Människans roll i autonoma marina system - CASE från BAOUT projektet

Presentation

- Pernilla Ulfvengren – KTH - Industrial engineering and management, INDEK (Technology & Management)

  – Maskiningenjör
  – Tekn Dr Human-Machine systems
  – Docent Industrial engineering and sociotechnical systems

  – Tidigare främst erfarenhet från flyget och säkerhet
  – KTH partner in HILAS, MASCA, PROSPERO and ORION

(Centre for Innovative Human Systems, TCD- Irland – kordinator)

Säkra förbättringar

- Identifiera relevanta risker och behov av förbättring
- Säkerställ att förändring faktiskt leder till implementering
- och att det hade önskad effekt dvs inte introducerade nya risker.

- I syfte att öka säkerheten i sociotekniska system

Senaste projekt - TRV

- FLYT 365 – vattenburen kollektivtrafik Stockholmsregionen
- INFRA - Systemanalys av nätverk av problem associerade till flygbuller

- APIS – Implementering av drönare i Stockholmsregionen – akustisk simuleringsplattform som beslutsstöd inför val och bedömning av rutter och landningsplatser map buller/störning.

Agenda

- Några ord om safety assurance – försäkran om säkerhet
- BOAUT projektet
- Resultat
- Idéer om behov av framtida forskningsområden

Assuring autonomy programme (Univ. of York)

- Body of Knowledge (BoK) som på sikt ska bli källan till referenser om ”assurance and regulation of Robotics and Autonomous Systems (RAS)
- Syftar till att undanröja hinder för innovation och implementering
- Kritiska hinder definierar vad som måste lösas för att undvika följande risker:
  - Ett säkert system kommer inte ”rullas ut”
  - Ett icke-säkert system ”rullas ut”
  - Tillämpning av säker teknik går långsamt
  - Utveckling avstannar inom något område
  - Incidenter och olyckor kommer slå tillbaka på utvecklingen

Kritiska hinder

- **Adaption** – hur ska RAS som anpassar beteende genom lärande garanteras fortsatt vara säkert nog? Hur ser regelverket ut?
- **Bounding behaviour** – där RAS opererar säkert, inom givna gränser, hur ska dessa gränser definieras i design och drift?
- **Explanations** – vilka beslut som fattas av RAS behöver förklaras för användare eller tillsyn? Hur ska detta göras på ett lämpligt sätt givet att flera beslut per sekund kan ske?
- **Handover** – i (semi) autonoma system där kontrollen överförs till en mänsklig operatör – hur försäkras att operatören har erforderlig situationsmedvetenhet för uppdraget?
- **Incident and accident investigations** – hur ska regelverket se ut och vilken information behövs göras tillgänglig?
- **Validation and verification** – vad är effektiva sätt att säkerställa dessa steg inom simulering, riskbedömning osv.?

BOAUT

BOundaries of AUTonomy

2022-03-13

VINDÖ MARIN carmenta KTH Ondheim systems

Project sponsor

ASSURING AUTONOMY INTERNATIONAL PROGRAMME

Project - Maritime Autonomous Surface Ship

Background

- Maritime Autonomous Surface Ships (MASS) are being trialled across the globe, with the International Maritime Organisation (IMO) bringing them onto the agenda actively since 2017.
- The value of the market is substantial, with large key players such as ABB, Honeywell International, Kongsberg, etc.
- Still, substantial confusion about what MASS means and how they will be handled.
- Focus is on overcoming the challenges of human-machine interactions in e.g. remote control, and “easing into” autonomy.

Project

- Boundaries Of AUTonomy (BOAUT)
  - Explored the role of ROC in assuring the safety of MASS.
- The challenge
  - Is it possible to assure that, **with the support of ROC**, autonomous ships (MASS) can operate as safely as manned ships?

Background

- Remote operations centres (ROC) are envisioned to be an important part of a maritime ecosystem involving MASS.
- Also in regard to ensuring their safety.

Reasons for keeping humans out-of-the-loop but still in the system?

Humans biggest risk?

Humans saves the day (everyday..)?

Added value of humans in autonomous systems?
Increasing the level of automation can be counterproductive?

Autonomy LOA: Väljer metod, utför åtgärd och ignorerar människan

Human operator's new role ”safety assurance”?

Project - Levels of Autonomy

	Human Control from Ship Bridge	Remote Human Control	Fully Autonomous
Human Role	Seafarers are on board to operate and control shipboard systems and functions. Some operations may be automated and at times be unsupervised but with seafarers on board to take control.	The ship is controlled and operated from another location.	Monitoring and emergency management.

Project - Levels of Autonomy

	Human		Fully Autonomous
Human Role	Seafar shipbo may be with se	ed	Monitoring and emergency management.

What does support by ROC even mean then, considering MASS should operate autonomously? (BOAUT did not focus on the road to autonomy.)

BOAUT- F. Asplund

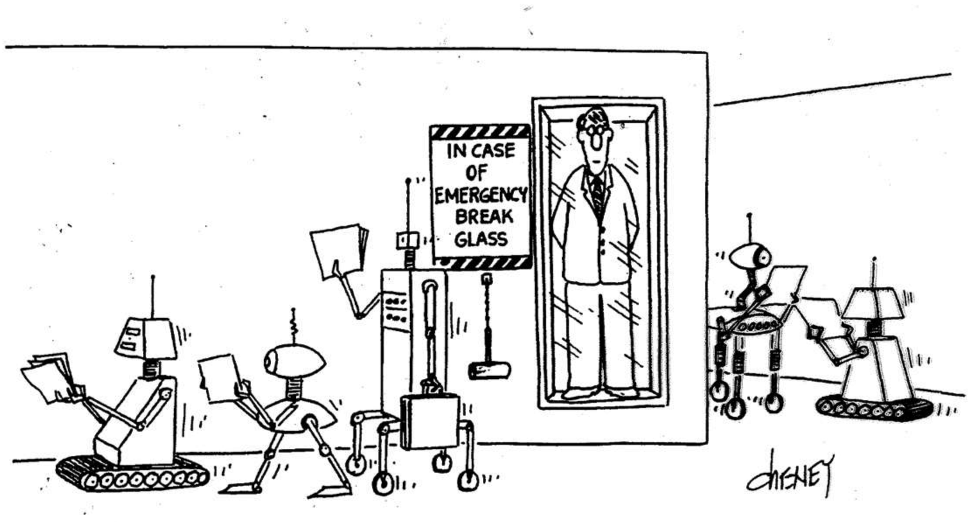


FIGURE 3.5
Ultimate functional allocation when using a "capability" criterion. (*Source:* Cheney, 1989. New Yorker Magazine, Inc.)

Human role and function now and beyond?

Humans are the biggest risk?

Humans saves the day (everyday..)?

Added value of humans' new role in autonomous systems?

Human operator's abilities for "safety assurance"?

- Improvise and solve problems
- Contain information for a long period of time and acquire relevant information for the situation.
- Recognize problems and situations, perform assessments
  - Fitts's list 1951 "MABA-MABA"

4/19/2023 17

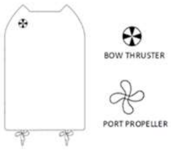
Project

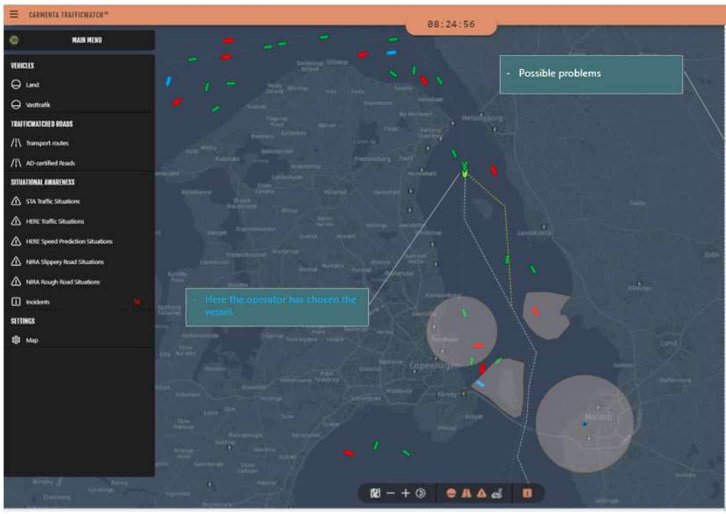
- Method
  - Defined detailed scenarios of interest
  - Configured a control application demonstrator
  - Tested scenarios with relevant operators
- Goal
  - Define a safety case based on outcomes

2023-04-19 18

BOAUT- F. Asplund

Project - Scenarios





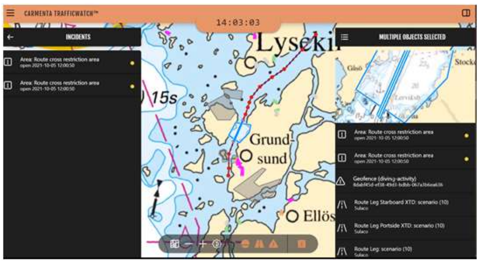
Vessel Capability

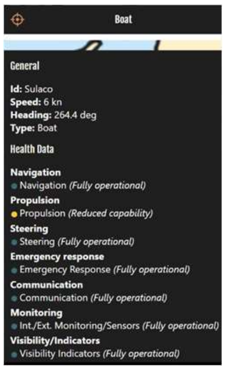
Basic Functionality
- Propulsion
- Steering
- Communication
- Navigation
- Internal Monitoring
- Visibility/Indicators
- Emergency response

Mission Functionality
- Automated Mooring and Disembarking
- Passenger comfort

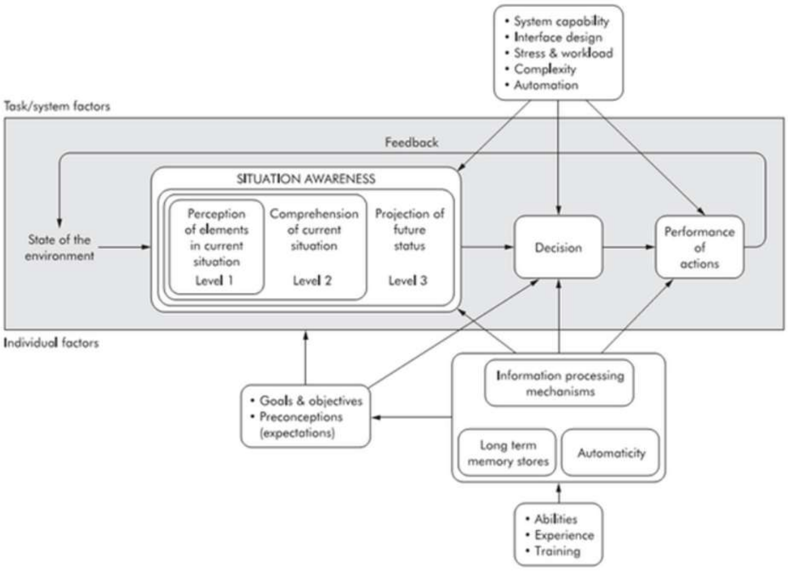
Sensor Data : Suggested AIS, Radar, and Camera.
~~(MASS has identified fog in area, …)~~
Default messages

Project - Demonstrator





SA



Perception of elements in current situation

– Is the demonstrator conveying sufficient information in relation to scenario objective?
  > *Performance of demonstrator interface - presenting information clearly*
  > *Operability of demonstrator interface – operator may acquire information*
  > *Acceptability – demonstrator fulfills requirement for operator to perceive elements of current situation*

– The demonstrator visualizing the situation intended in the scenario in a clear way.
– Operator perceives route, traffic, route change and alerts for risk, health issues etc.
– Operator does not get confused with irrelevant information

Criteria: evidence of the above from validation activity
– Activity: talk aloud and control questions ”what do you see”, any information missing ”what do you not see”

Comprehension of current situation

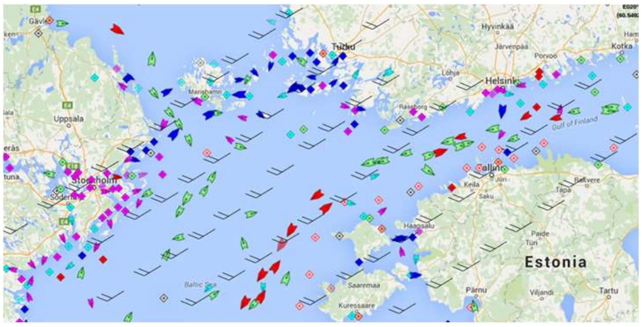
- Is the operator making sense of the situation in the scenario x with the information provided and acquired?
  - > *Performance of demonstrator – present MASS performance in relation to other vessels, restricted areas, knows type of vessels (as well as own)*
  - > *Operability of demonstrator interface – operator may acquire information to verify reasons behind MASS decisions*
  - > *Acceptability – demonstrator fulfills requirement for operator to comprehend current situation*
- Procedures
  - > *traffic maximum cross time exceedence*
  - > *reduced maneuverability, loss of speed, loss of heading (off-heading alarm)*
  - > *New information and risk contours conflict detection*

Projection of future status

- Operator can assess the situation and accept or not accept the BOAUT concept decision and reasons behind it.
  - > *Performance of demonstrator – allows operator to assess decisions made by MASS by verifying local SA information/data, global SA information/data and use competence and experience to assess risk and compare with situation at hand and decision made.*
  - > *Operability of demonstrator interface – operator finds information confirming their idea of what is about to happen.*
  - > *Acceptability – demonstrator fulfills requirement for operator to project future status*

Findings - Generic (1)

- A key concern for ROC operators are of course their situational awareness.
- Thus, much feedback on GUI design, configurability, prediction support, incident handling, notifications, …
- Good response to GUI, but, while important, not our focus.



Findings - Generic (2)

- Mixed traffic
- High respect for ship captain (i.e. also fleet management operator), but not all vessels operated by a competent, trained individual.
- Frequently broken “rules”

- Other domains might have to learn from autonomous shipping, rather than the other way around.

Findings - Generic (3)

- Fleet Management
- Traffic Management

- Not that easy to delegate the handling of unknown unknowns.
- Traffic management might have to handle emergencies.

Findings - Autonomy is a Change of Perspective

- MASS will have to be highly capable
- Frequently forgotten by operators, as they became bounded by their experiences.

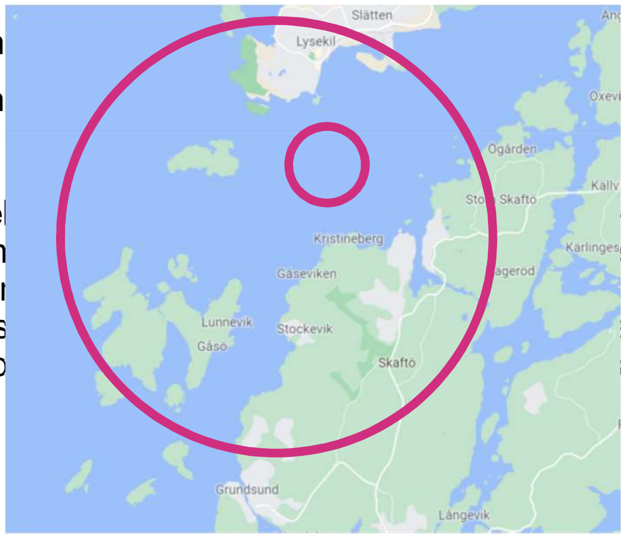
- The explanation provided by the MASS for its behaviour, rather than its actions, **is critical**.

Findings - Autonomy is a Change of Perspective

ame

Findings - To Confront or Avoid Danger?

- On
- On
- Rel - both ren d going fro for ROC sys and pro .



Findings - To Confront or Avoid Danger?

- One narrow focus on go/no-go situations.
- One broader focus on trade-offs.

- Related to envisioned role, but regardless – both fleet management control and traffic management need going from broader to narrow.
- Suggests the need for ROC systems capable of directing the attention and providing information as a situation evolves.

Findings -
An Expert with Support, or the Support to the Experts?

- Go/no-go situations
- Trade-off reasoning

- Confusion on the need for expertise, and the wish to avoid liability, can be handled by strict rules and step-by-step action lists.
- A case could be made for this approach at lower levels of automation, but seems to be the opposite of why the ROC operator would be there.
- A possible lock-in?

Findings - An Expert with Support, or the Support to the Experts?

h
d
or
tor

Being the human in the room?

- What about the kid?
- What kid?

Preliminary results

- Overall demonstrator was percieved as realistic for the purpose of discussing remote operator's role in autonomous system.
- Information elements:
  - Operator's found details irrelevant to their task as distracting – the level of detail of the charts: Example: no need to have depth marked, rather safety depths only
- Current situation:
  - I see other boats – but no-one is relevant to me (good).
  - Alarms and health system help me to understand the situation and why maneuverability has changed
  - Perhaps change color of boat when changes occur
- Projection of future status:
  - Would like to know final destination to assess route change…
  - I need information that explains reasons behind route change…
  - Difficult to understand the intention of a group of canoes (mixed operation environment)

Findings - Summary

- AI transparency and explainability
- How to appropriately direct the attention of operators through graphical interfaces.
- Maritime infrastructure for vessel traffic services

- Blame when introducing ROC as a support to MASS, as operators will fail.
- Supporting AI reasoning with human insights.
- How to include information from untrusted sources (such as the public) in common operational pictures.

Human centered autonomy?

- Are the design challenges of fully autonomous vessels different than classic ”ironies of automation”?
- Ergonomics/Human Factors traditionally identifies automation issues as residual risk or lack of human centered automation.
- Is there a difference in design for the humans in a safety assurance role for MASS?

- How to design support for ROC operator to identify and manage MASS breakdown?
  - Human role and function definition.
  - Human Cyber-physical system **design**
  - Maritime context

2023-04-19 37

GAP – Maritime and ”normal accidents”

- In a phased development with the design philosophy that humans have a role in safety assurance in going for fully autonomous systems, **then normal operation for these operators is near, or actual autonomous system breakdown.**

- How to design support for managing MASS breakdown?

- Maritime operational context – fast small private vessels, waves, currents, wind, obstacles (depth, land, other..)
- Maritime transportation system, reliability, ”sea worthiness”, regulation, maintenance etc.

2023-04-19 38

Operatören i kärnkraftverk eller flygplan när det händer… kommer vara ROCs enda uppgift?

- Operatören, som ska övervaka systemet och ingripa om det händer något oförutsett, kan uppleva sitt jobb som
- ***”99% monotoni, då det inte händer någonting, och 1% plötslig skräck,***
- då det inträffar en störning”
- som händer så sällan att man inte har fått utbildning på det eller har någon erfarenhet av.

Från presentation av Prof. Lena Mårtensson

Människors främlingskap inför datorer

(Sheridan, 1980)

- Människor bekymrar sig för att:
- Avlägsna sig från de operationer de ska utföra, interagera med dator i stället för systemet;
- Förlora sin perceptuellt-motoriska skicklighet, fingerfärdigheten;
- ***Ej förstå vad som händer i datorn***
- Människan ska övervaka datorerna
- ”*medan ett mänskligt fel tidigare kunde passera obemärkt och lätt kunde korrigeras, kan nu ett sådant fel framkalla en katastrof*”.

Från presentation av Prof. Lena Mårtensson

Clumsy automation...

(Wiener, 1989)

- Piloters kommentarer efter att ha flugit B-757 ett år:
  - ”What it is doing now?”
  - ”Why did it do that?”
  - ”What will it do next?”
  - ”How am I going to stop it from doing that?”
  - ”How am I going to make it do what **I** want?”

Från presentation av Prof. Lena Mårtensson

Sarter, Woods och Wickens (1994)

- ”automationen kan ibland uppfattas, som att den inte fungerar, när komplexiteten i dess algoritmer leder den att göra saker som en pilot inte förväntar sig, därför att piloten inte förstår de förutsättningar som kan ha triggat flygplanets beteende. Dessa former av
  - ***”automation-induced-surprises”***
- Kan, när de inträder leda till att piloten ingriper med en mindre optimal och osäker handling”.

Från presentation av Prof. Lena Mårtensson

Human role and function - Intervene in case of MASS breakdown

A ROC operator can contribute to safety assurance by acting in hazardous situations when the MASS:

- cannot signal that it is outside of its mission limits.
- can signal that it is outside of its mission limits.
- has a limited local situational awareness due to limited sensor information (compared to ROC global situational awareness)

- Classic automation challenges of ”ironies of automation” applies?

Ironies of automation (Bainbridge, 1983):

- The first paradox – control handed over to operators only when automation fails
  - with less skill than when performed regularly, see second paradox.
- The second paradox – automation allows fewer opportunities to train tasks
  - the more reliable the system is.
- The third paradox, leaving the operator only to supervise
  - vigilance becomes a challenge.
- The fourth paradox, humans are good at tackling the unexpected but requires time and attention
  - unfit for time critical emergencies.

Design for autonomy and safety assurance?

- Function and role is to identify and manage MASS breakdown:
  - Identifying hazards
  - Identifying hazardous behavior
  - Identifying potential deviation from required behaviour
  - Mitigating potential deviations
- Design for Situation Awareness (SA)
- Design for ”clumsy automation”

Maritime context - to intervene or not

- If the operator, in time to act, understands what the MASS is doing and why and assess the behaviour safe, no action is taken.
- If the operator, in time to act, understands what the MASS is doing and why but assess this to be an unsafe behaviour, then the ROC operator will intervene.
- If the operator understands too late or fails to identify (lack of SA) that the MASS is deviating from required behaviour then the ROC operator will not intervene.
- Damage control is still possible, even if operator failed to intervene, by warning others of what is about to happen and activate emergency response resources.

Maritime context
- Cassandra problem, error and blame

- The operator is assumed to intervene after identifying and mitigating what is understood as a deviation from required behaviour to be safe.
- If the ROC operator intervene in time, then safety is assured, or not, depending on if intervention is successful.
- After an investigation it is concluded that the MASS had it under control, or not, and the human contribution to safety assurance was either:
  - > *counterproductive or*
  - > *Saved the day!*
- The operator does not intervene after assessing the situation safe.
  - – Safety is assured. The operator did the right call by not intervening.
  - – Safety is not assured. After an investigation it is concluded that the MASS did not have it under control and a human safety assurance effort had been required.

2023-04-19 47

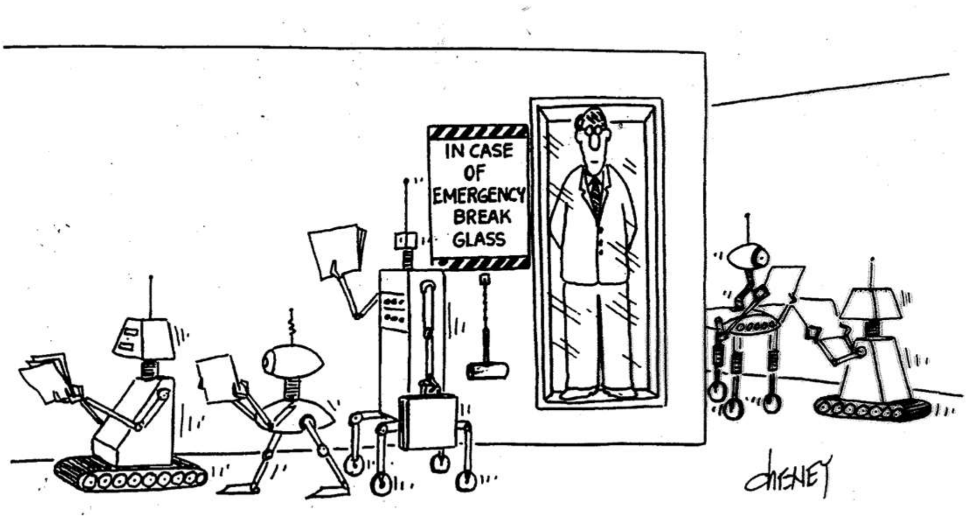


FIGURE 3.5
Ultimate functional allocation when using a "capability" criterion. (*Source:* Cheney, 1989. New Yorker Magazine, Inc.)

Utgått från autonoma fartyg i systemet

- Har denna bild ngn bäring på framtiden?
- Are there **Reasons for keeping humans out-of-the-loop but still in the system?**
- **Är det ens försvarbart?** Givet vad vi vet om människors förutsättningar i högautomatiserade system?
- Kan vi argumentera att människan skulle kunna vara en säkerhetsgarant i dessa system?
- Is **Human centered autonomy s**omething new or just old wine in new bottles?
- If **Normal operations is MASS breakdown!**
- What is then a …n**ew functional and task analysis and HF guidelines for engineers of autonomy?**

Best practice methods for the future?

- Work function and task analysis
- Participatory approach
- Safety case, scenarios
- Cathcing a moving target
- Systems engineering
- Concept generation
- Concept validation
- E-OCVM for MASS (fully autonomous and maritime context)

- F. Asplund och P. Ulfvengren, "Work functions shaping the ability to innovate: insights from the case of the safety engineer," *Cognition, Technology & Work*, 2019
- F. Asplund och P. Ulfvengren, "Engineer-Centred Design Factors and Methodological Approach for Maritime Autonomy Emergency Response Systems," *Safety*, vol. 8, no. 3, s. 54-54, 2022.