1]. Numbers reported by firefighters in Germany and the United States indicate that 25% (DE) and 10-13% (US) of injuries are caused by training-related accidents, thus making training one of the leading causes for injuries [2, 3, 4]. Here, XR offers an appealing proposition: repeated, self-paced practice with dynamic guidance in controlled conditions, independent of location and often with reduced hazard exposure [5, 6, 7, 8, 9, 10].

Despite this promise, many XR initiatives in mission-critical domains remain pilots: compelling demonstrations that do not translate into sustained use. Our experience suggests that the limiting factor is rarely graphics fidelity. Instead, deployments succeed or fail on organisational realities: instructor workflows, stakeholder legitimacy, logistics (shipping, updates, support), maintenance ownership, and whether the technology is perceived as threatening or augmenting embodied expertise.

In this paper, we share an applied perspective on how XR can move from pilot to a durable *training layer* within an emergency service organisation.

CHI'26: XR4CE workshop, April 14, 2026, Barcelona, Spain

*Corresponding author.
LHP30@cantab.ac.uk (L. Pietschmann)
0000-0001-6069-4567 (L. Pietschmann); 0009-0000-7312-7123 (F. Pickhardt)

2. A portfolio view: XR across training domains

Rather than treating XR as a single-purpose simulator, we view it as a modular, integrated layer that can support different competency types across domains:

Procedural and spatial knowledge We deployed a VR application for training the assembly of an emergency scaffolding system (EGS) as part of a live training programme over multiple cohorts ($n = 221$). In a between-subjects comparison against conventional preparation, VR-supported cohorts showed a statistically significant improvement of 21% in course pass rates. Participants broadly accepted VR as a valuable supplement, with minimal signs of cybersickness, while still expressing concerns around cost, haptics, and the fear of replacing hands-on drills [11, 12].

Situational, time-constrained judgement under uncertainty (Flood Defense / dyke inspection). Flood protection training involves spatial assessment and independent recognition of damage patterns that evolve over time. Unlike purely procedural rehearsal, the emphasis is on interpretation, judgement, and documentation under variable conditions. Our Flood Defence focused work explores how VR can support repeated inspection and defence exercises independent of weather and seasonal constraints, while remaining embedded in real training courses (ongoing case study)[13].

Hazmat preparation (CBRN). CBRN training often faces unavoidable tension: realism requires exposure to stressful conditions, yet exposure carries risk and cost. XR can provide a low-risk entry point for familiarisation with environments, equipment, and procedures, and for rehearsing decision points without physical hazard. We are implementing XR modules for CBRN-related training as part of a broader organisational XR portfolio [14, 15].

Section Leader Training. XR training for emergency response yields its strongest results in application-oriented scenarios with clear KPIs, as our scaffolding, CBRN and flood defence study demonstrates. Leadership training, however, remains harder to evaluate. Nonetheless, XR uniquely enables repeated safe exposure to high-stress decision-making, building cognitive and emotional resilience that conventional formats often fail to replicate. Developing evaluation frameworks for these less tangible leadership outcomes is a key frontier for the field[16, 17].

Across these domains, the common challenge is not merely "does it work?", but "does it *stick*?": Can the system be used repeatedly, by different instructors, with different learners, across sites, and beyond the novelty phase?

3. Beyond the pilot: five lessons for durable XR training

We summarise five lessons from scaling XR training inside a real emergency service organisation.

3.1. Lesson 1: Co-creation is legitimisation, not just ideation

Early iteration with instructors, subject matter experts, and pilot users is often framed as a way to elicit requirements. In our experience, it also serves another important goal: *legitimisation.* Iterative review cycles with those who own and perform the training (instructors) and those who embody operational expertise (SMEs, experienced responders) establish shared ownership and reduce later resistance. This is especially important when introducing new forms of assessment, guidance, or feedback, where perceived "outsider" design can trigger scepticism.

Practically, we found that short iteration loops with concrete artefacts (walkthrough builds, scenario slices, guidance prototypes) outperformed abstract workshops. Pilot users surface friction that experts may miss (e.g., onboarding, comfort settings, navigation habits), while SMEs protect realism and training relevance.

3.2. Lesson 2: Instructor buy-in is necessary but insufficient

Instructor buy-in is a gatekeeper for deployment, but sustained adoption requires alignment across a wider set of actors: training administrators, IT/operations support, procurement, and learners themselves. For example, our EGS deployment showed that some trainees expressed reservations about spending on VR rather than "real equipment" and worried that VR could erode the primacy of hands-on practice. These perceptions matter. They shape whether participants use the system at home, recommend it to peers, and accept it as legitimate preparation.

Thus, buy-in must be treated as a multi-stakeholder design target, supported by communication and evidence, not only by interface quality.

3.3. Lesson 3: Position XR as an optional add-on, not a replacement

One of The most consistent adoption levers was *framing*. While the ultimate goal remains the integral embedding of XR into existing training curricula, our experience shows that adoption must begin by positioning XR as a supplementary tool, demonstrating clear added value alongside established practices before deeper integration can be justified and sustained.

In mission-critical training cultures, competence is often equated with embodied, material practice. Systems perceived as threatening that practice trigger resistance. In contrast, XR was readily accepted when positioned as an optional layer for preparation and reflection, with clear boundaries: XR *augments* drills; it does not replace them.

In the EGS context, training protocols required maintaining the standard course structure; VR was introduced strictly as additional preparation time. Explicitly reassuring trainees that the programme would not become "fully virtual" lowered barriers and improved acceptance. This positioning is also important for instructors, who often hold responsibility for safety and certification integrity.

3.4. Lesson 4: Design for organisational logistics, not ideal usage

Scaling XR across decentralised sites turns "deployment" into a first-class design problem. Seemingly mundane factors become dominant: device shipping and returns, account management, resets, updates, breakage, quick-start materials, and the level of available technical support. Failure to plan for these can derail usage even when the application itself is effective.

We recommend treating logistics as part of the system: design for intermittent use, minimal setup, robust onboarding, and graceful operation under partial support (e.g., offline-first or low-dependency modes where feasible). Importantly, the goal is not to eliminate friction entirely, but to make adoption viable under realistic organisational capacity.

3.5. Lesson 5: Evidence unlocks political and financial viability

In organisations with limited budgets and high accountability, qualitative enthusiasm is rarely enough. In our experience, measurable outcomes, such as improved course pass rates in a real certification context, changed the conversation: from "interesting innovation" to "operationally relevant investment". Evidence also helps resolve the add-on vs. replacement debate by showing that XR can improve readiness without displacing hands-on drills.

We argue that XR programmes should plan evaluation as an adoption strategy, not an afterthought: define feasible metrics early (pass rates, completion, instructor workload, safety proxies, usage time), and use them to guide iteration and scaling decisions.

4. From tool to training layer

Taken together, these lessons support a reframing. The most robust path is not building isolated simulations, but establishing an XR *training layer* that: (i) supports different competency types (procedural rehearsal, situational judgement, hazard familiarisation), (ii) integrates alongside existing curricula without threatening embodied practice, (iii) is maintained through realistic logistics and ownership structures, and (iv) builds legitimacy through co-creation and evidence.

This framing also encourages reuse and transfer: design patterns (onboarding, guidance, difficulty control, scenario packaging, instructor modes) can carry across domains such as EGS, flood response, and CBRN, even as content differs.

5. From Training Tool to Operational Support

Our cases focus on pre-operational training rather than live field deployment. However, we argue that training is a realistic entry point toward operational XR in emergency services, because it exposes many of the same sociotechnical conditions that later shape in-field augmentation. Across our deployments, successful embedding depended on practitioner legitimacy, early iteration with subject matter experts and pilot users, low-friction logistics, and framing XR as an optional add-on rather than a replacement. We expect these same conditions to be of relevance for operational assistance tools used under stress and uncertainty. In this sense, training deployments do not yet provide evidence for live augmentation, but they do offer grounded insight into what it would take to make operational XR acceptable, maintainable, and trustworthy in practice. We therefore see our contribution as helping to bridge the training-to-operations gap: not by claiming field validation, but by identifying organisational preconditions that future XR augmentation tools will need to meet.

6. Conclusion

XR can meaningfully augment emergency service training, but lasting impact depends on institutional embedding rather than technical novelty. Based on multi-domain deployment experience inside a national emergency organisation, we propose focusing on XR as a resilient training layer: co-created for legitimacy, framed as an optional add-on, engineered for real (logistical) conditions, and supported by measurable evidence. We hope these reflections help the community move beyond pilots toward durable, scalable XR training ecosystems for high-stakes professional practice.

Acknowledgments

The authors wish to thank the whole THW XR development team. Moreover, this study would not have been possible without the continued support from THW (particularly the instructors, SME's, and participants). This research is supported in part by the Federal Ministry of the Interior and Community of Germany.

Declaration on Generative AI

In line with ACM's and CEUR-WS's Policy on Authorship and GenAI, supervised generative AI (ChatGPT 5.4) supported the development of this text with paraphrasing and spellchecking.

References

[1] B. T. H. (THW), Präventionsbericht 2023, 2023.
[2] F. U. Mitte, Unfallstatistik 2023, 2023.
[3] C. Campbell, S. Hall, United states firefighter injuries, National Fire Protection Association (2023).
[4] J. Garcia, M. C. Bazzocchi, K. Fite, J. D. Ocampo, M. Martinez, Review and statistical analysis of us structural firefighting injuries: Their causes and effects, Fire 7 (2024) 46.
[5] L. Pietschmann, Use of virtual reality for emergency service training, 2023. Manuscript in preparation for submission to ACM CHI Conference on Human Factors in Computing Systems 2024 – LBW track.

[6] A. Farr, L. Pietschmann, P. Zürcher, T. Bohné, Skill retention after desktop and head-mounted-display virtual reality training, Experimental Results 4 (2023) 5. URL: https://doi.org/10.1017/exp.2022.28. doi:10.1017/exp.2022.28.
[7] L. Pietschmann, P.-D. Zuercher, E. Bubík, Z. Chen, H. Pfister, T. Bohné, Quantifying the impact of xr visual guidance on user performance using a large-scale virtual assembly experiment, 2023. URL: https://arxiv.org/abs/2308.03390. doi:10.48550/ARXIV.2308.03390, in press. Preprint available via https://arxiv.org/abs/2308.03390.
[8] L. Pietschmann, M. Schimpf, Z.-T. Chen, H. Pfister, T. Bohné, Enhancing user performance and human factors through visual guidance in ar assembly tasks, in: Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, CHI EA '25, Association for Computing Machinery, New York, NY, USA, 2025. URL: https://doi.org/10.1145/3706599.3720094. doi:10.1145/3706599.3720094.
[9] L. Pietschmann, T. Bohné, M. Tsapali, Extended reality visual guidance for industrial environments: A scoping review, in: 2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS), IEEE, Cambridge, UK, 2022, pp. 211–215. URL: https://doi.org/10.1109/ichms56717.2022.9980683. doi:10.1109/ichms56717.2022.9980683.
[10] L. Pietschmann, Human-Computer Interaction in Extended Reality: Exploring the Impact of Visual Guidance on User Performance and Human Factors, Ph.D. thesis, University of Cambridge, United Kingdom, 2024.
[11] L. Pietschmann, F. Pickhardt, Use of virtual reality for emergency service training, in: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, CHI EA '24, Association for Computing Machinery, New York, NY, USA, 2024. URL: https://doi.org/10.1145/3613905.3651119. doi:10.1145/3613905.3651119.
[12] R. Lovreglio, X. Duan, A. Rahouti, R. Phipps, D. Nilsson, Comparing the effectiveness of fire extinguisher virtual reality and video training, Virtual Reality 25 (2021) 133–145.
[13] J. M. Mol, W. W. Botzen, J. E. Blasch, After the virtual flood: Risk perceptions and flood preparedness after virtual reality risk communication, Judgment and Decision Making 17 (2022) 189–214.
[14] J. Göllner, A. Peer, C. Meurers, G. Wurzer, C. Schönauer, H. Kaufmann, C. Bösch, Virtual reality cbrn defence, in: Meeting Proceedings of the Simulation and Modelling Group Symposium, volume 171, NATO, Vienna, AUSTRIA, 2019, pp. 1–25.
[15] F. Lamberti, F. De Lorenzis, F. G. Prattico, M. Migliorini, An immersive virtual reality platform for training cbrn operators, in: 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), IEEE, 2021, pp. 133–137.
[16] C. H. Wijkmark, M. M. Metallinou, I. Heldal, Remote virtual simulation for incident commanders—cognitive aspects, Applied Sciences 11 (2021) 6434.
[17] A. Mossel, C. Schoenauer, M. Froeschl, A. Peer, J. Goellner, H. Kaufmann, Immersive training of first responder squad leaders in untethered virtual reality, Virtual Reality 25 (2021) 745–759.