

## ARTICLE INFO

|           |               |
|-----------|---------------|
| Received  | 09 March 2026 |
| Revised   | 15 April 2026 |
| Accepted  | 17 April 2026 |
| Published | 30 June 2026  |

## MONITORAGGIO DELLO STATO UMANO NELLA FORMAZIONE MARITTIMA

Il caso On Watch

## HUMAN STATE MONITORING IN MARITIME TRAINING

The On Watch case

Mario Ivan Zignego, Alessandro Bertirotti, Paolo Gemelli, Laura Pagani

## ABSTRACT

Nei domini critici per la sicurezza la formazione è un'infrastruttura socio-tecnica (spazi, interfacce, procedure e fasi di preparazione e di riesame finale) che sostiene l'apprendimento e la cultura della sicurezza. In questa prospettiva il paper propone il quadro operativo del Designer-in-the-Loop (DITL) per introdurre gli Human State Monitoring Systems (HSMS) nella formazione marittima, rendendo verificabili i compromessi tra innovazione digitale e rischi di sorveglianza, distorsioni e disuguaglianze. In tale quadro, On Watch è assunto come caso di studio poiché consente di sperimentare operativamente il framework in un contesto applicativo concreto, integrando segnali multimodali e traducendoli in indicatori probabilistici non prescrittivi, privi di graduatorie e di automatismi decisionali, a supporto del riesame finale. Una sperimentazione pilota, articolata in tre contesti – ECDIS, plancia cooperativa e colloquio – valuta la fattibilità, l'utilità e l'accettabilità mediante indicatori chiave di prestazione tecnico-operativi e socio-distributivi replicabili.

In safety-critical domains, training is a socio-technical infrastructure (spaces, interfaces, procedures, and preparation, execution, and final review phases) that supports learning and the development of a safety culture. From this perspective, the paper proposes an operational framework for Designer-in-the-Loop (DITL) to introduce Human State Monitoring Systems (HSMS) into maritime training, making the trade-offs between digital innovation and surveillance-related risks, such as bias and inequality, explicit and verifiable. Within this framework, On Watch is adopted as a case study, enabling operational testing in a concrete application context by integrating multimodal signals and translating them into non-prescriptive probabilistic indicators, without rankings or automated decision-making, in support of the final review phase. A pilot experiment, structured across three contexts – ECDIS, cooperative bridge, and interview – evaluates feasibility, usefulness, and acceptability through replicable technical-operational and socio-distributive key performance indicators.

## KEYWORDS

formazione marittima per la sicurezza, monitoraggio dello stato umano, governance dei dati, tecnologie digitali, monitoraggio multimodale

maritime safety training, human state monitoring, digital technologies, data governance, multimodal monitoring

**Mario Ivan Zignego**, Full Professor at the Department of Architecture and Design (DAD) at the University of Genoa (Italy), conducts research in nautical design and digital innovation. E-mail: [mario.ivan.zignego@unige.it](mailto:mario.ivan.zignego@unige.it)

**Alessandro Bertirotti**, Adjunct Professor at the Department of Architecture and Design (DAD) at the University of Genoa (Italy), conducts research in anthropology and design. E-mail: [alessandro.bertirotti@edu.unige.it](mailto:alessandro.bertirotti@edu.unige.it)

**Paolo Gemelli**, Adjunct Professor at the Department of Architecture and Design (DAD) at the University of Genoa (Italy), conducts research in naval architecture and training. E-mail: [paolo.gemelli@edu.unige.it](mailto:paolo.gemelli@edu.unige.it)

**Laura Pagani** is a PhD Candidate at the Department of Architecture and Design (DAD), University of Genoa (Italy). She conducts research in human-centred artificial intelligence and socio-technical systems for safety training in the maritime domain. E-mail: [laura.pagani@edu.unige.it](mailto:laura.pagani@edu.unige.it)



Nei domini critici per la sicurezza la formazione è un'infrastruttura socio-tecnica: spazi (aule, simulatori, plance), artefatti (checklist, procedure) e fasi – soprattutto per la preparazione e il riesame finale – che abilitano l'apprendimento, la prevenzione degli errori e la costruzione di una cultura della sicurezza. Nel settore marittimo le performance dipendono dalla cooperazione, dalla consapevolezza situazionale e dalla gestione del carico cognitivo; per questo l'infrastruttura formativa è 'abilitante' per la continuità operativa e la riduzione del rischio. In tale scenario la digitalizzazione (sensori, analisi dati, simulazione, AI/IoT) può migliorare qualità e tracciabilità della formazione (SDG 9), ma può anche innescare violazione della privacy, distorsioni e disuguaglianze (SDG 10) se non progettata e governata (Salas et alii, 2012; Aguinis and Kraiger, 2009; ISO, 2019; Fig. 1). Questo rende cruciale progettare strumenti che sostengano la riflessione, non il mero adeguamento passivo.

In tale ottica il contributo assume che il nodo non è se adottare tecnologie basate sui dati nella formazione sui domini critici per la sicurezza, ma come progettarle affinché l'innovazione resti coerente con i principi dell'Industria 5.0 e con un paradigma centrato sull'uomo: la tecnologia deve aumentare la capacità del sistema di apprendere dall'esperienza e prevenire l'errore, senza trasformare l'infrastruttura formativa in un apparato di controllo. Ne deriva che il valore degli Human State Monitoring Systems (HSMS) non dipende solo dall'accuratezza dell'inferenza, ma anche dalla capacità dell'infrastruttura di rendere espliciti gli scopi, i limiti e le responsabilità lungo l'intero ciclo d'uso. A tal fine il progetto deve incorporare vincoli tecnici e organizzativi che impediscano riusi incompatibili e letture prescrittive.

L'ipotesi di fondo è che tale equilibrio sia raggiungibile solo spostando il focus dalla conformità ex post (norme e politiche aggiunte 'a posteriori') al governo del sistema integrato fin dalla progettazione: minimizzazione e separazione dei dati identificativi, consenso informato e revocabile, facoltà del soggetto di sottrarsi al monitoraggio, politiche di conservazione, tracciabilità e verifica, limiti d'uso espliciti e verificabili lungo l'intero ciclo vita del sistema. In altre parole la governance non è un capitolo 'normativo' a margine, ma una componente progettuale che condiziona interfacce, flussi informativi e fasi operative entro cui la tecnologia prende significato (Mittelstadt, 2019; NIST, 2023; OECD, 2019).

All'interno di questo quadro si collocano il caso studio On Watch e il prototipo Human State Monitoring System (HSMS), sviluppato e testato in contesti didattici e simulativi progressivi (aula con simulatore ECDIS – Electronic Chart Display and Information System, simulazione di plancia con ruoli cooperativi, colloqui / selezione), in collaborazione con l'Università di Genova e la FAIMM Accademia della Marina Mercantile, nell'ambito del programma 'ITS 4.0' (MIM, 2025). HSMS integra segnali multimodali (voce, indizi facciali e posturali, respirazione, testo e interazioni operative) e li traduce in indicatori probabilistici di stato (ad esempio attivazione e carico) da utilizzare come supporto informativo nelle fasi di preparazione e di riesame finale. In coerenza con l'obiettivo di evitare derive disciplinari e disuguaglianze il sistema è deliberatamente non prescrittivo: non produce classifiche individuali, non abilita graduatorie, non au-

tomatizza decisioni operative e non è progettato per finalità di controllo del lavoratore, ma restituisce visualizzazioni e segnali configurabili per scenario e fase, con tracciabilità delle fonti e controllo umano (Ronca et alii, 2023; Liu et alii, 2020; Figg. 2, 3).

L'originalità del contributo è duplice. In primo luogo propone una lettura della 'sicurezza formativa' come infrastruttura progettuale e gestionale, dove spazi, interfacce e fasi costituiscono un'unità inscindibile: è in tale unità che si producono o si riducono errori, sovraccarico informativo, false sicurezze e fallimenti di coordinamento. In secondo luogo formalizza un quadro operativo Designer-in-the-Loop per rendere misurabile e verificabile il compromesso tra innovazione digitale e riduzione delle disuguaglianze: l'adozione degli HSMS viene governata ex ante (attraverso la valutazione dei rischi di uso improprio, delle distorsioni e delle barriere d'accesso) ed ex post (tramite KPI tecnico-operativi e socio-distributivi) così che i compromessi non restino impliciti, ma diventino oggetto di progetto, verifica e miglioramento iterativo.

Alla luce di queste premesse il contributo si propone di chiarire come l'impiego di HSMS nella formazione marittima possa aumentare la qualità del riesame finale e dell'apprendimento dall'errore senza produrre effetti regressivi di sorveglianza o disuguaglianza. In particolare il paper: a) inquadra criticamente il dibattito su Affective Computing e HSMS nei domini critici per la sicurezza, evidenziando limiti epistemici e rischi d'uso improprio; b) descrive On Watch come caso di studio, mettendo in primo piano l'architettura dell'esperienza, le scelte di interfaccia e gli scenari d'uso; c) formalizza il quadro metodologico Designer-in-the-Loop con requisiti e KPI tecnico-operativi e socio-distributivi per una valutazione ex ante ed ex post replicabile e trasferibile. Il paper è organizzato come segue: stato dell'arte; caso studio; architettura dell'esperienza; interazione e interfacce; framework Designer-in-the-Loop; protocollo e KPI; risultati preliminari e discussione; barriere/limiti e trasferibilità; SDG; conclusioni.

Le domande di ricerca sono: 1) quali condizioni progettuali e di governance consentono di integrare HSMS nella formazione dei domini critici per la sicurezza senza trasformarli in dispositivi prescrittivi o disciplinari; 2) quali scelte di architettura dell'esperienza e di interfaccia rendono il sistema verificabile, controllabile e non prescrittivo nei diversi scenari; 3) quale set minimo di requisiti, soglie di qualità e KPI permette una valutazione replicabile e supporta la trasferibilità verso altri domini critici per la sicurezza.

Per rispondere a tali domande il paper combina una rassegna critica della letteratura internazionale con l'analisi del caso On Watch e la descrizione operativa del protocollo sperimentale, esplicitando requisiti, vincoli e indicatori necessari a replicare lo studio in altri contesti.

**Stato dell'arte** | L'adozione dell'Affective Computing e degli Human State Monitoring Systems nella formazione in contesti critici per la sicurezza incrocia tre filoni: inferenza multimodale di stati affettivi / cognitivi; formazione e valutazione delle prestazioni in ambienti ad alta criticità; studi su monitoraggio e sorveglianza, equità e governance dei dati. La rassegna è intenzionalmente critica: l'obiettivo non è legittimare attribuzioni di punteggio automatico, ma chiarire limiti epistemici, rischi d'uso

improprio e requisiti progettuali per rendere tali sistemi verificabili e trasferibili (Picard, 1997; Poria et alii, 2017; D'Mello and Kory, 2015; Ajunwa, Crawford and Schultz, 2017; Eurofound, 2020; Zuboff, 2019). In questa prospettiva la multimodalità e la natura situata dei processi percettivi / decisionali suggeriscono di trattare i risultati come stime probabilistiche e contestuali, a supporto della costruzione condivisa di senso nelle fasi di preparazione e di riesame finale e non come verdetti prescrittivi (Gazzaniga, 2005).

La letteratura sul rilevamento delle emozioni conferma la spinta verso approcci multimodali, ma evidenzia eterogeneità di protocolli, comparabilità limitata, disallineamenti tra insiemi di dati e problemi di generalizzazione, imponendo cautela contro false precisioni in domini critici per la sicurezza (Zeng et alii, 2009; Koelstra et alii, 2012). Nel filone dei fattori umani l'utilità degli HSMS emerge soprattutto quando sono integrati nella funzione pedagogica di preparazione e di riesame finale, in cui la qualità della riflessione dipende da artefatti, ruoli e sequenze comunicative più che dalla sola disponibilità di dati (Hart and Staveland, 1988; Endsley, 1995; Hancock and Matthews, 2019).

Sul piano applicativo alcune buone pratiche internazionali nell'ambito dell'addestramento basato su simulatori mostrano come l'uso di indicatori di carico e di stress sia più robusto quando è ancorato a scenari, a marker e a un riesame finale strutturato, e quando le misure sono trattate come evidenze contestuali e non come punteggi individuali. Studi recenti nella simulazione sui ponti evidenziano il valore della triangolazione tra dati psicofisiologici e osservazioni operative per interpretare carico di lavoro, stress e attenzione e per discutere criticamente limiti e affidabilità delle misure (Wulvik, Dybvik and Steinert, 2020; Ronca et alii, 2023; Xue et alii, 2024). Questi casi indicano criteri di selezione trasferibili: a) scenario ad alta criticità con ruoli espliciti; b) raccolta multimodale con punti di verifica della qualità; c) restituzione non prescrittiva nel riesame finale; d) governance che limita riusi disciplinari.

Infine nel dibattito su oggetti e servizi 'potenziati dall'intelligenza artificiale' l'attenzione si sposta verso l'esperienza d'uso e le responsabilità progettuali: configurabilità, controllo umano effettivo e integrazione della filiera dei dati diventano condizioni per evitare effetti regressivi e rendere i sistemi verificabili (Arquilla and Paracoli, 2023; Osello et alii, 2024; Villani, Romagna and Oddi, 2024), mentre studi su design dell'errore, alienazione e pressione valutativa mettono in guardia dal rischio di trasformare strumenti formativi in dispositivi di controllo; contributi sull'inclusività della multimodalità suggeriscono inoltre di leggere i segnali come evidenze situate da discutere nel riesame finale piuttosto che come mere etichette (Porfirione, Ferrari Tumay and Leggiero, 2024; Vacanti et alii, 2025; Rosmino, 2024; Arquilla and Caruso, 2025; Zehr Gantz et alii, 2025). Ne deriva il requisito di prevenire impieghi ai fini di sorveglianza lungo l'intera filiera, tramite il governo di un sistema integrato fin dalla progettazione e valutazioni anche socio-distributive; l'AI Act rafforza questa impostazione in termini di rischio, trasparenza e responsabilità organizzativa (European Union, 2024).

**Caso studio: On Watch e HSMS** | Il contributo si fonda sul caso studio On Watch e sul prototipo

HSMS, concepito come piattaforma sperimentale per supportare la formazione e il riesame finale in contesti critici per la sicurezza. La sperimentazione pilota è stata sviluppata con la FAIMM Accademia della Marina Mercantile e l'Università di Genova nell'ambito dell'ITS 4.0 Challenge 2025 (MIM, 2025). La sperimentazione segue una strategia a complessità crescente in tre contesti: a) aula con simulatore ECDIS; b) plancia cooperativa con ruoli differenziati; c) colloquio / selezione, con configurazioni più restrittive per ridurre lo stigma e il rischio d'uso improprio. Nello scenario plancia la prova principale usa un'emergenza 'uomo in mare' con più ruoli (Fig. 4); il simulatore 'integrale di plancia' impiegato per lo scenario cooperativo è illustrato in Figura 5.

All'interno di On Watch il prototipo HSMS integra segnali multimodali (voce, indizi facciali e posturali, parametri respiratori, testo e interazioni) e li traduce in indicatori probabilistici (ad esempio attivazione e carico) a supporto delle fasi di preparazione e di riesame finale. Il sistema è deliberatamente non prescrittivo: non produce una graduatoria né automatizza le decisioni, ma restituisce visualizzazioni configurabili con incertezza e tracciabilità (Ronca et alii, 2023; Wulvik, Dybvik and Steinert, 2020; Xue et alii, 2024). L'uso di segnali sensibili richiede minimizzazione, pseudonimizzazione, accesso controllato, tempi definiti di conservazione dei dati e procedure di verifica, con

un'informativa chiara e un consenso revocabile, nonché la facoltà di non aderire al monitoraggio; tali vincoli sono incorporati nelle fasi di preparazione e di riesame finale e nell'interfaccia utente, secondo una logica integrata di governo del sistema fin dalla progettazione.

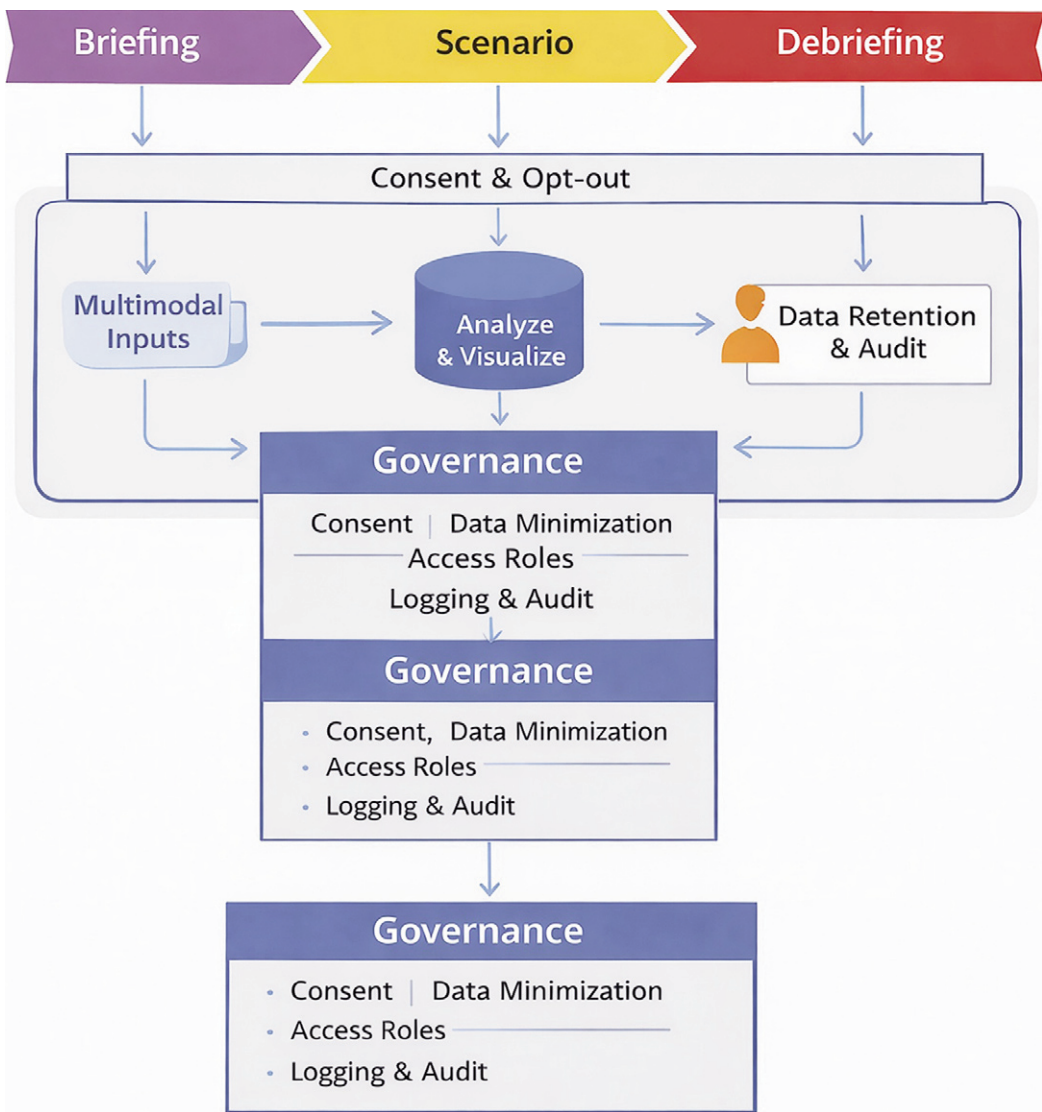
**Architettura dell'esperienza formativa** | Nei contesti critici per la sicurezza l'efficacia della formazione dipende meno dalla singola esercitazione e più dall'architettura dell'esperienza che la incornicia: preparazione, esecuzione, riflessione guidata. Nel caso On Watch tale architettura è stata progettata come una sequenza intenzionale di preparazione, scenario e riesame finale, in cui il monitoraggio multimodale non è sovrapposto all'attività, ma collocato in punti specifici del flusso per massimizzare l'utilità formativa e minimizzare gli effetti regressivi (Fanning and Gaba, 2007; Rudolph et alii, 2006; Cheng et alii, 2014). In questa prospettiva la percezione-azione e la riflessione post-evento capitalizzano il modo in cui il cervello 'riempie' le lacune e ricostruisce scenari dinamici, per esempio la soppressione saccadica e l'integrazione con la memoria e le aspettative. Di qui la scelta di concentrare gli HSMS sul riesame finale, fase in cui segnali eterogenei possono essere 'riancorati' temporalmente a eventi e decisioni operative, riducendo la 'falsa precisione' della lettura in tempo reale e valorizzando il controllo umano.

Nella fase di preparazione il formatore allinea gli obiettivi didattici e lo scenario, chiarisce ruoli e responsabilità (in particolare nei contesti cooperativi) ed esplicita le condizioni di uso del dato come parte del governo del sistema integrato fin dalla progettazione: finalità formative e di ricerca, esclusione di scopi disciplinari, minimizzazione, pseudonimizzazione, accesso controllato, possibilità di revoca e facoltà di non aderire al monitoraggio, tempi di conservazione dei dati e tracciabilità. La preparazione assume anche una funzione di 'contratto psicologico' per ridurre i timori di sorveglianza e orientare l'interpretazione degli esiti restituiti dal sistema, come supporto al riesame finale.

Nel corso dello scenario l'obiettivo non è 'misurare l'emozione' come etichetta, ma osservare come gli stati emotivi e cognitivi emergano e varino in condizioni di pressione operativa e di cooperazione. La raccolta dati è impostata come filiera sincronizzata, finalizzata all'analisi successiva alla sessione: comportamento operativo e interazioni, segnali facciali e posturali, parametri vocali e prosodici, contenuti verbali e dinamiche comunicative. In coerenza con l'obiettivo di evitare interferenze operative, gli HSMS possono produrre segnali di attenzione calibrabili per fase e scenario, progettati per ridurre falsi allarmi e non dirottare il comportamento dell'operatore.

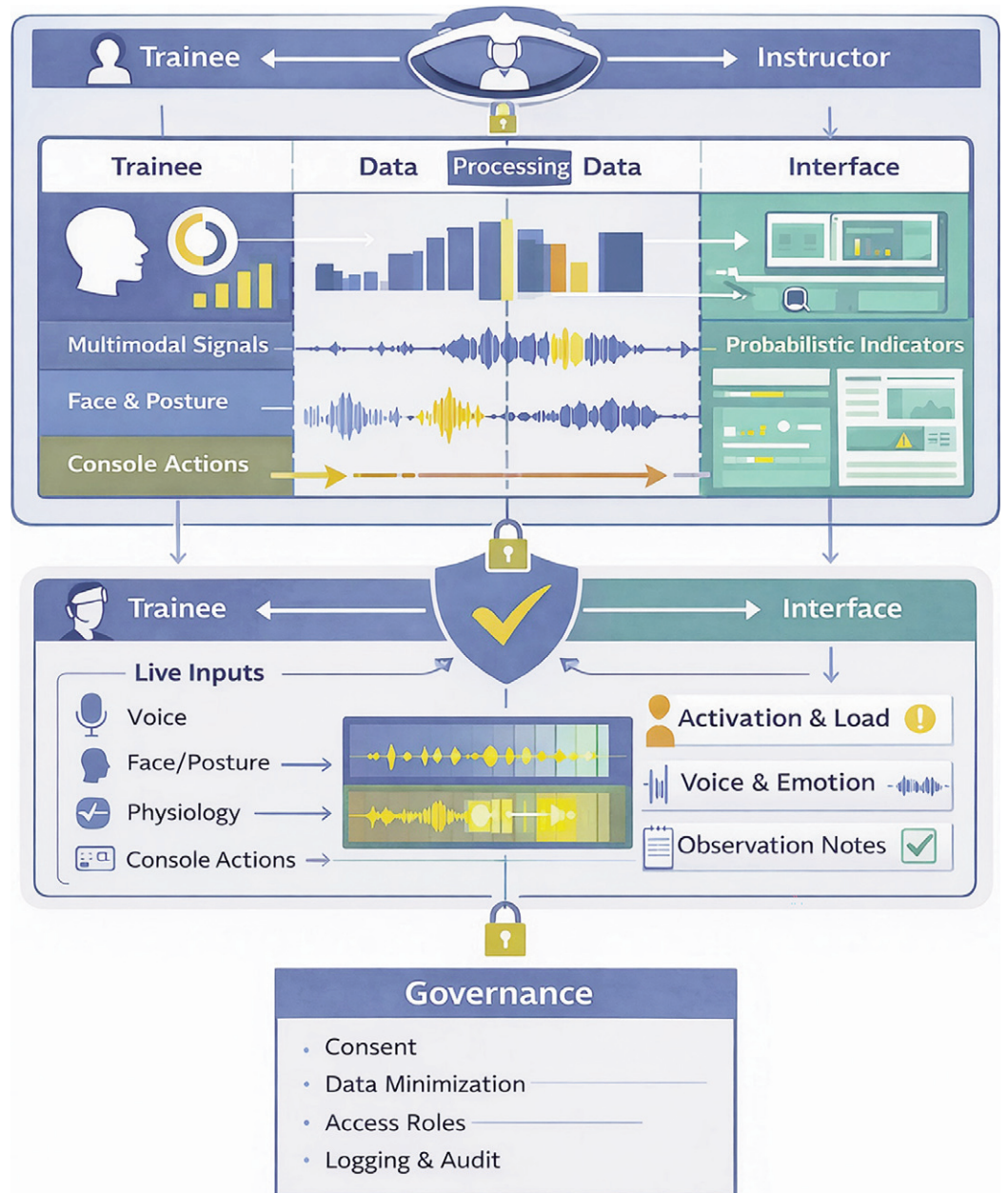
Il riesame finale è la fase in cui l'infrastruttura diventa effettivamente 'empatica': si costruisce una lettura situata dell'esperienza e si mettono in relazione decisioni, comunicazioni e momenti critici. Gli HSMS entrano pienamente in questa fase perché consentono di ancorare la discussione a evidenze temporali e multimodali, senza trasformarle in punteggi individuali. In chiave di prevenzione di derive verso logiche di sorveglianza, il riesame finale è anche un presidio di governance poiché definisce cosa può essere discusso, chi può accedere alle visualizzazioni, quali dati sono pertinenti allo scopo didattico e come vengono conservati o eliminati.

**Interazione e interfacce** | Nel caso On Watch l'interazione è progettata con un assunto esplicito: in un dominio critico per la sicurezza, l'interfaccia utente non deve 'chiudere' l'interpretazione, ma abilitare una lettura verificabile dei dati durante la preparazione e il riesame finale. In linea con questo principio le interfacce degli HSMS sono con-



**Fig. 1** | Architecture of the training experience – briefing, scenario, debriefing – and HSMS and governance intervention points – consent / opt-out, minimisation, retention, and audit (credit: the Authors; AI-generated with ChatGPT, 2026).

**Fig. 2 |** HSMS socio-technical ecosystem: actors (trainee / instructor), multimodal data flow, processing, and levels of interface output with governance-by-design constraints (credit: the Authors; AI-generated with ChatGPT, 2026).



cepite come strumenti di interpretazione orientati a rendere visibili gli andamenti e le discontinuità temporali, collegarli a eventi operativi e comunicativi, esplicitare l'incertezza e la provenienza dell'inferenza e preservare il controllo umano effettivo, prevenendo derive prescrittive o disciplinari (Parasuraman and Riley, 1997).

Il prototipo HSMS adotta un set limitato di viste, pensate per il riesame finale (e solo marginalmente per la fase in tempo reale), evitando KPI sintetici unidimensionali che tendono a trasformarsi in valutazioni. La progettazione dell'interfaccia utente può sfruttare principi gestaltici (prossimità, somiglianza, buona continuità e chiusura) per raggruppare eventi, marcatori e indicatori lungo la linea temporale, riducendo il carico cognitivo e favorendo il riconoscimento di configurazioni senza tradursi in attribuzioni di punteggi. Ad esempio il raggruppamento per ruolo e fase (somiglianza), l'allineamento del marcatore con traiettorie visive continue (buona continuità) e la chiusura delle lacune informative con le note del formatore (chiusura) supportano la discussione senza scivolare in graduatorie. La vista di default è una linea temporale con indicatori probabilistici (ad esempio attivazione e carico) e marcatori di scenario; una vista degli eventi consente, selezionando un intervallo, di visualizzare azioni operative, comunicazioni e cambi di ruolo, con la possibilità di annotazioni del formatore. Una vista multimodale basata sulle evidenze rende accessibili, quando necessario e autorizzato, le fonti che hanno contribuito alla stima (caratteristiche vocali e prosodiche e indizi facciali e posturali aggregati), corredate da marcatura temporale e metadati di qualità. Nei contesti cooperativi l'interfaccia utente supporta una lettura per ruolo, evitando confronti competitivi (Fig. 6).

Un requisito chiave è la configurabilità per scenario e fase: scenario ECDIS (con enfasi su carico e decisioni), plancia cooperativa (con enfasi su coordinazione e comunicazioni) e colloquio e selezione (con configurazioni più restrittive per ridurre il rischio di stigmatizzazione). La configurazione influisce sulla granularità temporale, sulla visibilità delle fonti e sulle regole di accesso, cioè su chi può vedere cosa e quando. Tale impostazione costituisce una misura di governo del sistema integrato fin dalla progettazione: l'interfaccia utente incorpora limiti d'uso e vincoli di accesso come parte integrante dell'esperienza, rendendo più difficile la perdita di fun-

zionalità (Fig. 7); l'ambiente ECDIS utilizzato nel pilot è mostrato nella Figura 8.

Per garantire una responsabilità tracciabile il prototipo HSMS rende verificabili la provenienza e la qualità delle modalità, le trasformazioni (dal segnale alla caratteristica estratta fino all'indicatore) e le registrazioni degli accessi, sottoposte a verifica. La non-prescrittività deriva dall'assenza di graduatorie, dalla visibilità dell'incertezza e dal controllo della visibilità delle fonti.

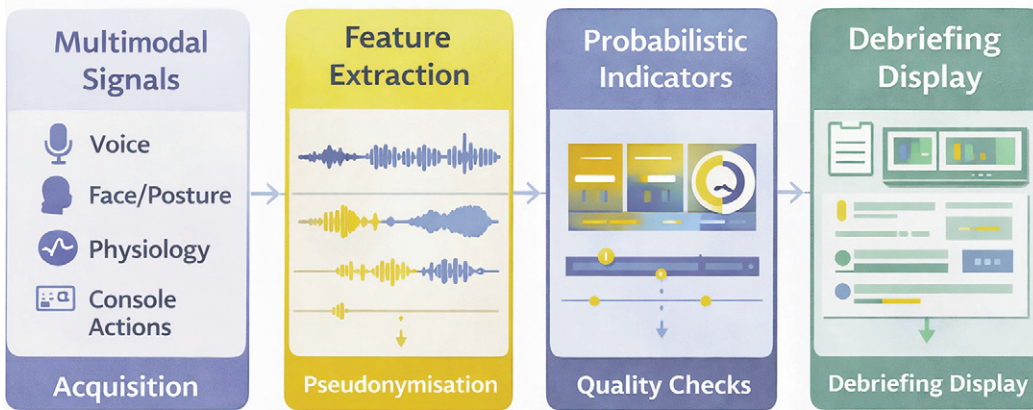
**Framework Designer-in-the-Loop** | Il framework Designer-in-the-Loop (DITL) proposto in questo contributo nasce da una premessa tipica della 'design research' applicata ai sistemi complessi: quando una tecnologia entra in un contesto sensibile alla sicurezza non 'si aggiunge' semplicemente a un processo esistente: ridefinisce invece l'infrastruttura socio-tecnica che sostiene l'apprendimento. Nel caso di un HSMS ciò significa intervenire contemporaneamente su rituali (preparazione / riesame finale), artefatti (dashboard, log, report), aspettative (cosa è legittimo dedurre dai dati) e relazioni di potere (chi osserva, chi è osservato, chi decide, chi può contestare). Per questo il divario tra innovazio-

ne (SDG 9) e riduzione delle disuguaglianze (SDG 10) non può essere trattato come un'appendice normativa: deve diventare un requisito progettuale, osservabile e verificabile nella pratica.

Nella fase ex-ante, prima di attivare qualsiasi raccolta dati, l'adozione viene impostata come un esercizio di 'definizione progettuale dei vincoli': si chiariscono scopo e limiti (formazione e riesame finale, non disciplina), si mappa la filiera informativa e si traducono rischi in requisiti di esperienza e interfaccia: configurabilità per scenario e fase, assenza di assegnazione di punteggi valutativi, indicatori probabilistici con incertezza visibile, tracciabilità delle fonti e controllo umano effettivo, oltre all'integrazione delle regole già in fase di progettazione (Fig. 9).

Nella fase ex-post il DITL chiede di misurare l'impatto in modo bifocale: da un lato si valutano esiti tecnico-operativi coerenti con l'SDG 9 tra cui la qualità del debriefing rispetto alla ricostruzione degli eventi, all'individuazione di fasi critiche e alla lezione appresa, all'utilità per il formatore e alla stabilità e affidabilità delle stime in condizioni realistiche, dall'altro lato si valutano esiti socio-distributivi coerenti con l'SDG 10: accettabilità e fiducia, per-

### Data processing pipeline



**Fig. 3** | Data processing pipeline: multimodal acquisition, feature extraction, generation of probabilistic indicators, quality checks, and visualisation for debriefing, with pseudonymisation (credit: the Authors; AI-generated with ChatGPT, 2026).

cezione di sorveglianza o pressione valutativa e possibili disparità negli esiti tra gruppi a parità di scenario. Quando emergono criticità è importante il ritorno alla fase progettuale attraverso modifiche all'interfaccia utente, indicatori e sequenza preparazione / riesame finale e rafforzamento dei vincoli di accesso. In questo modo i compromessi con gli SDG 9 e 10 diventano oggetto di iterazione progettuale e di verifica.

**Protocollo di validazione e replicabilità** | Il protocollo mira a rendere replicabile la sperimentazione, valutando l'effetto dell'infrastruttura sull'apprendimento e sulla prevenzione di derive prescrittive o inique. Campione e sessioni (pilot esplorativo): tre partecipanti per scenario, tre sessioni registrate, durata media 6 minuti; obiettivo: fattibilità, integrazione nel riesame finale e accettabilità, non generalizzazione statistica. Il protocollo adotta una strategia a complessità crescente, articolata in tre contesti (aula / ECDIS, plancia cooperativa, colloquio), ciascuno strutturato in preparazione standardizzata, esecuzione dello scenario e riesame finale guidato con HSMS. Per soddisfare il requisito di replicabilità la Tabella 1 sintetizza, per ciascuna fase, i dati in ingresso, gli esiti attesi, i vincoli, i KPI, le soglie di qualità e le condizioni di validità del protocollo.

La validazione coinvolge allievi / partecipanti dei percorsi formativi FAIMM e, nello scenario plancia, team con ruoli distinti. I criteri di selezione e inclusione sono documentati; dove rilevante, vengono registrate le variabili contestuali (esperienza, lingua / accenti, condizioni che possano influire sulla voce o sulla mimica) per valutare l'affidabilità dei risultati e monitorare eventuali disparità. Per ogni sessione vengono acquisiti dati coerenti con la minimizzazione e lo scopo formativo: audio e video, dove previsti, parametri fisiologici non invasivi, oltre a registri di interazione ed eventi di scenario. La filiera è documentata in acquisizione / sincronizzazione, estrazione delle caratteristiche rappresentative e generazione di indicatori probabilistici con metadati di qualità. Il protocollo specifica i punti in cui si applicano la pseudonimizzazione, il controllo degli accessi e la tracciatura delle operazioni.

Per rendere riproducibile la valutazione, la filiera include un controllo qualità che segnala come 'non affidabili' i segmenti con qualità del segnale scarsa, escludendoli dalle sintesi e rendendoli visibili lungo la linea temporale come intervalli a bassa confidenza. I KPI sono definiti come set mini-

mo: KPI tecnici = quota di segmenti affidabili / non affidabili, stabilità intra-sessione, coerenza temporale con marcatore di scenario; KPI formativi = numero di momenti critici discussi e lezioni apprese emerse nel riesame finale, qualità della ricostruzione (ancoraggio a eventi / indicatori); KPI socio-distributivi = accettabilità (fiducia, percezione di sorveglianza / pressione), chiarezza dei limiti d'uso e verifica di possibili differenze sistematiche tra gruppi a parità di scenario.

La valutazione segue il set di KPI sopra definito; il pacchetto di replicazione include: schede scenario, testi della fase di preparazione, schema del riesame finale, configurazioni HSMS per scenario, definizioni degli KPI tecnici e socio-distributivi e procedure di governance (Fig. 10).

**Risultati preliminari e discussione critica** | I risultati preliminari sono coerenti con la natura sperimentale del caso On Watch: l'obiettivo non è dimostrare un'accuratezza assoluta dell'inferenza, bensì verificare se l'infrastruttura proposta riesca a generare valore formativo senza innescare effetti regressivi. Sul piano tecnico-operativo, le sessioni confermano che la qualità dei segnali non è uniforme e dipende da condizioni pratiche quali rumore, occlusioni, distanza, illuminazione o simultaneità di conversazioni. Ne consegue la necessità di gestire segmenti non affidabili e di trattare i risultati come indicatori probabilistici, associandoli a metadati di qualità, riducendo così la falsa precisione e gli automatismi interpretativi. L'ancoraggio temporale degli indicatori facilita la ricostruzione dei passaggi chiave e rende discutibile la relazione tra gli andamenti (attivazione / carico) e gli eventi operativi. L'effetto osservato è coerente con il ruolo della memoria e della plasticità sinaptica: la ricostruzione guidata, ancorata a marcatori temporali, facilita il potenziamento di tracce pertinenti e la ristrutturazione di schemi operativi, rispetto a ricostruzioni vaghe o puramente verbali.

Sul piano formativo l'evidenza più consistente riguarda la qualità della fase di riesame finale: l'uso del prototipo HSMS come supporto ha consentito di strutturare la discussione su momenti critici identificabili, riducendo ricostruzioni vaghe e favorendo le lezioni apprese. Nei contesti cooperativi, la lettura per ruolo e per fase supporta la discussione sulla coordinazione, senza produrre giudizi individuali; la presenza di evidenze temporali facilita una narrativa sull'errore come risorsa, più che come colpa.

Dal punto di vista socio-distributivo l'accettabilità dipende dalla cornice interpretativa nella fase della preparazione e dalle proprietà dell'interfaccia: quando sono esplicitati scopo e limiti e quando le visualizzazioni restano non prescrittive, la percezione di sorveglianza si riduce. In contesti percepiti come valutativi aumenta la sensibilità allo stigma e all'uso improprio, confermando la necessità di configurazioni più restrittive e di un governo del sistema fin dalla fase di progettazione. Un rischio da monitorare riguarda possibili effetti iniqui 'silenziosi': differenze nei segnali vocali e negli stili comunicativi possono produrre falsi alert differenziali; il contesto DITL impone di trattare tali evidenze come criteri di redesign (riduzione della salienza in condizioni di bassa qualità e rafforzamento della triangolazione).

**Limiti, barriere e criticità** | L'introduzione di HSMS nella formazione nei domini critici per la sicurezza incontra barriere che non sono solo tecniche, ma soprattutto socio-organizzative. Il contributo adotta intenzionalmente un punto di vista 'orientato al design': l'attenzione è posta sull'architettura dell'esperienza, sulle scelte di interfaccia e sul governo del sistema fin dalla fase di progettazione, più che sulla competizione tra modelli di inferenza. Ne consegue un limite: i risultati non vanno letti come una validazione definitiva dell'accuratezza o della capacità diagnostica. Inoltre la generalizzabilità dei risultati dipende dal contesto di impiego (simulazione o contesti controllati), dall'assetto di acquisizione e dalla cultura organizzativa in cui il sistema viene introdotto.

La barriera più critica è culturale: in molte organizzazioni la formazione si muove su un crinale tra apprendimento e valutazione; il monitoraggio può essere percepito come sorveglianza, soprattutto in assenza di una cultura dell'errore come risorsa. Il rischio di derive legate all'ipersorveglianza rende indispensabile esplicitare scopo e limiti, incorporando vincoli (Eurofound, 2020; Ajunwa, Crawford and Schultz, 2017). Una seconda barriera riguarda la dimensione distributiva: differenze culturali e linguistiche, accenti e stili comunicativi, neurodiversità o condizioni individuali possono alterare la qualità del segnale e l'interpretabilità dei risultati, con il rischio di stigma e disparità. Per questo l'equità e l'accessibilità cognitiva della interfaccia utente devono essere integrate nella fase progettuale e monitorate tramite KPI socio-distributivi (Buolamwini and Gebu, 2018; Raji and Buolamwini, 2019).

Sul piano economico-operativo la diffusione richiede infrastrutture, manutenzione e competenze interdisciplinari; un'adozione parziale, in particolare senza governance e senza formazione dei formatori, può aumentare i rischi. Esistono inoltre barriere normative e di responsabilità relative alla gestione dei dati sensibili, alla chiarezza sull'accesso e sulla conservazione e alla definizione delle responsabilità d'uso. La trasferibilità è plausibile a condizioni minime: riesame finale strutturato, governance integrata nel progetto e interfacce non prescrittive ove siano chiaramente indicate l'incertezza e la tracciabilità dei dati (European Union, 2016, 2024).

**SDG: benefici, sinergie e compromessi** | Il caso On Watch consente di leggere in termini operativi il profilo di benefici, sinergie e compromessi connesso all'introduzione di HSMS non prescrittivi nella formazione marittima. Sul piano dei benefici diretti il contributo si colloca anzitutto nell'orizzonte dell'SDG 9, poiché propone una forma di innovazione infrastrutturale della formazione fondata su strumenti capaci di supportare in modo verificabile le fasi di preparazione e di riesame finale, senza tradursi in automatismi decisionali o in dispositivi valutativi.

In relazione all'SDG 10 il valore del framework risiede nella riduzione dei rischi distributivi attraverso un'impostazione di governance integrata fin dalla progettazione, che delimita chiaramente gli scopi del sistema, esclude l'attribuzione di punteggi valutativi, rende visibili l'incertezza e la tracciabilità delle inferenze, regola l'accesso ai dati e introduce audit e indicatori socio-distributivi per monitorare possibili effetti iniqui. A ciò si aggiunge un contributo rilevante all'SDG 4, nella misura in cui il sistema migliora la qualità didattica del riesame finale e rafforza l'apprendimento dall'errore, e all'SDG 8, poiché tutela il lavoro dignitoso limitando le pressioni valutative e gli usi disciplinari dei dati e preservando l'autonomia, la dignità e condizioni eque nei processi di formazione e valutazione.

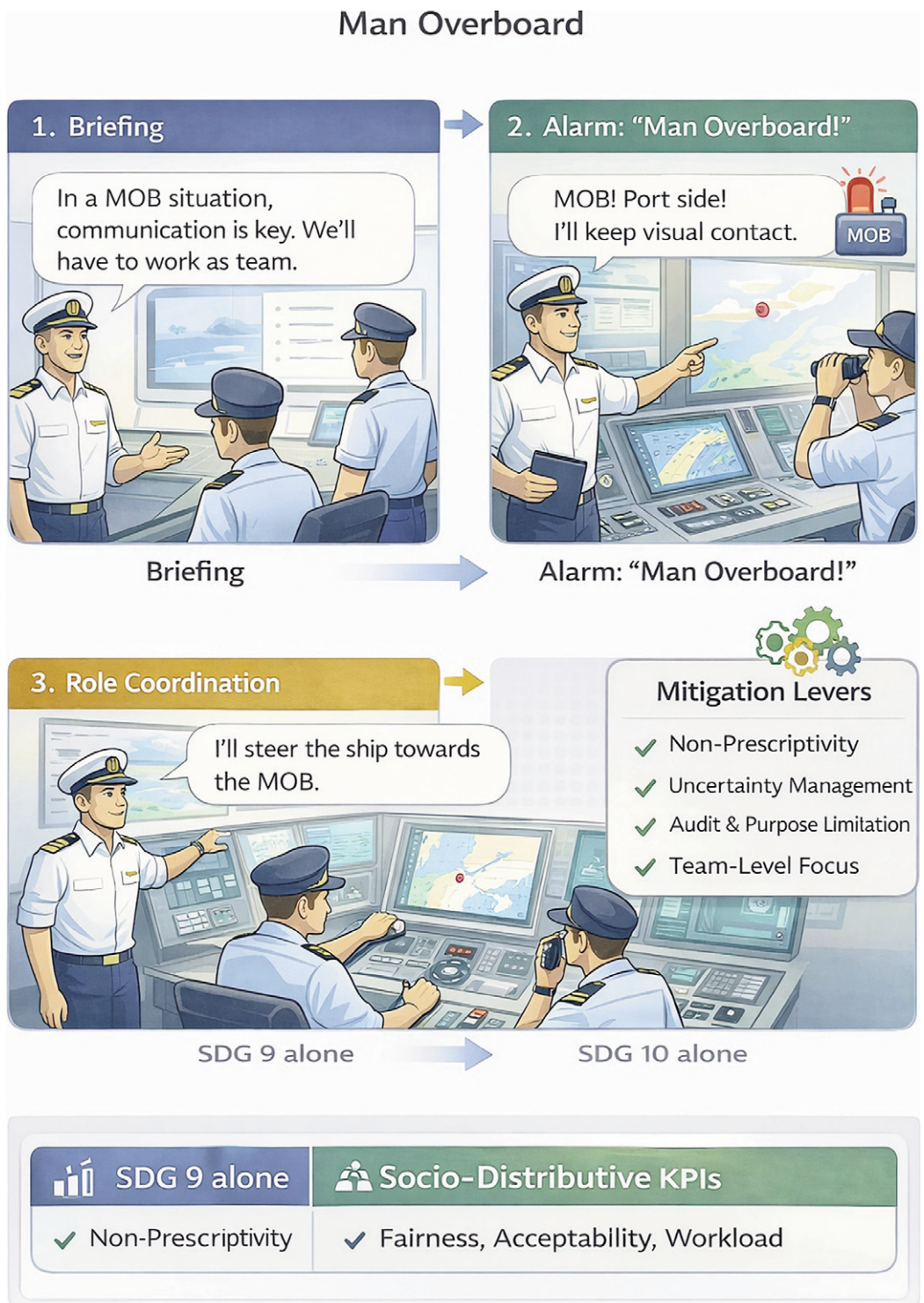
Accanto a questi effetti diretti emergono ricadute indirette che riguardano ulteriori obiettivi di sviluppo sostenibile. In particolare il framework può

contribuire all'SDG 3, sostenendo il benessere degli operatori e la prevenzione del sovraccarico; agli SDG 12 e 13 attraverso una logica di sobrietà del dato fondata su minimizzazione, contenimento dei tempi di conservazione e riduzione degli sprechi digitali; all'SDG 16 mediante il rafforzamento di accountability e trasparenza organizzativa; all'SDG 17 grazie alla costruzione di partnership tra università, accademia e stakeholder del settore.

Tali benefici non eliminano tuttavia la presenza di compromessi e rischi residui, che possono incidere anche su altri SDG, in particolare quando persistano accessi diseguali alle infrastrutture, distorsioni nei segnali multimodali o costi computazionali non trascurabili. In questo senso le possibili criticità che possono riverberarsi sugli SDG 1, 2, 5, 7, 11, 14 e 15 non vanno considerate come effetti ester-

ni o secondari, ma come condizioni da governare progettualmente. Proprio qui il framework Designer-in-the-Loop mostra la propria rilevanza, poiché rende tali compromessi osservabili, misurabili e governabili sia ex ante sia ex post, attraverso soglie di qualità, audit, verifiche iterative e aggiustamenti progettuali continui (Fig. 11). Nel loro insieme i risultati mostrano dunque una chiara convergenza con gli SDG 3, 4, 8, 12 e 16, ma anche la necessità di presidiare con attenzione le tensioni legate a sorveglianza, pressione valutativa, pregiudizio e disparità, affinché l'innovazione digitale resti coerente con un paradigma formativo centrato sulla persona.

**Conclusioni** | Questo contributo propone di spostare il baricentro del dibattito su AI e formazione nei domini critici della sicurezza: non è da chiedersi



**Fig. 4** | Storyboard of the 'man overboard' scenario: briefing, alarm, role coordination, and debriefing as the structure of the training experience (credit: the Authors; AI-generated with ChatGPT, 2026).

'quanto è accurato il riconoscimento dello stato umano', ma 'quale infrastruttura socio-tecnica rende l'innovazione verificabile, utile e giusta'. L'analisi del caso On Watch mostra che l'introduzione di un Human State Monitoring System in contesti formativi ad alta criticità produce valore soprattutto quando è progettata come architettura dell'esperienza – preparazione, scenario e riesame finale – e quando le interfacce sono concepite come strumenti di costruzione del significato, anziché come dispositivi di valutazione.

L'aspetto più rilevante emerso dallo studio è che la sostenibilità dell'innovazione non si gioca nella sola performance del modello, bensì nella capacità del progetto di evitare la deriva verso la sor-

veglianza e le disuguaglianze. Per tale motivo il governo del sistema fin dalla fase di progettazione, che garantisce assenza di ranking, chiara e limitata definizione degli scopi, accesso ai dati, incertezza visibile e tracciabilità delle fonti, è un requisito progettuale intrinseco.

In questo contesto l'innovazione infrastrutturale può produrre impatti positivi solo se accompagnata da vincoli che proteggano la dignità ed l'equità. La sua sostenibilità richiede il passaggio da un 'io progettuale' a un 'noi progettuale': il prototipo HSMS è, in tal senso, un artefatto che prende significato entro pratiche sociali (preparazione e riesame finale) e regimi di governance negoziati con formatori e discenti; solo un'esecuzione co-

progettata può conciliare l'utilità (SDG 9) e l'equità (SDG 10).

La trasferibilità dell'approccio non è legata a una specifica manovra o a un simulatore, ma a tre elementi replicabili: la progettazione dell'esperienza formativa come infrastruttura socio-tecnica centrata sulla preparazione e sul riesame finale; un set di scelte di 'interaction design' che mantengano l'interpretazione aperta e verificabile; il framework Designer-in-the-Loop come metodo ex-ante ed ex-post che integri KPI tecnico-operativi e socio-distributivi. Gli sviluppi futuri riguardano: a) robustezza e gestione delle variazioni nel set di dati; b) consolidamento delle metriche e delle procedure di equità contestuale e delle versioni a bassa inva-



Fig. 5 | Full mission bridge simulator used in the cooperative scenario for analysing coordination dynamics and team cognitive workload (credit: the Authors, 2026).

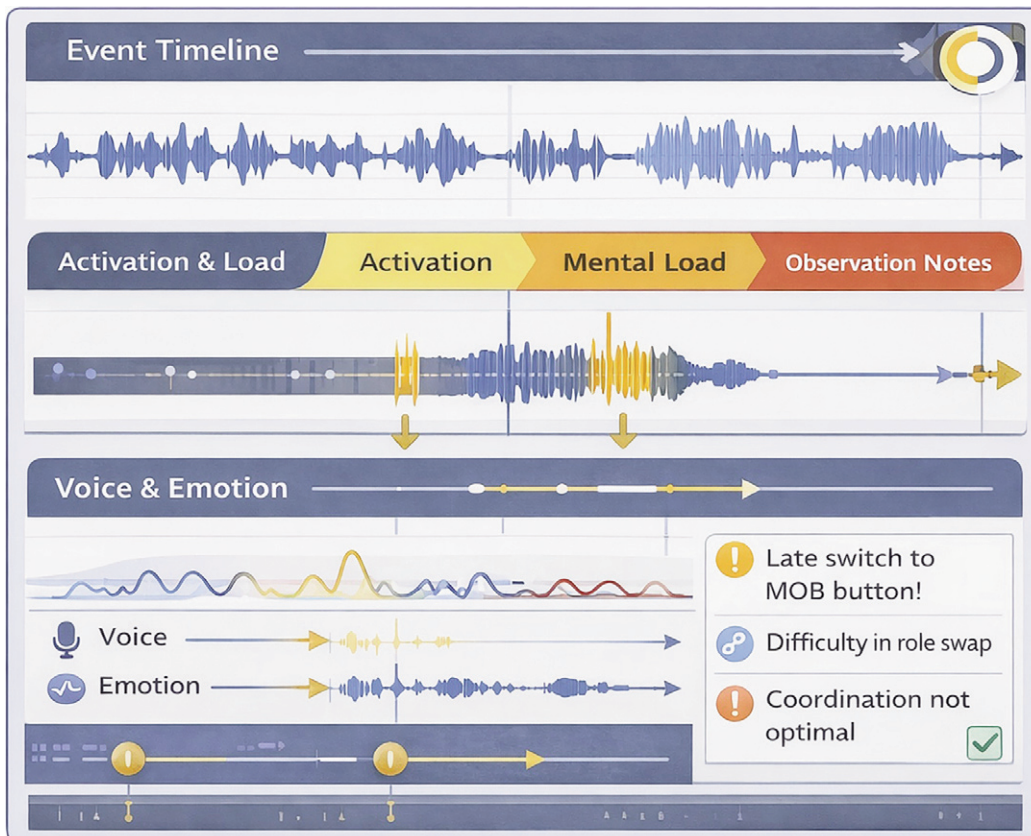


Fig. 6 | Non-prescriptive debriefing interface: event timeline, activation / workload indicators, vocal / emotional signals, and notes panel for guided discussion (credit: the Authors; AI-generated with ChatGPT, 2026).

**Fig. 7** | Configuration model by scenario and phase: enabling functions across briefing, live, and debriefing in the three settings of ECDIS, bridge, and interview (credit: the Authors; AI-generated with ChatGPT, 2026).

| Scenario configuration model |          |          |              |            |
|------------------------------|----------|----------|--------------|------------|
|                              | Briefing | Scenario | Live Session | Debriefing |
| ECDIS Simulation             | ✓        | ✓        | ✓            | ✓          |
| Bridge Simulation            | ✓        | ✓        | ✓            | ✓          |
| Interview Exercise           | ✓        | ✓        | ✓            | ✓          |

sività; c) formalizzazione delle pratiche di verifica, segnalazione degli incidenti e sospensione dell'uso in caso di criticità. In prospettiva la comunità scientifica è chiamata a valutare gli HSMS nella formazione non solo con metriche predittive, ma con un impianto che consideri esperienza, contesto e governance come variabili costitutive.

In safety-critical domains, training is a socio-technical infrastructure: spaces (classrooms, simulators, bridges), artefacts (checklists, procedures), and phases – especially those related to preparation and final review – that enable learning, error prevention, and the development of a safety culture. In the maritime sector, performance depends on cooperation, situational awareness, and cognitive load management; for this reason, the training infrastructure enables operational continuity and risk reduction. In this context, digitalisation (sensors, data analysis, simulation, AI/IoT) can improve the quality and traceability of training (SDG 9), but it can also trigger privacy violations, biases, and inequalities (SDG 10) if not properly designed and governed (Salas et alii, 2012; Aguinis and Kraiger, 2009; ISO, 2019; Fig. 1). This makes it crucial to design tools that support reflection, rather than mere passive compliance.

From this perspective, the paper assumes that the key issue is not whether to adopt data-driven technologies in training for safety-critical domains, but how to design them so that innovation remains consistent with the principles of Industry 5.0 and with a human-centred paradigm: technology must enhance the system's capacity to learn from experience and prevent error, without turning the training infrastructure into a control apparatus. It follows that the value of Human State Monitoring Systems (HSMS) does not depend solely on inference accuracy, but also on the ability of the infrastructure to make explicit the purposes, limits, and responsibilities throughout the entire usage cycle. For this reason, the design must incorporate technical and organisational constraints that prevent incompatible reuse and prescriptive interpretations.

The underlying hypothesis is that this balance can only be achieved by shifting the focus from ex post compliance (rules and policies added afterwards) to the governance of the integrated system from the design stage: minimisation and separation of identifiable data, informed and revocable consent, the subject's right to opt out of monitoring, retention policies, traceability and verification,

and explicit and verifiable limits of use throughout the system lifecycle. In other words, governance is not a 'regulatory' section at the margins, but a design component that shapes interfaces, information flows, and operational phases within which the technology acquires meaning (Mittelstadt, 2019; NIST, 2023; OECD, 2019). Within this framework are situated the On Watch case study and the Human State Monitoring System (HSMS) prototype, developed and tested in progressively complex educational and simulation contexts (classroom with ECDIS simulator – Electronic Chart Display and Information System, bridge simulation with cooperative roles, interviews / selection), in collaboration with the University of Genoa and the FAIMM Italian Merchant Marine Academy, within the ITS 4.0 programme (MIM, 2025). HSMS integrates multimodal signals (voice, facial and postural cues, breathing, text, and operational interactions) and translates them into probabilistic state indicators (e.g., activation and load) to support informed interpretation during the preparation and final review phases. To avoid disciplinary drift and inequalities, the system is deliberately non-prescriptive: it does not produce individual rankings, does not enable league tables, does not automate operational decisions, and is not designed for worker surveillance purposes, but provides visualisations and signals configurable by scenario and phase, with traceability of sources and human oversight (Ronca et alii, 2023; Liu et alii, 2020; Figg. 2, 3).

The originality of the paper is twofold. First, it proposes an interpretation of training safety as a design and management infrastructure, where spaces, interfaces, and phases constitute an inseparable unit: it is within this unit that errors, information overload, false sense of security, and coordination failures are either produced or reduced. Second, it formalises an operational Designer-in-the-Loop framework to make measurable and verifiable the trade-off between digital innovation and the reduction of inequalities: the adoption of HSMS is governed ex ante (through the assessment of risks of misuse, biases, and access barriers) and ex post (through technical-operational and socio-distributive KPIs), so that trade-offs do not remain implicit but become objects of design, verification, and iterative improvement.

In light of these premises, the paper aims to clarify how the use of HSMS in maritime training can enhance the quality of the final review and learning from error without producing regressive effects in terms of surveillance or inequality. In particular, the paper: a) critically frames the debate on

Affective Computing and HSMS in safety-critical domains, highlighting epistemic limits and risks of misuse; b) describes On Watch as a case study, foregrounding the experience architecture, interface choices, and usage scenarios; c) formalises the Designer-in-the-Loop methodological framework with requirements and technical-operational and socio-distributive KPIs for a replicable and transferable ex ante and ex post evaluation. The paper is structured as follows: state of the art; case study; experience architecture; interaction and interfaces; Designer-in-the-Loop framework; protocol and KPIs; preliminary results and discussion; barriers / limitations and transferability; SDGs; conclusions.

The research questions are: 1) which design and governance conditions allow the integration of HSMS into training in safety-critical domains without turning them into prescriptive or disciplinary devices; 2) which experience architecture and interface choices make the system verifiable, controllable, and non-prescriptive across different scenarios; 3) which minimum set of requirements, quality thresholds, and KPIs enables a replicable evaluation and supports transferability to other safety-critical domains.

To address these questions, the paper combines a critical review of the international literature with the analysis of the On Watch case and the operational description of the experimental protocol, making explicit the requirements, constraints, and indicators necessary to replicate the study in other contexts.

**State of the art** | The adoption of Affective Computing and Human State Monitoring Systems in training within safety-critical contexts intersects three strands: multimodal inference of affective / cognitive states; training and performance assessment in high-criticality environments; and studies on monitoring and surveillance, equity, and data governance. The review is intentionally critical: the objective is not to legitimise automatic scoring, but to clarify epistemic limits, risks of misuse, and design requirements needed to make such systems verifiable and transferable (Picard, 1997; Poria et alii, 2017; D'Mello and Kory, 2015; Ajunwa, Crawford and Schultz, 2017; Eurofound, 2020; Zuboff, 2019). From this perspective, the multimodal nature and the situated character of perceptual / decision-making processes suggest treating results as probabilistic and contextual estimates, supporting the shared construction of meaning in the preparation and final review phases, rather than as

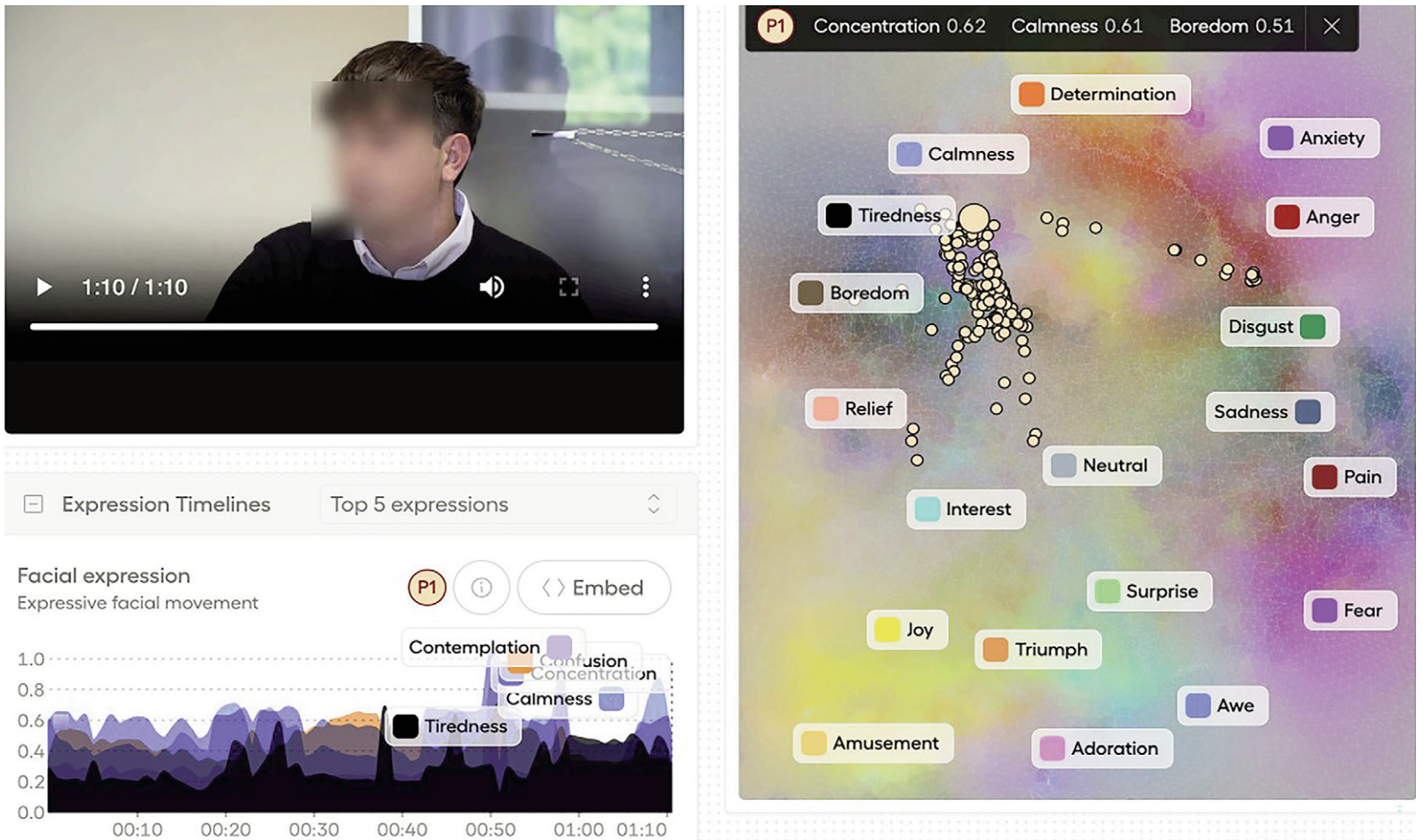


Fig. 8 | ECDIS workstation used in the On Watch pilot for individual simulation and integration of HSMS indicators in debriefing (credit: the Authors, 2026).

prescriptive verdicts (Gazzaniga, 2005). The literature on emotion detection confirms the shift towards multimodal approaches. However, it highlights the heterogeneity of protocols, limited comparability, misalignment between datasets, and generalisation issues, all of which require caution against false precision in safety-critical domains (Zeng et alii, 2009; Koelstra et alii, 2012). Within the human factors strand, the usefulness of HSMS emerges primarily when they are integrated into the pedagogical function of preparation and final review, where the quality of reflection depends more on artefacts, roles, and communicative sequences than on the mere availability of data (Hart and Staveland, 1988; Endsley, 1995; Hancock and Matthews, 2019).

At the application level, several international best practices in simulator-based training show that the use of workload and stress indicators is more robust when anchored to scenarios, markers, and a structured final review, and when measurements are treated as contextual evidence rather than as individual scores. Recent studies in bridge simulation highlight the value of triangulating psychophysiological data with operational observations to interpret workload, stress, and attention, and to critically discuss the limits and reliability of the measures (Wulvik, Dybvik and Steinert, 2020; Ronca et alii, 2023; Xue et alii, 2024). These cases suggest transferable selection criteria: a) high-criticality scenario with explicit roles; b) multimodal data collection with quality checkpoints; c) non-prescriptive feedback in the final review; d) governance that limits disciplinary reuse.

Finally, in the debate on objects and services enhanced by artificial intelligence, attention shifts towards the user experience and design responsibilities: configurability, effective human control, and integration of the data supply chain become conditions for avoiding regressive effects and making systems verifiable (Arquilla and Paracolli, 2023; Osello et alii, 2024; Villani, Romagna and Oddi, 2024), while studies on error design, alienation, and evaluative pressure warn against the risk of transforming training tools into control devices; contributions on the inclusivity of multimodality also suggest interpreting signals as situated evidence to be discussed in the final review rather than as mere labels (Porfirione, Ferrari Tumay and Leggiero, 2024; Vacanti et alii, 2025; Rosmino, 2024; Arquilla and Caruso, 2025; Zehr Gantz et alii, 2025). This underscores the need to prevent surveillance uses across the entire data supply chain, through the governance of an integrated system from the design stage and through socio-distributive evaluations. The AI Act reinforces this approach regarding risk, transparency, and organisational accountability (European Union, 2024).

**Case study: On Watch and HSMS** | The paper is grounded in the On Watch case study and the HSMS prototype, conceived as an experimental platform to support training and final review in safety-critical contexts. The pilot experimentation was developed with FAIMM – Italian Maritime Academy and the University of Genoa within the ITS 4.0 Challenge 2025 framework (MIM, 2025). The experimentation follows a strategy of increasing

complexity across three contexts: a) classroom with an ECDIS simulator; b) cooperative bridge with differentiated roles; c) interview / selection, with more restrictive configurations to reduce stigma and the risk of misuse. In the bridge scenario, the main test uses a ‘man overboard’ emergency with multiple roles (Fig. 4); the ‘full mission bridge’ simulator used for the cooperative scenario is illustrated in Figure 5.

Within On Watch, the HSMS prototype integrates multimodal signals (voice, facial and postural cues, respiratory parameters, text, and interactions). It translates them into probabilistic indicators (for example, activation and workload) to support the preparation and final review phases. The system is deliberately non-prescriptive: it does not produce rankings nor automate decisions, but provides configurable visualizations with uncertainty and traceability (Ronca et alii, 2023; Wulvik, Dybvik and Steinert, 2020; Xue et alii, 2024). The use of sensitive signals requires minimisation, pseudonymisation, controlled access, defined data retention periods, and verification procedures, together with clear information and revocable consent, as well as the option not to participate in monitoring; these constraints are embedded in the preparation and final review phases and in the user interface, according to an integrated system governance logic from the design stage.

**Architecture of the training experience** | In safety-critical contexts, the effectiveness of training depends less on the single exercise and more on the architecture of the experience that frames it: pre-

paration, execution, guided reflection. In the On Watch case, this architecture was designed as an intentional sequence of preparation, scenario, and final review, in which multimodal monitoring is not superimposed on the activity, but positioned at specific points in the flow to maximise training utility and minimise regressive effects (Fanning and Gaba, 2007; Rudolph et alii, 2006; Cheng et alii, 2014). From this perspective, perception-action and post-event reflection capitalise on the brain's tendency to 'fill in' gaps and reconstruct dynamic scenarios, for example, through saccadic suppression and integration with memory and expectations. Hence, the choice to concentrate HSMS on the final review, a phase in which heterogeneous signals can be temporally 're-anchored' to operational events and decisions, reduces the 'false precision' of real-time interpretation and enhances human control.

In the preparation phase, the trainer aligns learning objectives and the scenario, clarifies roles and responsibilities (particularly in cooperative contexts), and explicitly states the conditions of data use as part of the governance of the integrated system from the design stage: training and research purposes, exclusion of disciplinary uses, minimization, pseudonymisation, controlled access, possibility of withdrawal and the option not to participate in monitoring, data retention periods, and traceability. Preparation also serves as a 'psychological contract' to reduce fears of surveillance and to guide the interpretation of the system's outputs as support for the final review. During the scenario, the objective is not to 'measure emotion' as a label, but to observe how emotional and cognitive states emerge and vary under conditions of operational pressure and cooperation. Data collection is structured as a synchronised pipeline for post-session analysis: operational behaviour and interactions; facial and postural signals; vocal and prosodic parameters; verbal content; and communication dynamics. To avoid operational interference, HSMS may generate attention cues configurable by phase and scenario, designed to reduce false alarms without diverting the operator's behaviour.

The final review is the phase in which the infrastructure becomes effectively 'empathetic': a situated reading of the experience is constructed, and decisions, communications, and critical moments are brought into relation. HSMS fully come into play in this phase because they enable anchoring the discussion to temporal and multimodal

evidence without turning them into individual scores. From the perspective of preventing drift towards surveillance logics, the final review also serves as a governance safeguard, defining what can be discussed, who can access the visualisations, which data are relevant to the educational purpose, and how they are stored or deleted.

**Interaction and interfaces** | In the On Watch case, interaction is designed with an explicit assumption: in a safety-critical domain, the user interface must not 'close' interpretation, but enable a verifiable reading of data during preparation and final review. In line with this principle, HSMS interfaces are conceived as interpretative tools aimed at making temporal trends and discontinuities visible, linking them to operational and communicative events, making uncertainty and the provenance of inference explicit, and preserving effective human control, thus preventing prescriptive or disciplinary drift (Parasuraman and Riley, 1997).

The HSMS prototype adopts a limited set of views, designed for the final review (and only marginally for the real-time phase), avoiding synthetic one-dimensional KPIs that tend to turn into evaluation. User interface design may leverage Gestalt principles (proximity, similarity, good continuity, and closure) to group events, markers, and indicators along the timeline, reducing cognitive load and supporting the recognition of patterns without translating into score attribution. For example, grouping by role and phase (similarity), aligning markers with continuous visual trajectories (good continuity), and closing informational gaps with the trainer's notes (closure) support discussion without sliding into rankings.

The default view is a timeline with probabilistic indicators (for example, activation and workload) and scenario markers; an events view allows, by selecting a time interval, the visualisation of operational actions, communications, and role changes, with the possibility of trainer annotations. A multimodal, evidence-based view provides access, when necessary and authorised, to the sources that contributed to the estimate (vocal and prosodic features and aggregated facial and postural cues), accompanied by timestamps and quality metadata. In cooperative contexts, the user interface supports a role-based reading, avoiding competitive comparisons (Fig. 6).

A key requirement is configurability by scenario and phase: ECDIS scenario (with emphasis on

workload and decision-making), cooperative bridge (with emphasis on coordination and communication), and interview / selection (with more restrictive configurations to reduce the risk of stigmatisation). Configuration affects temporal granularity, source visibility, and access rules – that is, who can see what and when. This approach constitutes a governance measure of the integrated system from the design stage: the user interface embeds usage limits and access constraints as an integral part of the experience, making functional drift more difficult (Fig. 7). The ECDIS environment used in the pilot is shown in Figure 8.







To ensure traceable accountability, the HSMS prototype makes the provenance and quality of modalities, the transformations (from signal to extracted feature to indicator), and access logs verifiable. Non-prescriptiveness derives from the absence of rankings, the visibility of uncertainty, and the control of source visibility.

**Designer-in-the-Loop framework** | The Designer-in-the-Loop (DITL) framework proposed in this paper stems from a premise typical of design research applied to complex systems: when a technology enters a safety-sensitive context, it does not simply 'add' to an existing process; rather, it redefines the socio-technical infrastructure that supports learning. In the case of an HSMS, this means intervening simultaneously on rituals (preparation / final review), artefacts (dashboards, logs, reports), expectations (what it is legitimate to infer from data), and power relations (who observes, who is observed, who decides, who can challenge). For this reason, the gap between innovation and the reduction of inequalities cannot be treated as a regulatory appendix: it must become a design requirement, observable and verifiable in practice.

In the ex-ante phase, before activating any data collection, adoption is framed as an exercise in the 'design definition of constraints': purposes and limits are clarified (training and final review, not disciplinary use), the information pipeline is mapped, and risks are translated into experience and interface requirements: configurability by scenario and phase, absence of evaluative scoring, probabilistic indicators with visible uncertainty, traceability of sources, and effective human control, in addition to the integration of rules already at the design stage (Fig. 9).

In the ex-post phase, DITL requires impact to be measured in a bifocal way: on the one hand,

### Governance-by-Design in Safety-Critical Environments

| Risk  | Design Requirement    | Verification Evidence   | When Assessed     |
|---|-----------------------|---|-------------------|
|  | ✓ No ranking          |  Audit logs              | Ex-ante / Ex-post |
|  | ✓ Uncertainty visible |  UI constraint           | Ex-ante / Ex-post |
|  | ✓ Role-based access   |  Socio-distributive KPIs | Ex-post           |

**Fig. 9** | Governance-by-design: risk mapping (surveillance drift, bias, stigma / pressure), design requirements, verification evidence, and timing of assessment, ex-ante / ex-post (credit: the Authors; AI-generated with ChatGPT, 2026).

| Phase (DITL)       | Input data                                  | Operational steps / parameters   | Outputs                                   | Constraints & governance                       | KPIs (operational)  | Quality gates / thresholds   | Validity & replication conditions           |
|--------------------|---|--|---|--|---|--|---|
| Ex-ante scoping    | Scenario briefs; roles; data-map            | Define purpose limitation; select modalities per scenario; configure views | Approved protocol + configs               | No ranking; opt-out; access control; retention | Protocol completeness; risk register coverage                                 | N/A  | Same scripts / configs used across sessions |
| Briefing           | Briefing script; consent form               | Explain purpose/limits; assign roles; calibrate devices; verify opt-out    | Participants ready; log consent           | Consent revocable; pseudonymisation planned    | Briefing completeness; comprehension check                                    | Consent recorded; devices pass check                                       | Standardised briefing text / time           |
| Scenario execution | ECDIS task / Bridge MOB / Interview prompts | Run 6-min scenario; capture audio / video / biosignals; mark events        | Synced multimodal dataset + event markers | No live scoring; limited live cues             | Data completeness; event marker accuracy                                      | Unreliable segments flagged if missing modality >30% or audio clipping >5% | Same scenario script; same room / setup     |
| Debriefing         | Timeline + markers + indicators             | Review critical moments; annotate; discuss lessons learned                 | Debrief notes; lessons learned list       | Role-based visibility; evidence view on-demand | no. of critical moments discussed; lessons learned count; anchoring to events | Show uncertainty + quality metadata  | Same debrief guide; same UI views           |
| Ex-post evaluation | Questionnaires + logs                       | Compute KPIs; check disparities; update design                             | KPI report + redesign actions             | Audit log; purpose limitation                  | Acceptability; perceived surveillance; disparity checks                       | Flag if group-level differences > pre-defined margin                       | Report template; same questionnaires        |

Tab. 1 | Methodological synthesis for replicability (On Watch / HSMS).

technical-operational outcomes consistent with SDG 9 are assessed, including the quality of the debriefing in relation to the reconstruction of events, the identification of critical phases, and lessons learned, as well as usefulness for the trainer and the stability and reliability of estimates under realistic conditions; on the other hand, socio-distributive outcomes consistent with SDG 10 are evaluated, including acceptability and trust, perceived surveillance or evaluative pressure, and possible disparities in outcomes between groups under the same scenario. When critical issues emerge, returning to the design phase is essential through modifications to the user interface, indicators, and the preparation / final review sequence, as well as strengthening access constraints. In this way, trade-offs with SDGs 9 and 10 become the object of design iteration and verification.

**Validation and replicability protocol** | The protocol aims to make the experimentation replicable by assessing the effect of the infrastructure on learning and on the prevention of prescriptive or inequitable drift. Sample and sessions (exploratory pilot): three participants per scenario, three recorded sessions, average duration 6 minutes; objective: feasibility, integration into the final review, and acceptability, not statistical generalisation. The protocol adopts a strategy of increasing complexity, structured across three contexts (classroom/ECDIS, cooperative bridge, interview), each organised into standardised preparation, scenario execution, and a guided final review with HSMS. To meet the requirement of replicability, Table 1 summarises, for each phase, the input data, expected outcomes, constraints, KPIs, quality thresholds, and protocol validity conditions.

Validation involves trainees / participants from FAIMM training programmes and, in the bridge scenario, teams with distinct roles. Selection and inclusion criteria are documented; where relevant, contextual variables (experience, language/accent, conditions that may affect voice or facial expression) are recorded to assess the reliability of results and monitor potential disparities. For each session, data consistent with minimisation and training purposes are collected: audio and video (where applicable), non-invasive physiological pa-

rameters, interaction logs, and scenario events. The pipeline is documented in acquisition / synchronisation, feature extraction, and generation of probabilistic indicators with quality metadata. The protocol specifies the points at which pseudonymisation, access control, and operation logging are applied.

To ensure reproducibility of the evaluation, the pipeline includes a quality control process that flags segments with poor signal quality as ‘unreliable,’ excluding them from summaries and making them visible along the timeline as low-confidence intervals. KPIs are defined as a minimum set: technical KPIs = proportion of reliable / unreliable segments, intra-session stability, temporal consistency with scenario markers; training KPIs = number of critical moments discussed and lessons learned emerging in the final review, quality of reconstruction (anchoring to events / indicators); socio-distributive KPIs = acceptability (trust, perception of surveillance / evaluative pressure), clarity of usage limits, and verification of possible systematic differences between groups under the same scenario.

Evaluation follows the KPI set defined above. The replication package includes: scenario sheets, preparation-phase scripts, final-review structure, HSMS configurations per scenario, definitions of technical and socio-distributive KPIs, and governance procedures (Fig. 10).

**Preliminary results and critical discussion** | The preliminary results are consistent with the experimental nature of the On Watch case: the objective is not to demonstrate absolute inference accuracy but rather to verify whether the proposed infrastructure can generate training value without triggering regressive effects.

From a technical-operational perspective, the sessions confirm that signal quality is not uniform and depends on practical conditions such as noise, occlusions, distance, lighting, or simultaneous conversations. This leads to the need to manage unreliable segments and to treat results as probabilistic indicators, associating them with quality metadata, thereby reducing false precision and interpretative automatisms. The temporal anchoring of indicators facilitates the reconstruction of key steps and makes the relationship between trends

(activation / workload) and operational events open to discussion. The observed effect is consistent with the role of memory and synaptic plasticity: guided reconstruction, anchored to temporal markers, facilitates the strengthening of relevant traces and the restructuring of operational schemas, compared to vague or purely verbal reconstructions. From a training perspective, the most consistent evidence concerns the quality of the final review phase: the use of the HSMS prototype as support enabled the discussion to be structured around identifiable critical moments, reducing vague reconstructions and fostering lessons learned. In cooperative contexts, role- and phase-based reading supports discussion on coordination without producing individual judgments. The presence of temporal evidence facilitates a narrative of error as a resource rather than as a source of blame.

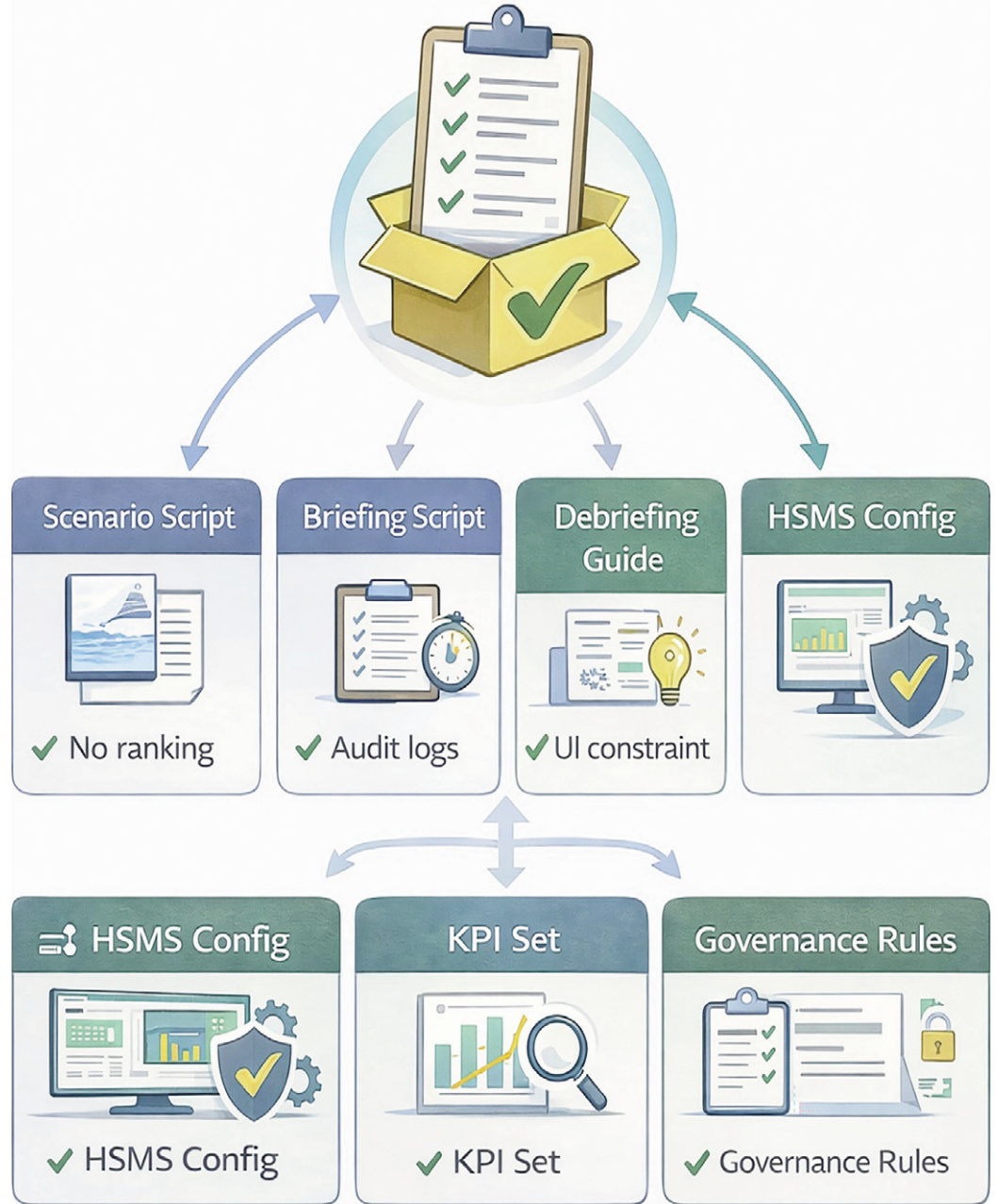
From a socio-distributive perspective, acceptability depends on the interpretative framing in the preparation phase and on interface properties. When purpose and limits are explicitly stated, and visualisations remain non-prescriptive, perceptions of surveillance decrease. In contexts perceived as evaluative, sensitivity to stigma and misuse increases, underscoring the need for more restrictive configurations and for system governance from the design stage. A risk to be monitored concerns potential ‘silent’ inequitable effects: differences in vocal signals and communication styles may produce differential false alerts; the DITL context requires treating such evidence as criteria for redesign (reducing salience under low-quality conditions and strengthening triangulation).

**Limitations, barriers, and critical issues** | The introduction of HSMS training in safety-critical domains encounters barriers that are not only technical but, above all, socio-organisational. This paper intentionally adopts a design-oriented perspective: attention is focused on the architecture of the experience, interface choices, and system governance from the design stage, rather than on competition between inference models. This implies a limitation: the results should not be interpreted as a definitive validation of accuracy or diagnostic capability. Moreover, the generalisability of the re-

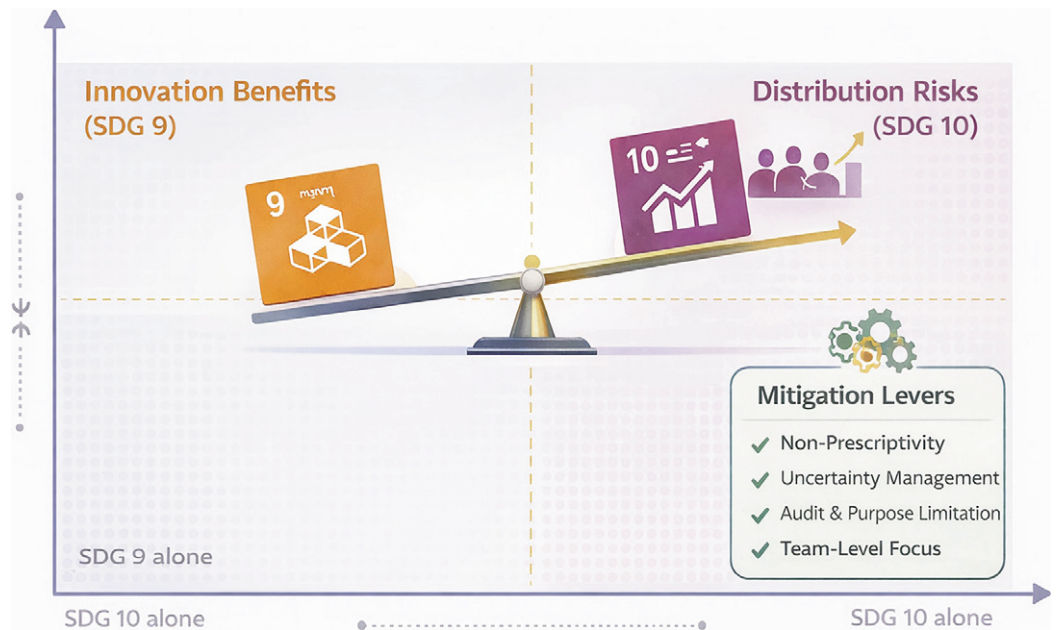
sults depends on the context of use (simulation or controlled environments), the acquisition setup, and the organisational culture into which the system is introduced.

The most critical barrier is cultural: in many organisations, training operates along a boundary between learning and evaluation; monitoring may be perceived as surveillance, especially in the absence of a culture that treats error as a resource. The risk of drift towards hypersurveillance makes it essential to explicitly define purpose and limits, incorporating constraints (Eurofound, 2020; Ajunwa, Crawford and Schultz, 2017). A second barrier concerns the distributive dimension: cultural and linguistic differences, accents and communication styles, neurodiversity or individual conditions may affect signal quality and the interpretability of results, with the risk of stigma and disparities. For this reason, equity and cognitive accessibility of the user interface must be integrated into the design phase and monitored through socio-

## What You Need to Replicate



**Fig. 10** | Package of minimum artefacts to replicate the study and adoption: scenario / briefing / debriefing scripts, HSMS configuration, KPI set, governance rules (credit: the Authors; AI-generated with ChatGPT, 2026).



**Fig. 11** | Design trade-off between innovation benefits (SDG 9) and distributive risks (SDG 10), with mitigation levers including non-prescriptiveness, uncertainty management, audit / purpose limitation, and team focus (credit: the Authors; AI-generated with ChatGPT, 2026).

distributive KPIs (Buolamwini and Gebru, 2018; Raji and Buolamwini, 2019).

From an economic and operational perspective, diffusion requires infrastructure, maintenance, and interdisciplinary skills; partial adoption, particularly without governance and trainer training, may increase risks. There are also regulatory and accountability barriers related to the management of sensitive data, clarity on access and retention, and the definition of responsibilities for use. Transferability is plausible under minimum conditions: a structured final review, governance integrated into the design, and non-prescriptive interfaces that clearly indicate uncertainty and data traceability (European Union, 2016, 2024).

**SDGs: benefits, synergies, and trade-offs** | The On Watch case enables operational interpretation of the profile of benefits, synergies, and trade-offs associated with the introduction of non-prescriptive HSMS in maritime training. In terms of direct benefits, the paper primarily falls within the scope of SDG 9, as it proposes an infrastructural innovation in training based on tools that can verifiably support the preparation and final review phases, without translating into automated decision-making or evaluative devices.

With regard to SDG 10, the value of the framework lies in the reduction of distributive risks through a governance approach integrated from the design stage, which clearly defines the purposes of the system, excludes the attribution of evaluative scores, makes uncertainty and the traceability of inferences visible, regulates access to data, and introduces audits and socio-distributive indicators to monitor potential inequitable effects. This is complemented by a relevant contribution to SDG 4, insofar as the system improves the educational quality of the final review and strengthens learning from error, and to SDG 8, as it protects decent work by limiting evaluative pressures and disciplinary uses of data, and by preserving autonomy, dignity, and fair conditions in training and evaluation processes.

Alongside these direct effects, indirect impacts emerge that concern additional sustainable development goals. In particular, the framework may contribute to SDG 3 by supporting operator well-being and the prevention of overload; to SDGs 12 and 13 through a data-sobriety logic based on minimisation, limited retention periods, and the re-

duction of digital waste; to SDG 16 by strengthening accountability and organisational transparency; and to SDG 17 through the development of partnerships between universities, academies, and sector stakeholders.

However, these benefits do not eliminate trade-offs or residual risks, which may also affect other SDGs, particularly when unequal access to infrastructure persists, distortions in multimodal signals occur, or computational costs are non-negligible. In this sense, potential critical issues that may impact SDGs 1, 2, 5, 7, 11, 14, and 15 should not be considered external or secondary effects, but rather conditions to be governed through design. It is precisely here that the Designer-in-the-Loop framework demonstrates its relevance, as it makes such trade-offs observable, measurable, and governable both ex ante and ex post, through quality thresholds, audits, iterative verification, and continuous design adjustments (Fig. 11). Overall, the results therefore show a clear convergence with SDGs 3, 4, 8, 12, and 16, while also highlighting the need to carefully manage tensions related to surveillance, evaluative pressure, bias, and disparities, so that digital innovation remains aligned with a person-centred training paradigm.

**Conclusions** | This paper proposes shifting the centre of gravity of the debate on AI and training in safety-critical domains: the question is not 'how accurate is human state recognition', but rather 'which socio-technical infrastructure makes innovation verifiable, useful, and fair'. The analysis of the On Watch case shows that the introduction of a Human State Monitoring System in highly critical training contexts generates value above all when it is designed as an architecture of the experience – preparation, scenario, and final review – and when interfaces are conceived as tools for meaning-making rather than as evaluative devices.

The most significant aspect emerging from the study is that the sustainability of innovation does not depend solely on model performance, but on the project's ability to avoid drift towards surveillance and inequalities. For this reason, system governance from the design stage – ensuring the absence of rankings, a clear and limited definition of purposes, controlled data access, visible uncertainty, and traceability of sources – is an intrinsic design requirement. In this context, infrastructural innovation can produce positive impacts only if

accompanied by constraints that protect dignity and equity. Its sustainability requires a shift from an 'I of design' to a 'we of design': the HSMS prototype is, in this sense, an artefact that gains meaning within social practices (preparation and final review) and governance regimes negotiated with trainers and learners. Only a co-designed implementation can reconcile usefulness (SDG 9) and equity (SDG 10).

The transferability of the approach is not tied to a specific manoeuvre or simulator, but to three replicable elements: the design of the training experience as a socio-technical infrastructure centred on preparation and final review; a set of interaction design choices that keep interpretation open and verifiable; and the Designer-in-the-Loop framework as an ex-ante and ex-post method integrating technical-operational and socio-distributive KPIs. Future developments concern: a) robustness and management of variations in the dataset; b) consolidation of metrics and procedures for contextual fairness and low-invasiveness versions; c) formalisation of verification practices, incident reporting, and suspension of use in case of critical issues. Looking ahead, the scientific community is called to evaluate HSMS in training not only through predictive metrics but also through a framework that considers experience, context, and governance as constitutive variables.

## Acknowledgements

The research project 'On Watch – Human State Monitoring in maritime training for safety', developed during the 2025-2026 biennium and still ongoing, has been carried out thanks to funding from FAIMM – Italian Maritime Academy of Genoa, as a partner in training and simulation, granted to the University of Genoa (Department of Architecture and Design, DAD, Italy). The institutions involved are the University of Genoa and FAIMM. The Working Group consists of M. I. Zignego (scientific coordination), A. Bertirotti (socio-organisational analysis), P. Gemelli (training and simulation setting), and L. Pagani (HSMS protocol, KPIs, and analysis).

## References

- Aguinis, H. and Kraiger, K. (2009), "Benefits of Training and Development for Individuals and Teams, Organizations, and Society", in *Annual Review of Psychology*, vol. 60, pp. 451-474. [Online] Available at: doi.org/10.1146/annurev.psych.60.110707.163505 [Accessed 28 March 2026].
- Ajunwa, I., Crawford, K. and Schultz, J. M. (2017), "Limitless Worker Surveillance", in *California Law Review*, vol. 105, pp. 735-776. [Online] Available at: doi.org/10.15779/Z38BR8MF94 [Accessed 28 March 2026].
- Arquilla, V. and Caruso, C. (2025), "Un approccio di integrazione nativa dell'inclusione nel meta-design – L'esperienza COMeta | Towards native integration of inclusivity in

meta-design – The COMeta experience", in *Agathón | International Journal of Architecture, Art and Design*, vol. 17, pp. 360-373. [Online] Available at: doi.org/10.69143/2464-9309/17252025 [Accessed 28 March 2026].

Arquilla, V. and Paracolle, A. (2023), "Design sull'esperienza dell'utente e sostenibilità degli oggetti con intelligenza artificiale | User experience design and sustainability of AI-infused objects", in *Agathón | International Journal of Architecture, Art and Design*, vol. 13, pp. 259-268. [Online] Available at: doi.org/10.19229/2464-9309/13222023 [Accessed 28 March 2026].

Buolamwini, J. and Gebru, T. (2018), "Gender Shades – Intersectional Accuracy Disparities in Commercial Gender Classification", in Friedler, S. A. and Wilson, C. (eds), *Pro-*

ceedings of the 1st Conference on Fairness, Accountability and Transparency, New York, NY, USA, February 23-24, 2018, Proceedings of Machine Learning Research, vol. 81, pp. 77-91. [Online] Available at: [proceedings.mlr.press/v81/buolamwini18a.html](https://proceedings.mlr.press/v81/buolamwini18a.html) [Accessed 28 March 2026].

Cheng, A., Eppich, W., Grant, V., Sherbino, J., Zendejas, B. and Cook, D. A. (2014), "Debriefing for technology-enhanced simulation – A systematic review and meta-analysis", in *Medical Education*, vol. 48, issue 7, pp. 657-666. [Online] Available at: [doi.org/10.1111/medu.12432](https://doi.org/10.1111/medu.12432) [Accessed 28 March 2026].

D'Mello, S. K. and Kory, J. (2015), "A Review and Meta-Analysis of Multimodal Affect Detection Systems", in *ACM Computing Surveys*, vol. 47, issue 3, article 43, pp. 1-36. [Online] Available at: [doi.org/10.1145/2682899](https://doi.org/10.1145/2682899) [Accessed 28 March 2026].

Endsley, M. R. (1995), "Toward a Theory of Situation Awareness in Dynamic Systems", in *Human Factors | The Journal of the Human Factors and Ergonomics Society*, vol. 37, issue 1, pp. 32-64. [Online] Available at: [doi.org/10.1518/001872095779049543](https://doi.org/10.1518/001872095779049543) [Accessed 28 March 2026].

Eurofound (2020), *Employee monitoring and surveillance – The challenges of digitalization*, Publications Office of the European Union, Luxembourg. [Online] Available at: [doi.org/10.2806/424580](https://doi.org/10.2806/424580) [Accessed 28 March 2026].

European Union (2024), *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance)*, document 32024R1689. [Online] Available at: [data.europa.eu/eli/reg/2024/1689/oj](https://data.europa.eu/eli/reg/2024/1689/oj) [Accessed 28 March 2026].

European Union (2016), *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons about the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*, document 32016R0679. [Online] Available at: [data.europa.eu/eli/reg/2016/679/oj](https://data.europa.eu/eli/reg/2016/679/oj) [Accessed 28 March 2026].

Fanning, R. M. and Gaba, D. M. (2007), "The role of debriefing in simulation-based learning", in *Simulation in Healthcare | Journal of the Society for Simulation in Healthcare*, vol. 2, issue 2, pp. 115-125. [Online] Available at: [doi.org/10.1097/SIH.0b013e3180315539](https://doi.org/10.1097/SIH.0b013e3180315539) [Accessed 28 March 2026].

Gazzaniga, M. S. (2005), "Forty-five years of split-brain research and still going strong", in *Nature Reviews Neuroscience*, vol. 6, issue 8, pp. 653-659. [Online] Available at: [doi.org/10.1038/nrn1723](https://doi.org/10.1038/nrn1723) [Accessed 28 March 2026].

Hancock, P. A. and Matthews, G. (2019), "Workload and Performance – Associations, Insensitivities, and Dissociations", in *Human Factors*, vol. 61, issue 3, pp. 374-392. [Online] Available at: [doi.org/10.1177/0018720818809590](https://doi.org/10.1177/0018720818809590) [Accessed 28 March 2026].

Hart, S. G. and Staveland, L. E. (1988), "Development of NASA-TLX (Task Load Index) – Results of Empirical and Theoretical Research", in Hancock, P. A. and Meshkati, N. (eds), *Human Mental Workload*, North-Holland, Oxford, pp. 139-183. [Online] Available at: [doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9) [Accessed 28 March 2026].

ISO – International Organization for Standardization (2019), *ISO 9241-210:2019 – Ergonomics of human-system interaction – Part 210 – Human-centred design for interactive systems*. [Online] Available at: [org/standard/77520.html](https://org/standard/77520.html) [Accessed 28 March 2026].

Koelstra, S., Mühl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A. and Patras, I. (2012), "DEAP – A Database for Emotion Analysis – Using Physiological Signals", in *IEEE Transactions on Affective Computing*, vol. 3, issue 1, pp. 18-31. [Online] Available at: [doi.org/10.1109/T-AFFC.2011.15](https://doi.org/10.1109/T-AFFC.2011.15) [Accessed 28 March 2026].

Liu, Y., Lan, Z., Cui, J., Krishnan, G., Sourina, O., Konovessis, D., Ang, H. E. and Mueller-Wittig, W. (2020),

"Psychophysiological evaluation of seafarers to improve training in maritime virtual simulator", in *Advanced Engineering Informatics*, vol. 44, article 101048, pp. 1-9. [Online] Available at: [doi.org/10.1016/j.aei.2020.101048](https://doi.org/10.1016/j.aei.2020.101048) [Accessed 28 March 2026].

MIM – Ministero dell'Istruzione e del Merito (2025), *JOB&Orienta 2025 – Programma del Ministero dell'Istruzione e del Merito*. [Online] Available at: [mim.gov.it/documents/20182/8952594/Job%26Orienta%2B2025%2B-%2BProgramma%2Bdel%2BMinistero%2Bdel%27Istruzione%2Be%2Bdel%2BMerito.pdf/def37280-0067-4a6b-b1cc-09bdc31897ba?i=1764260799177&version=1.1](https://mim.gov.it/documents/20182/8952594/Job%26Orienta%2B2025%2B-%2BProgramma%2Bdel%2BMinistero%2Bdel%27Istruzione%2Be%2Bdel%2BMerito.pdf/def37280-0067-4a6b-b1cc-09bdc31897ba?i=1764260799177&version=1.1) [Accessed 28 March 2026].

Mittelstadt, B. (2019), "Principles alone cannot guarantee ethical AI", in *Nature Machine Intelligence*, vol. 1, issue 11, pp. 501-507. [Online] Available at: [doi.org/10.1038/s42256-019-0114-4](https://doi.org/10.1038/s42256-019-0114-4) [Accessed 28 March 2026].

NIST – National Institute of Standards and Technology (2023), *Artificial Intelligence Risk Management Framework (AI RMF 1.0) – NIST AI 100-1*. [Online] Available at: [doi.org/10.6028/NIST.AI.100-1](https://doi.org/10.6028/NIST.AI.100-1) [Accessed 28 March 2026].

OECD (2019), *Recommendation of the Council on Artificial Intelligence*. [Online] Available at: [one.oecd.org/document/C/MIN\(2019\)3/FINAL/en](https://one.oecd.org/document/C/MIN(2019)3/FINAL/en) [Accessed 28 March 2026].

Osello, A., Ugliotti, F. M., Rimella, N. and Loddo, F. (2024), "Modelli digitali e linguaggio naturale – Nuove prospettive per interpretare la complessità | Digital models and natural language – New perspectives for interpreting complexity", in *Agathón | International Journal of Architecture, Art and Design*, vol. 16, pp. 212-219. [Online] Available at: [doi.org/10.19229/2464-9309/16182024](https://doi.org/10.19229/2464-9309/16182024) [Accessed 28 March 2026].

Parasuraman, R. and Riley, V. (1997), "Humans and Automation – Use, Misuse, Disuse, Abuse", in *Human Factors | The Journal of the Human Factors and Ergonomics Society*, vol. 39, issue 2, pp. 230-253. [Online] Available at: [doi.org/10.1518/001872097778543886](https://doi.org/10.1518/001872097778543886) [Accessed 28 March 2026].

Picard, R. W. (1997), *Affective Computing*, The MIT Press, Cambridge, MA. [Online] Available at: [doi.org/10.7551/mitpress/1140.001.0001](https://doi.org/10.7551/mitpress/1140.001.0001) [Accessed 28 March 2026].

Porfirione, C., Ferrari Tumay, X. and Leggiero, I. (2024), "Conoscenza, innovazione e cambiamento – Il potere dell'errore nel design e nei sistemi complessi | Knowledge, innovation, and change – The power of error in design and complex systems", in *Agathón | International Journal of Architecture, Art and Design*, vol. 16, pp. 232-241. [Online] Available at: [doi.org/10.19229/2464-9309/16202024](https://doi.org/10.19229/2464-9309/16202024) [Accessed 28 March 2026].

Poria, S., Cambria, E., Bajpai, R. and Hussain, A. (2017), "A review of affective computing – From unimodal analysis to multimodal fusion", in *Information Fusion*, vol. 37, pp. 98-125. [Online] Available at: [doi.org/10.1016/j.inffus.2017.02.003](https://doi.org/10.1016/j.inffus.2017.02.003) [Accessed 28 March 2026].

Raji, I. D. and Buolamwini, J. (2019), "Actionable Auditing – Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products", in *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19)*, Honolulu, HI, USA, January 27-29, 2019, Association for Computing Machinery, New York, NY, pp. 429-435. [Online] Available at: [doi.org/10.1145/3306618.3314244](https://doi.org/10.1145/3306618.3314244) [Accessed 28 March 2026].

Ronca, V., Uflaz, E., Turan, O., Bantan, H., MacKinnon, S. N., Lommi, A., Pozzi, S., Kurt, R. E., Arslan, O., Kurt, Y. B., Erdem, P., Akyuz, E., Vozzi, A., Di Flumeri, G., Aricò, P., Giorgi, A., Capotorto, R., Babiloni, F. and Borghini, G. (2023), "Neurophysiological Assessment of An Innovative Maritime Safety System in Terms of Ship Operators' Mental Workload, Stress, and Attention in the Full Mission Bridge Simulator", in *Brain Sciences*, vol. 13, issue 9, article 1319, pp. 1-19. [Online] Available at: [doi.org/10.3390/brainsci13091319](https://doi.org/10.3390/brainsci13091319) [Accessed 28 March 2026].

Rosmino, A. (2024), "Corpi, menti e design – Un approccio integrato per l'innovazione museale | Bodies, minds, and design – An integrated approach to museum innovation", in *Agathón | International Journal of Architecture,*

*Art and Design*, vol. 16, pp. 278-289. [Online] Available at: [doi.org/10.19229/2464-9309/16232024](https://doi.org/10.19229/2464-9309/16232024) [Accessed 28 March 2026].

Rudolph, J. W., Simon, R., Dufresne, R. L. and Raemer, D. B. (2006), "There's no such thing as 'nonjudgmental' debriefing – A theory and method for debriefing with good judgment", in *Simulation in Healthcare | Journal of the Society for Simulation in Healthcare*, vol. 1, issue 1, pp. 49-55. [Online] Available at: [doi.org/10.1097/01266021-200600110-00006](https://doi.org/10.1097/01266021-200600110-00006) [Accessed 28 March 2026].

Salas, E., Tannenbaum, S. I., Kraiger, K. and Smith-Jentsch, K. A. (2012), "The Science of Training and Development in Organizations – What Matters in Practice", in *Psychological Science in the Public Interest*, vol. 13, issue 2, pp. 74-101. [Online] Available at: [doi.org/10.1177/1529100612436661](https://doi.org/10.1177/1529100612436661) [Accessed 28 March 2026].

Vacanti, A., De Chirico, M., Crippa, D. and Fagnoni, R. (2025), "Temporalità digitale – La responsabilità dell'Interaction Design nell'alienazione digitale | Digital temporality – The responsibility of Interaction Design in digital alienation", in *Agathón | International Journal of Architecture, Art and Design*, vol. 17, pp. 336-347. [Online] Available at: [doi.org/10.69143/2464-9309/17232025](https://doi.org/10.69143/2464-9309/17232025) [Accessed 28 March 2026].

Villani, T., Romagna, G. and Oddi, A. (2024), "Ottimizzare la fruibilità nei musei – Gestione integrata di dati sui modi d'uso dello spazio e dei contenuti culturali | Optimising usability in museums – Integrated management of data on the use of space and cultural content", in *Agathón | International Journal of Architecture, Art and Design*, vol. 16, pp. 220-231. [Online] Available at: [doi.org/10.19229/2464-9309/16192024](https://doi.org/10.19229/2464-9309/16192024) [Accessed 28 March 2026].

Wulvik, A. S., Dybvik, H. and Steinert, M. (2020), "Investigating the relationship between mental state (workload and affect) and physiology in a control room setting (ship bridge simulator)", in *Cognition, Technology & Work*, vol. 22, issue 1, pp. 95-108. [Online] Available at: [doi.org/10.1007/s10111-019-00553-8](https://doi.org/10.1007/s10111-019-00553-8) [Accessed 28 March 2026].

Xue, H., Haugseggen, Ø., Røds, J.-F., Batalden, B.-M. and Prasad, D. K. (2024), "Assessment of stress levels based on biosignal during the simulator-based maritime navigation training and its impact on sailing route reliability", in *Transportation Research Interdisciplinary Perspectives*, vol. 24, article 101047, pp. 1-14. [Online] Available at: [doi.org/10.1016/j.trip.2024.101047](https://doi.org/10.1016/j.trip.2024.101047) [Accessed 28 March 2026].

Zehr Gantz, S., Massari, S., McDonagh, D. and Vokoun, J. (2025), "Affrontare l'insicurezza alimentare – Empatia e Design Thinking per l'apprendimento trasformativo | Addressing food insecurity – Leveraging empathy and Design Thinking to achieve transformative learning", in *Agathón | International Journal of Architecture, Art and Design*, vol. 17, pp. 384-397. [Online] Available at: [doi.org/10.69143/2464-9309/17272025](https://doi.org/10.69143/2464-9309/17272025) [Accessed 28 March 2026].

Zeng, Z., Pantic, M., Roisman, G. I. and Huang, T. S. (2009), "A Survey of Affect Recognition Methods – Audio, Visual, and Spontaneous Expressions", in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, issue 1, pp. 39-58. [Online] Available at: [doi.org/10.1109/TPAMI.2008.52](https://doi.org/10.1109/TPAMI.2008.52) [Accessed 28 March 2026].

Zuboff, S. (2019), *The Age of Surveillance Capitalism – The Fight for a Human Future at the New Frontier of Power*, PublicAffairs, New York. [Online] Available at: [hbs.edu/faculty/Pages/item.aspx?num=56791](https://hbs.edu/faculty/Pages/item.aspx?num=56791) [Accessed 28 March 2026].

