ADVANCES IN MARITIME AI:

APPLICATIONS OF GENERATIVE AND EXPLAINABLE AI IN SYSTEMS AND TRAINING

EDITOR:
YUSUF TARIK MUTLU

BIDGE Publications

Advances in Maritime AI: Applications of Generative and Explainable AI in Systems and Training

Editor: Yusuf Tarık Mutlu

ISBN: 978-625-8821-20-8

1st Edition

Page Layout By: Gozde YUCEL

Publication Date: 24.06.2026

BIDGE Publications

Certificate No: 71374

www.bidgeyayinlari.com.tr - bidgeyayinlari@gmail.com
Krc Bilişim Ticaret ve Organizasyon Ltd. Şti.
Güzeltepe Mahallesi Abidin Daver Sokak Sefer Apartmanı No: 7/9 Çankaya / Ankara

İÇİNDEKİLER

OTONOM GEMİ MAKİNE DAİRELERİNDE AÇIKLANABİLİR YAPAY ZEKA (XAI): TRIZ VE TOPSIS TABANLI BİR SAVUNMA SANAYİİ ENTEGRASYON MODELİ

Yusuf Tarık MUTLU[1 2]
Muhammed Emin BAŞAK[3]

Giriş

Denizcilik sektörü ve savunma sanayii, gemi otomasyonu ve otonom sistemler bağlamında tarihinin en köklü dijital dönüşümlerinden birini yaşamaktadır. Geleneksel planlı bakım ve arıza sonrası müdahale stratejileri, yerini sensör verilerine dayalı proaktif karar destek sistemlerine ve kestirimci bakım (predictive maintenance - PdM) metodolojilerine bırakmaktadır (Simion et al., 2024). Özellikle TCG Anadolu gibi çok maksatlı amfibi hücum gemileri veya MİLGEM sınıfı korvetler gibi stratejik deniz muharip

[1] Araştırma Görevlisi, İstanbul Teknik Üniversitesi, Denizcilik Fakültesi, Gemi Makineleri İşletme Mühendisliği Bölümü, Orcid: 0009-0003-0400-8659
[2] Doktora Öğrencisi, Yıldız Teknik Üniversitesi, Gemi İnşaatı ve Denizcilik Fakültesi, Gemi İnşaatı ve Gemi Makineleri Mühendisliği Bölümü, Orcid: 0009-0003-0400-8659
[3] Prof. Dr., Yıldız Teknik Üniversitesi, Gemi İnşaatı ve Denizcilik Fakültesi, Gemi Makineleri İşletme Mühendisliği Bölümü, Orcid: 0000-0001-5520-8579

unsurlarda, tahrik ve enerji üretim sistemlerinin kesintisiz operasyonel hazırlığı hayati önem taşımaktadır. Bu tür karmaşık askeri unsurlarda, makine dairesinde yaşanabilecek beklenmedik bir arıza yalnızca lojistik bir gecikme değil, doğrudan ulusal güvenliği ve Mavi Vatan doktrinindeki güç projeksiyonunu olumsuz etkileyebilecek stratejik bir zafiyettir.

Son yıllarda, makine dairesi bileşenlerinin durum izlemesi (condition monitoring) ve erken arıza teşhisi için derin öğrenme (deep learning - DL) ve yapay sinir ağları gibi ileri düzey makine öğrenmesi teknikleri literatürde geniş yer bulmaya başlamıştır (Simion et al., 2024; Wang et al., 2025). Bu algoritmalar, sensörlerden gelen titreşim, sıcaklık ve basınç gibi çok boyutlu zaman serisi verilerini analiz ederek arızaları yüksek bir doğruluk payıyla önceden tespit edebilmektedir. Ancak, bu modellerin mimari derinliği arttıkça, karar verme süreçleri giderek daha karmaşık ve şeffaflıktan uzak bir "kara kutu" (black box) yapısına bürünmektedir (Zemmouchi-Ghomari, 2026). Otonom veya yarı-otonom bir sistemin, vardiya mühendisine "jeneratörde %90 ihtimalle arıza bekleniyor" uyarısını vermesi tek başına yeterli değildir; sistemin bu karara hangi spesifik anomaliye dayanarak vardığını açıklayamaması, kriz anlarında operatörün algoritmaya olan güvenini zedelemekte ve acil müdahale süresini uzatmaktadır.

Tam bu noktada, açıklanabilir yapay zeka (explainable AI - XAI) konsepti, denizcilik otomasyonundaki bu şeffaflık problemini çözmek için kritik bir köprü görevi görmektedir. XAI, karmaşık yapay zeka modellerinin ürettiği sonuçları, sahadaki personelin ve karar vericilerin anlayabileceği, güvenebileceği ve eyleme dökebileceği şeffaf bir formata dönüştürmeyi amaçlayan teknikler bütünüdür (Kalafatelis et al., 2025; Park et al., 2023).

Bu çalışma, denizcilik endüstrisindeki kestirimci bakım teknolojilerinde yaşanan "yüksek doğruluk fakat düşük şeffaflık" problemini ele almaktadır. Çalışma kapsamında, otonom makine dairesi konseptine yönelik güncel küresel patentler incelenerek, mevcut yapay zeka çözümlerindeki donanımsal ve algoritmik çelişkiler tespit edilmiştir. Literatürdeki bu boşluk, yaratıcı problem çözme teorisi (TRIZ) matrisi kullanılarak analiz edilmiş ve makine dairesi otomasyonuna özel XAI stratejileri (bölgesel modeller, tersine çevirme, 3d görselleştirme ve dinamik özellik seçimi) geliştirilmiştir. Geliştirilen bu mühendislik konseptleri, TOPSIS (technique for order preference by similarity to ıdeal solution) çok kriterli karar verme yöntemi ile TCG Anadolu vizyonu ve yerli savunma sanayii dinamikleri doğrultusunda önceliklendirilmiştir. Bölümün nihai amacı, teorik yapay zeka tartışmalarını makine dairesindeki operasyonel gerçeklerle buluşturan stratejik ve yerli bir XAI mimarisi sunmaktır.

Literatür Özeti ve Problemin Tanımı

Denizcilikte kara kutu zafiyeti ve XAI ihtiyacı

Gemi makine dairelerindeki kestirimci bakım stratejileri, son yıllarda klasik istatistiksel yöntemlerden derin öğrenme ve yapay sinir ağlarına doğru belirgin bir geçiş yapmıştır. Denizcilik uygulamalarında yapay zeka tabanlı hata tespiti ve teşhis algoritmalarının, geleneksel eşik değerli uyarı sistemlerine kıyasla çok daha hassas sonuçlar verdiğini ortaya koymuştur (Cheliotis et al., 2022). Ancak literatür, bu modellerin pratik operasyonel sahadaki uygulanabilirliğinin ciddi sınırlamalar içerdiğini de göstermektedir.

Derin öğrenme modellerinin karşılaştığı en temel kısıtlamalardan biri "kara kutu" doğalarıdır. Yüzlerce gizli katmandan oluşan bu mimariler, gemi hızındaki değişimler, yük durumları ve zorlu deniz koşulları gibi değişken çevre faktörleri altında isabetli arıza tahminleri yapsa da, modelin bu karara nasıl vardığı şeffaf değildir. Güncel literatürden kaynaklar, gemi motorlarının kestirimci bakımında sadece tahmin yapmanın yeterli olmadığını; modelin hangi sensör verisine dayanarak (örneğin egzoz gazı sıcaklığı mı yoksa soğutma suyu basıncı mı?) arıza uyarısı verdiğinin açıklanabilmesinin operasyonel güvenlik için zorunlu olduğunu vurgulamaktadır (Je-Gal et al., 2024; Kalafatelis et al., 2025).

Otonom ve yarı-otonom gemi sistemlerinde insan-makine işbirliğinin (human-machine collaboration) doğası, doğrudan "güven" unsuruna dayanır. Yapılan çalışmalar, otonom gemilerde karar şeffaflığının eksikliğinin, operatörlerin sistemin uyarılarını ya tamamen göz ardı etmesine ya da körü körüne güvenmesine (automation bias) yol açtığını; her iki durumun da kriz anlarında felaketle sonuçlanabileceğini belirtmektedir (Clancy, 2019; A. Madsen et al., 2023; A. N. Madsen et al., 2025). Bu bağlamda açıklanabilir yapay zeka teknikleri, yalnızca algoritmik bir optimizasyon aracı değil, makine dairesindeki vardiya personelinin karar alma süreçlerini destekleyen stratejik bir arayüz olarak değerlendirilmektedir.

Sınai durum ve problemin tanımı

Akademik literatürdeki bu "şeffaflık ve açıklanabilirlik" ihtiyacı, teknoloji geliştiricilerin sınai hamleleri (patentler) incelendiğinde daha da belirgin bir çelişki olarak karşımıza çıkmaktadır. Denizcilik otomasyonu ve filo yönetimi alanında son dönemde tescillenen küresel patentler, kestirimci bakımın donanım entegrasyonunda önemli bir mesafe kat edildiğini, ancak XAI boyutunun ihmal edildiğini göstermektedir.

Bu çalışma kapsamında incelenen güncel sınai durum, beş spesifik patent üzerinden yapılandırılmıştır. Örneğin, makine dairesi ekipmanlarının durum algılaması için k-means algoritması ve geri yayılım (back-propagation) sinir ağlarını kullanan sistemler (Hui et al., 2023) veya TCG Anadolu gibi stratejik unsurlarda da bulunan elektrikli tahrik sistemlerinin (eSiPOD vb.) dönüştürücü arızalarını teşhis etmek için adaptif seyrek filtreleme (sparse filtering) kullanan metotlar (Zicheng et al., 2023), yüksek ölçüm doğruluğuna sahiptir. Benzer şekilde, robotik insansız deniz araçları (İDA) filolarının bakım öngörüsü (H et al., 2023; Howard et al., 2023) ve makine dairesi için sanal gerçeklik tabanlı bilgi füzyonu (Yihuai et al., 2022) sistemleri, sensör verilerini birleştirerek harika tahmin modelleri sunmaktadır.

Ancak problemin odak noktası şudur: Bu sınai çözümlerin tamamı derin öğrenme veya kara kutu makine öğrenmesi yapılarına dayanmaktadır. Sistem, elektrik tahrik konvertöründe bir arıza öngördüğünde, vardiya mühendisine "hangi semistörün aşırı ısındığını" veya "hangi modülasyon indeksinin kritik eşiği aştığını" matematiksel veya görsel bir nedensellikle açıklayamamaktadır.

Bu durum, mühendislik tasarımında evrensel bir çelişki yaratmaktadır: Makine dairesi yapay zeka otomasyonlarında kullanılan derin öğrenme mimarileri, sistemin ölçüm doğruluğunu muazzam ölçüde artırırken; modellerin kara kutu doğası, personelin sistemi anlamasını ve acil durumlarda müdahale etmesini zorlaştırarak kullanım kolaylığını dramatik şekilde düşürmektedir. Bu çalışmanın temel problemi, otonom denizcilik sistemlerindeki bu teknik çelişkiyi çözümleyecek, TCG Anadolu gibi stratejik savunma sanayii unsurlarının operasyonel gerçeklerine uygun, şeffaf ve yerlileştirilebilir bir XAI karar destek mimarisi kurgulamaktır.

Araştırma Metodolojisi

Bu çalışmanın araştırma tasarımı, problemi küresel sınai veriler üzerinden teşhis eden ve matematiksel karar modelleriyle çözüm üreten üç aşamalı hibrit bir yaklaşıma dayanmaktadır. Birinci aşamada bibliyometrik patent analizi ve yapay zeka destekli metin madenciliği kullanılarak veri seti daraltılmış; ikinci aşamada yaratıcı problem çözme teorisi ile teknik çelişkiler modellenmiş; üçüncü aşamada ise TOPSIS yöntemiyle stratejik önceliklendirme kurgulanmıştır.

Veri toplama, filtreleme ve sınai odak seçimi

Küresel denizcilik sektöründe yapay zeka entegrasyonuna yönelik mevcut teknolojik eğilimleri belirlemek amacıyla Lens.org patent veri tabanı kullanılmıştır. Kapsamlı bir bibliyometrik tarama yapabilmek için otonomi, makine dairesi ve arıza teşhisi kavramlarını içeren spesifik bir arama denklemi (boolean query) oluşturulmuştur. Çalışmanın sınırlarını askeri ve ticari denizcilik uygulamalarıyla netleştirmek için aramalar "gemi, denizcilik, makine dairesi" (ship, marine, engine room) gibi zorunlu anahtar kelimelerle kısıtlanmıştır (Tablo 1).

Tablo 1 Lens.org patent arama kriterleri ve filtreleme parametreleri

Parametre	Açıklama
Arama Denklemi (Query)	("explainable AI" OR "XAI" OR "machine learning" OR "deep learning" OR "artificial intelligence") AND ("marine engine" OR "ship propulsion" OR "ship machinery" OR "engine room") AND ("predictive maintenance" OR "fault diagnosis" OR "condition monitoring" OR "automation") AND (ship OR vessel OR marine OR maritime)
Zaman Aralığı	2016 – 2026
Belge Türü	Onaylanmış Patent (Granted) ve Patent Başvurusu (Application)
Ham Sonuç	323 Adet Küresel Patent

Elde edilen 323 patentin tamamı CSV formatında dışa aktarılarak bir ön elemeye tabi tutulmuştur. İlgili patentlerin özet (abstract) metinleri üzerinde otomatik metin taraması uygulanarak, donanım ve arıza teşhisiyle doğrudan bağlantılı olmayan ilgisiz ve gürültülü veriler ayıklanmış; içerisinde engine, fault, diagnosis, maintenance, marine kelimelerinden en az birini barındıran ve atıf sayılarına göre en yüksek etkiye sahip 63 patentlik bir "çekirdek veri seti" (core dataset) oluşturulmuştur.

Çalışmanın stratejik kurgusuna hizmet edecek spesifik teknolojileri belirlemek için, bu 63 patent üzerinde yapay zeka destekli (LLM tabanlı) derin metin analizi uygulanmıştır. Bu seçim aşamasında patentler üç katı kritere göre değerlendirilerek sayı 5'e düşürülmüştür:

Donanım teması (fiziksel bağlam): Sadece teorik bir yazılım modeli değil; jeneratör, valf veya tahrik motoru gibi somut makine dairesi bileşenleriyle entegrasyon.

Kara kutu (black box) zafiyeti: Modelin doğruluğu yüksek olmasına rağmen, derin sinir ağları gibi şeffaflıktan uzak algoritmalar kullanması.

Çift kullanım (dual-use) potansiyeli: Teknolojinin sivil denizcilikten öte, TCG Anadolu ve MİLGEM gibi stratejik savunma unsurlarının operasyonel kurgusuna doğrudan uyarlanabilir olması.

TRIZ tabanlı mühendislik çelişkisi modeli

Seçilen 5 odak patentin ortak teknolojik zafiyeti belirlendikten sonra, sorunun çözümü için Genrich Altshuller tarafından geliştirilen yaratıcı problem çözme teorisi (TRIZ) kullanılmıştır. TRIZ metodolojisi, bir sistemi iyileştirmeye çalışırken sistemin başka bir özelliğinin kötüleşmesi durumunu "teknik çelişki" (technical contradiction) olarak tanımlar ve 39x39 parametre matrisi üzerinden 40 inovatif prensip ile çözümler sunar (Altshuller, 1996).

Bu çalışma kapsamında, derin öğrenme algoritmalarının arıza tespit yeteneğinin artması, TRIZ 39 parametre matrisindeki "iyileşen parametre 28: ölçüm doğruluğu (measurement accuracy)" olarak kodlanmıştır. Buna karşılık, modelin "kara kutu" yapısının vardiya personeli için sistemin anlaşılamaz hale gelmesine neden olması ise "kötüleşen parametre 33: kullanım kolaylığı (ease of operation)" olarak tanımlanmıştır. Bu iki parametrenin kesişimi incelenerek, makine dairesi otomasyonunda şeffaflığı sağlayacak XAI çözüm konseptleri (inovatif prensipler) elde edilmiştir.

Çok kriterli karar verme (TOPSIS) kurgusu

Araştırmanın son aşamasında, TRIZ matrisinden elde edilen teorik XAI çözüm konseptlerini pratik bir savunma sanayii yatırım haritasına dönüştürmek amacıyla TOPSIS (ideal çözüme benzerliğe göre tercih sıralama tekniği) modeli kurulmuştur. Hwang ve Yoon tarafından literatüre kazandırılan TOPSIS, alternatifler arasından pozitif ideal çözüme en yakın ve negatif ideal çözüme en uzak olanı seçme prensibine dayanır (Hwang & Yoon, 1981).

Değerlendirme kurgusunda, TCG Anadolu'nun 40 MW kurulu güce sahip karmaşık makine dairesi ve 50 günlük seyir görevleri (Millisavunma, 2023; T.C. Cumhurbaşkanlığı İletişim Başkanlığı, 2023) dikkate alınarak dört ana kriter belirlenmiştir. Karar matrisindeki alternatif puanlamaları ve kriter ağırlıkları; denizcilik alanında doktora seviyesinde araştırmalar yürüten ve uzakyol vardiya mühendisliği tecrübesine sahip uzman araştırmacı (sorumlu yazar) tarafından, operasyonel saha gerçeklikleri (kısıtlı personel, acil karar gereksinimi ve sınırlı hesaplama altyapısı) gözetilerek değerlendirilmiştir. Bu bağlamda kriterler ve ağırlıkları şu şekilde belirlenmiştir:

K1 - Operasyonel Anlaşılabilirlik (Ağırlık: %35 / fayda kriteri): Başmühendis ve vardiya mühendislerinin model çıktılarını kriz anında eyleme dönüştürebilme hızı.

K2 - Entegrasyon Kolaylığı (Ağırlık: %25 / fayda kriteri): Geliştirilecek modülün mevcut gemi durum izleme (CMS) ve SCADA altyapılarına donanımsal yük getirmeden yazılımsal olarak entegre edilebilmesi.

K3 - Hesaplama Maliyeti (Ağırlık: %20 / maliyet kriteri): XAI algoritmalarının gerçek zamanlı çalışabilmesi için gemi sistemlerinde talep edeceği veri işleme (CPU/GPU) gücü.

K4 - Yerlileştirme Potansiyeli (Ağırlık: %20 / fayda kriteri): Sistemlerin stratejik özerklik prensibi gereği yerli savunma sanayii (ASELSAN, HAVELSAN vb.) firmalarınca dışa bağımsız geliştirilebilirliği.

Oluşturulan karar matrisi üzerinden vektör normalizasyonu yapılmış, ağırlıklandırılmış matris değerleri üzerinden ideal (A*) ve negatif (A-) çözümlere olan Öklid uzaklıkları (S* ve S-) hesaplanarak alternatiflerin yakınlık katsayıları (C*) elde edilmiştir.

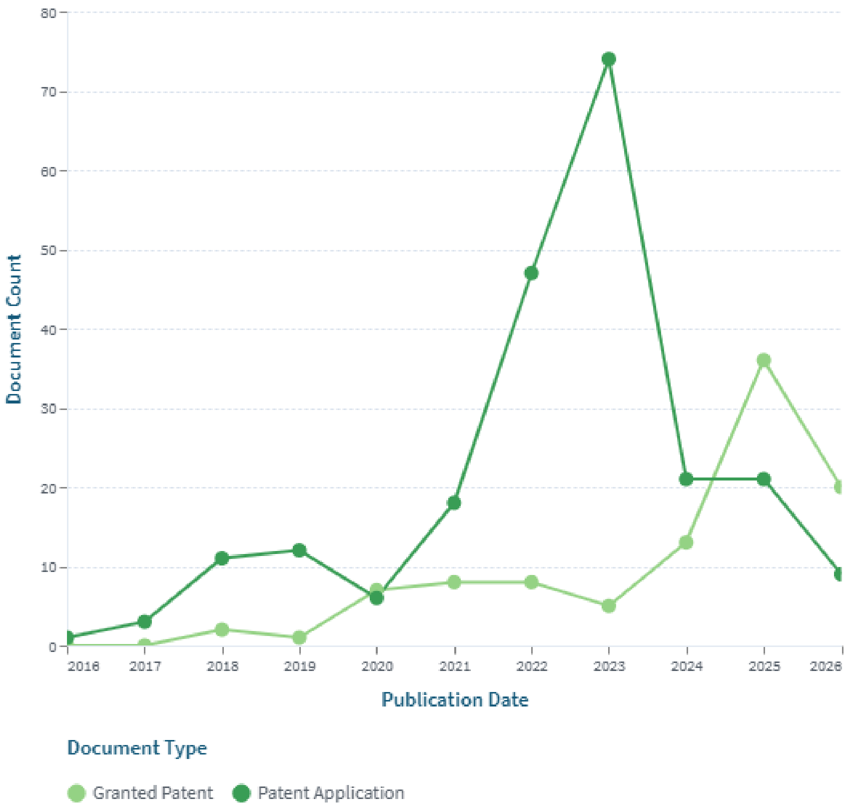
Bulgular ve Tartışma

Bu bölümde, araştırma metodolojisi kapsamında elde edilen sınai patent eğilimleri, TRIZ matrisi üzerinden tanımlanan teknik çelişkiler ve bu çelişkileri çözmek üzere kurgulanıp TOPSIS ile önceliklendirilen XAI stratejilerine ait analitik bulgular sunulmaktadır.

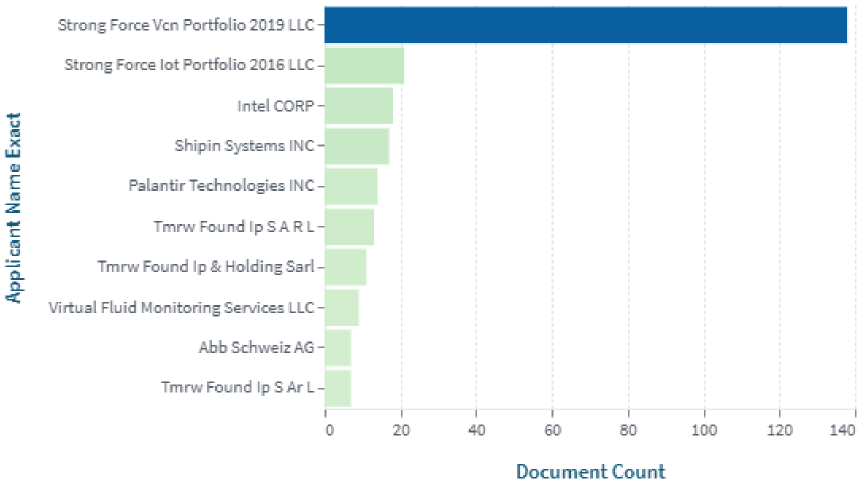
Sınai eğilimler ve patent analizi bulguları

Lens.org veri tabanı üzerinden gerçekleştirilen 2016-2026 yıllarını kapsayan patent analizi, denizcilikte kestirimci bakım teknolojilerindeki "kara kutu" yapılarının baskınlığını ortaya koymaktadır. İlgili veri setinden süzülen 63 odak patentin yıllara göre dağılımı incelendiğinde (Şekil 1), başvuruların doğrusal bir artış eğilimi göstermekten ziyade, belirli yıllarda dalgalanmalar yaşadığı görülmektedir. Özellikle 2023 yılında patent başvurularında belirgin bir zirve noktasına ulaşılmış, ancak takip eden dönemlerde düşüş trendi izlenmiştir. En çok başvuru yapan kurumlar ve şirketler profili (Şekil 2) incelendiğinde ise; listenin ilk sıralarını Strong Force Portfolio, Intel CORP ve Shipin Systems INS gibi aktörlerin oluşturduğu saptanmıştır. Bu veriler, denizcilikte otonomi ve kestirimci bakım teknolojilerindeki sınai bilgi birikiminin (know-how), teknoloji geliştiriciler ve belirli endüstriyel paydaşlar etrafında kümelendiğini göstermektedir. Ancak bulgular, bu sınai büyümenin algoritmik karmaşıklığı artırırken, son kullanıcı olan vardiya personelinin karar alma süreçlerindeki bilişsel yükünü (cognitive load) göz ardı etme eğiliminde olduğunu göstermektedir.

Şekil 1 Patentlerin yıllara göre dağılım



Şekil 2 Lider başvuru sahipleri (şirketler/kurumlar)



Teknolojik odaklanmanın yapısı Uluslararası Patent Sınıflandırma (IPCR) kodları üzerinden incelendiğinde, bu multidisipliner yapının teknik sınırları çok daha net görülmektedir. Süzülen odak patentlerin IPCR sınıflandırma dağılımını gösteren Şekil 3 incelendiğinde; en yoğun başvuru yapılan sınıfların geleneksel denizcilik kodlarının (B63B serisi) yanı sıra, kontrol ve izleme sistemleri ile güvenlik düzenlemelerini içeren G05B ve yapay sinir ağları ile makine öğrenmesi mimarilerini kapsayan G06N kodlarında kümelendiği görülmektedir. Özellikle kontrol sistemlerindeki izleme süreçlerinin G05B doğrudan gelişmiş öğrenme metotları G06N ile entegre edildiği bu tablo, gemi makinelerinde durum izleme süreçlerinin artık tamamen akıllı algoritmalar ekseninde şekillendiğini doğrulamaktadır. Bu dağılım, teknik teşhis süreçlerinin karmaşık ve şeffaflıktan uzak yazılım katmanlarına bağımlı hale geldiğini kanıtlamakta; dolayısıyla TCG Anadolu gibi stratejik muharip unsurlarda operatörün bu otonom kontrol süreçlerini anlamasını sağlayacak Açıklanabilir Yapay Zeka (XAI) ihtiyacının teknik bir zorunluluk olduğunu ortaya koymaktadır.

Şekil 3 Seçilen patentlerin uluslararası patent sınıflandırma (IPCR) kodlarına göre dağılımı

TRIZ çelişki matrisi ve XAI çözüm konseptleri

Sınai durum analizinde tespit edilen temel problem, TRIZ metodolojisi kullanılarak bir mühendislik çelişkisi formatında modellenmiştir. Makine dairesi otomasyonuna entegre edilen yapay zeka modellerinin yarattığı etki, iki ana parametre üzerinden detaylandırılmıştır:

İyileşen Parametre (Parametre 28 - ölçüm doğruluğu / measurement accuracy): Bir sistemin durumunu veya anormalliklerini tespit etme yeteneğindeki kesinliktir. İncelemeye alınan patentlerdeki derin öğrenme modelleri (örneğin adaptif seyrek filtreleme algoritmaları), makine dairesindeki çok boyutlu sensör verilerini eşzamanlı işleyerek insan kapasitesinin ötesinde bir teşhis imkânı sunmaktadır. Örneğin, TCG Anadolu gibi harp gemilerinde yer alan 11 MW kapasiteli elektrikli tahrik motorlarında (eSiPOD) oluşabilecek başlangıç seviyesi rulman yorulmalarını veya güç dönüştürücülerindeki anlık sinyal sapmalarını yüksek doğrulukla tespit edebilmektedir. Bu durum, sistemin "ölçüm doğruluğu" parametresinde belirgin bir iyileşme sağlamaktadır.

Kötüleşen Parametre (Parametre 33 - kullanım kolaylığı / ease of operation): Bir sistemi kontrol etmek, anlamak ve müdahale etmek için gereken insan çabasının ölçütüdür. Yapay zeka modellerinin çok katmanlı (hidden layers) yapısı, karar süreçlerini şeffaflıktan uzaklaştırmaktadır. Örneğin, bir invertör modülündeki toplam harmonik bozulma (THD) seviyesinin sınır değeri aştığı durumlarda, sistemin salt "Arıza olasılığı: %85" şeklinde bir çıktı üretmesi, operasyonel müdahale için yeterli nedenselliği sunmamaktadır. Modelin kararına temel teşkil eden değişkenleri açıklayamaması, vardiya personelinin sisteme duyduğu güveni azaltmakta, acil durumlarda reaksiyon süresini uzatmakta ve "kullanım kolaylığı" parametresini kısıtlamaktadır.

Altshuller'in 39x39 Çelişki Matrisinde, 28 numaralı iyileşen parametre ile 33 numaralı kötüleşen parametrenin kesişimi incelendiğinde, bu çelişkiyi aşmak için dört evrensel prensip (1, 13, 17, 34) önerilmektedir. Bu prensipler, denizcilik otomasyonu bağlamında XAI stratejilerine şu şekilde uyarlanmıştır:

- Prensip 1 (bölme) A1: LIME (bölgesel modeller): Karmaşık ana modelin, yalnızca o anki spesifik arıza uyarısı etrafında doğrusal ve yorumlanabilir yerel (lokal) bir modele indirgenerek operatöre sunulmasıdır.
- Prensip 13 (tersine çevirme) A2: Counterfactual (karşıt olgusal açıklamalar): Durumu olumsuz bir tespit yerine çözüm odaklı, tersine çevrilmiş bir senaryo ile açıklamaktır. Örneğin, sistemin doğrudan "THD oranının %5'in altına çekilmesi arıza olasılığını %80 azaltacaktır"

formatında eyleme dönüştürülebilir "ne olurdu (what-if)" senaryoları üretmesidir.

- Prensip 17 (başka boyuta geçme) A3: 3D SHAP (ısı haritaları): Sayısal veri çıktılarının çok boyutlu görsel bir düzleme taşınmasıdır. Makine dairesinin dijital ikizi (digital twin) üzerinde, arızaya sebep olan sensörlerin veya bileşenlerin SHAP değerlerine göre renklendirilerek (ısı haritası) gösterilmesidir.
- Prensip 34 (atma ve geri kazanma) A4: Dinamik özellik seçimi: Karar anında doğrudan etkiye sahip olmayan veri setlerinin (gürültünün) dışarıda bırakılarak, sadece modeli tetikleyen en kritik sensör parametrelerinin operatör ekranında tutulmasıdır.

TOPSIS stratejik önceliklendirme sonuçları

Geliştirilen dört XAI stratejisinin (A1, A2, A3, A4) askeri deniz gemilerine entegrasyonu, TCG Anadolu vizyonu çerçevesinde belirlenen dört temel kriter üzerinden TOPSIS yöntemiyle değerlendirilmiştir. Karar modelinde yer alan kriterler ve ağırlıkları (W) şu şekildedir: operasyonel anlaşılabilirlik (K1, %35, fayda), entegrasyon kolaylığı (K2, %25, fayda), hesaplama maliyeti (K3, %20, maliyet) ve yerlileştirme potansiyeli (K4, %20, fayda).

Çalışma kapsamında, denizcilik operasyonları ve makine dairesi dinamikleri bağlamında alternatiflerin 1-9 arası (1: çok düşük, 9: çok yüksek) Likert ölçeğiyle değerlendirildiği karar matrisi (decision matrix) Tablo 2'de sunulmaktadır.

Tablo 2 XAI Alternatifleri için TOPSIS karar matrisi

Alternatifler	K1: Operasyonel Anlaşılabilirlik	K2: Entegrasyon Kolaylığı	K3: Hesaplama Maliyeti	K4: Yerlileştirme Potansiyeli
A1	6	8	6	7
A2	8	6	7	6
A3	7	5	8	6
A4	5	7	3	8

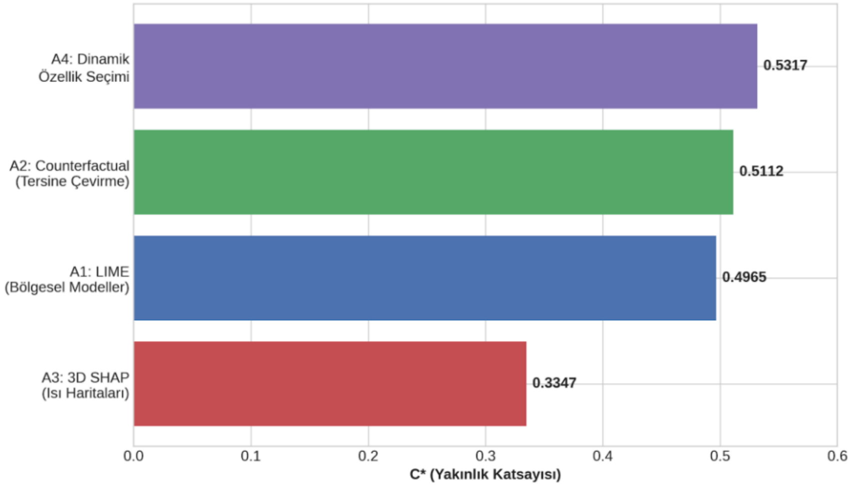
Oluşturulan karar matrisi üzerinden vektör normalizasyonu yapılmış ve kriter ağırlıkları uygulanmıştır. Pozitif ideal (A*) ve negatif ideal (A-) çözümlere olan Öklid uzaklıkları (S* ve S-) ile nihai yakınlık katsayıları (C*) hesaplanarak Tablo 3'te özetlenmiştir.

Tablo 3 XAI stratejileri için ideale uzaklık ve TOPSIS sıralama sonuçları

Sıra	Alternatifler	S* (İdeale Uzaklık)	S- (Negatife Uzaklık)	C* (Yakınlık Katsayısı)
1	A4: Dinamik Özellik Seçimi	0.0818	0.0929	0.5317
2	A2: Counterfactual (Tersine Çevirme)	0.0797	0.0834	0.5112
3	A1: LIME (Bölgesel Modeller)	0.0729	0.0719	0.4965
4	A3: 3D SHAP (Isı Haritaları)	0.1055	0.0531	0.3347

Elde edilen C* skorlarına göre, A4 (dinamik özellik seçimi) alternatifinin ideal çözüme en yakın strateji olduğu saptanmıştır. Bu durum Şekil 4'te sunulan sıralama grafiğinde de net biçimde görülmektedir.

Şekil 4 TOPSIS sıralaması C skorları



Analiz bulguları derinlemesine incelendiğinde, A4 alternatifinin ilk sıraya yerleşmesindeki temel rasyonelin "hesaplama maliyeti" (K3) ve "yerlileştirme potansiyeli" (K4) kriterlerindeki üstünlüğü olduğu anlaşılmaktadır. Görsel açıdan yüksek anlaşılabilirliğe sahip A3 (3D SHAP) yöntemi, gemideki kısıtlı işlemci gücü kaynaklarını (CPU/GPU) yoğun tüketeceği ve mevcut CMS/SCADA altyapılarına donanımsal entegrasyonu zor olduğu için sıralamada sonuncu olmuştur. Buna karşılık, dinamik özellik seçimi (A4), yapay zeka modelini hafifleterek gemi bilgisayarlarında gerçek zamanlı çalışabilme imkânı sunmakta ve dışa bağımlılık yaratmadan yerli algoritmalarla geliştirilebilme avantajı sağlamaktadır. İkinci sıradaki A2 (counterfactual) alternatifi ise, denizcilerin kriz anındaki karar verme dinamiğine en uygun "neden-sonuç" ilişkisini sunduğu için operasyonel anlaşılabilirlik (K1) bağlamında güçlü bir stratejik unsur olarak öne çıkmaktadır.

Sonuç ve Stratejik Öneriler

Bu çalışma, denizcilik endüstrisinde kestirimci bakım teknolojilerinin önündeki en büyük engellerden biri olan "kara kutu" zafiyetine analitik bir çözüm mimarisi sunmaktadır. Küresel patent eğilimlerinden yola çıkarak tanımlanan teknik çelişkiler, TRIZ metodolojisiyle çözümlenmiş ve TCG Anadolu vizyonuna uygun açıklanabilir yapay zeka stratejileri TOPSIS yöntemiyle önceliklendirilmiştir.

Analiz bulguları, otonom gemi makine dairelerinde yüksek doğruluklu (parametre 28) bir teşhis sistemi kurmanın, sistemin anlaşılabilirliğini (parametre 33) feda etmeden mümkün olduğunu kanıtlamaktadır. TOPSIS sıralama sonuçları, dinamik özellik seçimi (A4) stratejisinin hesaplama verimliliği ve yerlileştirme potansiyeli bakımından en rasyonel yatırım alanı olduğunu göstermektedir. Ancak denizcilik operasyonlarının fiziksel ve psikolojik dinamikleri, salt veri sadeleştirmenin ötesinde, nedenselliğe dayalı bir operatör desteği gerektirmektedir.

Bu bulgular ışığında, askeri deniz araçları için önerilen stratejik teknoloji yol haritası şu şekildedir:

Hibrit XAI mimarisinin kurulması: Tekil bir yöntem yerine, A4 (dinamik özellik seçimi) ve A2 (counterfactual / karşıt olgusal açıklamalar) stratejilerinin entegre edildiği hibrit bir yapı tercih edilmelidir. Bu mimaride, A4 stratejisi "uç bilişim" (edge-computing) katmanında çalışarak geminin karmaşık sensör ağındaki (THD seviyeleri, vibrasyon spektrumları vb.) veri gürültüsünü süzmeli; A2 stratejisi ise elde edilen bu konsantre veriyi vardiya mühendisine "what-if" senaryoları olarak sunmalıdır. Örneğin, sistemin "Eğer 2 numaralı tahrik motoru invertöründeki anahtarlama frekansı %10 düşürülürse, aşırı ısınma riski ortadan kalkacaktır" şeklindeki yönlendirmesi, personelin modele olan güvenini ve müdahale hızını artıracaktır.

Donanımsal özerklik ve entegrasyon: TCG Anadolu gibi stratejik unsurların 50 günü aşan kesintisiz seyir görevleri, sistemlerin dış bulut bağlantısına ihtiyaç duymadan, gemi üzerindeki mevcut SCADA ve durum izleme (CMS) altyapısında çalışmasını zorunlu kılmaktadır. TOPSIS analizinde maliyet kriterinin (K3) belirleyici etkisi, hafifletilmiş XAI modellerinin (A4 gibi) geliştirilmesinin stratejik bir öncelik olduğunu göstermektedir. Bu yaklaşım, yüksek işlemci gücü gerektiren yabancı menşeli grafik yoğunluklu çözümlere (A3 gibi) olan bağımlılığı ortadan kaldıracaktır.

Savunma sanayii ekosisteminde yerlileştirme: Karar matrisinde öne çıkan "yerlileştirme potansiyeli" (K4), önerilen XAI modellerinin Türk savunma sanayii paydaşları (ASELSAN, HAVELSAN, STM vb.) tarafından milli algoritmalarla geliştirilmesinin önünü açmaktadır. Açıklanabilir mimariler, siber güvenlik denetimlerine ve validasyon süreçlerine daha uygun oldukları için askeri standartlara adaptasyonları çok daha hızlı gerçekleşecektir.

Günümüzde TCG Anadolu gibi modern donanma unsurlarında gemi entegre kontrol ve izleme sistemi (EPKİS) veya muadili gelişmiş otomasyon altyapıları ile sensör bazlı anlık durum izleme yapılabilmektedir; ancak modelin kendi kararlarını operatöre açıklayabildiği açıklanabilir yapay zeka destekli dinamik kestirimci bakım modülleri henüz standart askeri donanım kapsamında yer almamaktadır. Oysa Mavi Vatan doktrini çerçevesinde inşa edilen yerli deniz sistemlerinin, okyanus ötesi görevlerde teknolojik özerkliklerini koruyarak tam operasyonel hazırlık seviyesinde tutulabilmesi için yapay zekada "şeffaflık ve operatör uyumu" bir lüks değil, stratejik bir gerekliliktir. Dolayısıyla bu çalışma ile sunulan TRIZ-TOPSIS tabanlı karar modeli; deniz kuvvetlerinin mevcut otomasyon sistemlerinin gelecekteki modernizasyon süreçleri, tam otonomizasyona geçiş adımları ve kullanılacak şeffaf karar destek mekanizmaları için bilimsel ve uygulanabilir bir vizyon belgesi niteliği taşıyarak sağlam bir temel oluşturmaktadır.

Kaynakça

Altshuller, G. (1996). And suddenly the inventor appeared: TRIZ, the theory of inventive problem solving. Technical Innovation Center, Inc.

Cheliotis, M., Lazakis, I., & Cheliotis, A. (2022). Bayesian and machine learning-based fault detection and diagnostics for marine applications. Ships and Offshore Structures, 17(12), 2686–2698. https://doi.org/10.1080/17445302.2021.2012015

Clancy, J. P. (2019). BREAKDOWNS IN HUMAN-AI PARTNERSHIP: REVELATORY CASES OF AUTOMATION BIAS IN AUTONOMOUS VEHICLE ACCIDENTS [University of North Carolina]. https://doi.org/10.17615/jpah-hc02

H, C. C., Brad, K., S, E.-T. T., Andrew, C., & Leon, F. J. R. (2023). Maintenance Prediction and Health Monitoring for Robotic Fleet Management (US20230109096A1). U.S. Patent and Trademark Office. https://lens.org/066-396-403-458-98X

Howard, C. C., S, E.-T. T., & Leon, F. J. R. (2023). Artificial-Intelligence-Based Preventative Maintenance for Robotic Fleet (US20230222454A1). U.S. Patent and Trademark Office. https://lens.org/073-109-907-911-687

Hui, C., Mengzhuo, H., Cong, G., & Xiangguo, Y. (2023). Cabin equipment state perception and fault diagnosis method based on K-means algorithm and BP neural network (CN115879018A). China National Intellectual Property Administration. https://lens.org/061-731-459-675-91X

Hwang, C.-L., & Yoon, K. (1981). Multiple Attribute Decision Making (Vol. 186). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-48318-9

Je-Gal, H., Park, Y.-S., Park, S.-H., Kim, J.-U., Yang, J.-H., Kim, S., & Lee, H.-S. (2024). Time-Series Explanatory Fault Prediction Framework for Marine Main Engine Using Explainable Artificial Intelligence. Journal of Marine Science and Engineering, 12(8), 1296. https://doi.org/10.3390/jmse12081296

Kalafatelis, A. S., Levis, G., Giannopoulos, A., Tsoulakos, N., & Trakadas, P. (2025). Explainable Predictive Maintenance of Marine Engines Using a Hybrid BiLSTM-Attention-Kolmogorov Arnold Network. Journal of Marine Science and Engineering, 14(1), 32. https://doi.org/10.3390/jmse14010032

Madsen, A., Brandsæter, A., & V Aarset, M. (2023). Decision Transparency for enhanced human-machine collaboration for autonomous ships. 14th International Conference on Applied Human Factors and Ergonomics. https://doi.org/10.54941/ahfe1003750

Madsen, A. N., Brandsæter, A., van de Merwe, K., & Park, J. (2025). Improving decision transparency in autonomous maritime collision avoidance. Journal of Marine Science and Technology, 30(1), 134–152. https://doi.org/10.1007/s00773-024-01043-x

Millisavunma. (2023). TCG Anadolu Çok Maksatlı Amfibi Hücum Gemisi (LHD). https://www.millisavunma.com/tcg-anadolu-cok-maksatli-amfibi-hucum-gemisi-lhd/

Park, M., Kim, H., Lee, S., & Lee, J. (2023). Anomaly Detection and Root Cause Analysis of Ship Main Engines: Explainable Artificial Intelligence-Based Methodology Considering Internal Sensors and External Environmental Factors. Journal of the Korean Institute of Industrial Engineers, 49(5), 379–394. https://doi.org/10.7232/JKIIE.2023.49.5.379

Simion, D., Postolache, F., Fleacă, B., & Fleacă, E. (2024). AI-Driven Predictive Maintenance in Modern Maritime Transport—

Enhancing Operational Efficiency and Reliability. Applied Sciences, 14(20), 9439. https://doi.org/10.3390/app14209439

T.C. Cumhurbaşkanlığı İletişim Başkanlığı. (2023). Erdoğan, deniz devini uğurladı… Türkiye'nin en büyük savaş gemisi Karadeniz'e doğru yola çıktı. https://www.iletisim.gov.tr/turkce/dis_basinda_turkiye/detay/erdogan-deniz-devini-ugurladi-turkiyenin-en-buyuk-savas-gemisi-karadenize-dogru-yola-cikti

Wang, B., Wang, Z., Yao, C., Chen, J., Lu, L., & Song, E. (2025). Multi-system condition monitoring of marine engines: A unified deep learning framework introducing physical prior knowledge. International Journal of Naval Architecture and Ocean Engineering, 17, 100698. https://doi.org/10.1016/j.ijnaoe.2025.100698

Yihuai, H., Jiawwei, J., & Lei, G. (2022). Remote Monitoring Method based on Information Fusion and Virtual Reality for Marine Engine Room (US20220366657A1). U.S. Patent and Trademark Office. https://lens.org/060-013-899-230-532

Zemmouchi-Ghomari, L. (2026). Explainable AI for predictive maintenance: A review and standardized evaluation framework. Management Science Letters, 16(1), 15–36. https://doi.org/10.5267/j.msl.2025.11.001

Zicheng, L., Lanlan, F., She, Y., & Dong, J. (2023). Motor Driving System Converter Fault Diagnosis Method Based On Adaptive Sparse Filtering (US20230163713A1). U.S. Patent and Trademark Office. https://lens.org/131-338-557-900-056

GENERATIVE AI IN MARITIME EDUCATION AND TRAINING: MAPPING APPLICATIONS AND EVALUATION METHODOLOGIES

Hüseyin Fatih NACAR[1]

Introduction

Maritime transport carries more than 80 percent of international cargo trade by volume (Li et al., 2023), and the safety of that system rests to an uncomfortable degree on how well its people are trained. Human error accounts for approximately 70 percent of maritime accident cases, and analyses tie this share to inadequate training (Karahalios, 2025). Maritime education and training (MET) is therefore not a peripheral service function but a safety barrier, and its content is under pressure: as ship systems become more automated, calls have emerged for MET to add digital and AI related competencies to those already required under the Standards of Training, Certification, and Watchkeeping (STCW) framework (Sharma et al., 2022).

The technology side of that pressure has changed character since late 2022. Generative artificial intelligence (GenAI) has been

[1] Research Assistant, Istanbul Technical University, Department of Marine Engineering, ORCID: 0009-0001-7178-141X

adopted at a speed with few precedents in the history of technology (Giannakos et al., 2025), and a position paper argued that large language models are not a passing phase and will persist despite local bans (Kasneci et al., 2023). For professional training in particular, these models promise support for competencies such as field specific language, report writing, and decision making (Kasneci et al., 2023), and the prevailing position in the educational literature is to integrate the tools rather than attempt to suppress them (Grassini, 2023), while warning against adoption without evidence of efficacy and pedagogical soundness (Giannakos et al., 2025). Whether and how this technology wave reaches a safety critical, internationally regulated training domain such as MET is, on this view, an empirical question rather than a foregone conclusion.

For education in general, the evidence base needed to answer such questions is well developed. A systematic review of artificial intelligence in higher education synthesised 146 articles as early as 2019 (Zawacki-Richter et al., 2019), and a scoping review of large language models in education included 118 empirical studies published between 2017 and 2022 (Yan et al., 2023). No comparable synthesis was identified for MET. Researchers in the field themselves report that few published works have addressed artificial intelligence in maritime education (Diahyleva et al., 2025), a recent review found no studies connecting the concepts of autonomous shipping and Society 5.0 to MET (Ahmad Fuad et al., 2025), and the one systematic review identified confines its scope to AI and machine learning techniques within adaptive instructional systems (Karimi et al., 2024). What has accumulated instead is a scatter of individual studies, and it is currently difficult to say which applications have been tried in maritime classrooms and simulators, with what methods they were evaluated, and where the untested territory lies.

The timing makes this more than a bibliographic inconvenience. Analyses of maritime autonomous surface ships (MASS) anticipate a restructuring of seafarer competence requirements and of the curricula that produce them (Wang et al., 2020), a review of STCW development concluded that seafarer training lags behind the speed of technological change (Yi et al., 2025), and interview research with maritime stakeholders judged that MET institutes are not yet ready to train operators of autonomous ships (Emad & Ghosh, 2023). Training providers facing this demand will make curriculum decisions either way; without a map of the existing evidence, those decisions risk being made blind.

This chapter provides such a map. It reports a narrative scoping synthesis of GenAI and related AI applications in MET, built on 44 publications spanning 2017 to 2026, and asks three connected questions: which application paradigms have been investigated in maritime training contexts, which methodology families have been used to evaluate them, and where the cross-tabulation of these two axes leaves combinations untested. The next section describes the review design and coding. The three results sections that follow report the application typology, the methodology inventory, and the coverage matrix. The discussion and conclusion then read the observed patterns and set out a research agenda.

Methods

Review design

This chapter reports a narrative scoping synthesis of research on generative artificial intelligence (GenAI) and related AI applications in maritime education and training (MET). A scoping approach was chosen because it fits a young and fast moving body of research, where the goal is to chart what has been studied and with which methods rather than to test a narrowly defined effect (Yan et al., 2023). As a book chapter contribution, no strict systematic protocol was employed; comparable design choices are documented in recent rapid and non systematic evidence summaries on GenAI in education (Giannakos et al., 2025; Lo, 2023). Such summaries trade steps of the full systematic process for timeliness (Lo, 2023); the same trade off applies here, and its consequences are examined in the Discussion section. Earlier narrative reviews of AI in education have applied and defended comparable qualitative review designs (Chen et al., 2020).

Search and selection

The literature was identified through iterative keyword searches and citation chaining rather than a single registered search string. Search terms combined the maritime education domain (maritime education, MET, seafarer competency, Maritime English, COLREGs, MASS, simulator training) with GenAI and AI terms (GenAI, ChatGPT, large language model, chatbot, intelligent tutoring system) and with review descriptors for the comparator literature (AI in education, scoping review, systematic review). Fifty candidate documents were retrieved. Six records were excluded: two duplicates published under different identifiers, one full journal issue from which the target article could not be isolated, and three records that could not be obtained in usable full text. The remaining 44 publications form the included study set.

In order to be included, a publication had to address GenAI or AI in an MET context, or to serve one of two defined comparator roles: syntheses and position papers on AI in education broadly, and empirical transfer studies from adjacent domains that share high stakes operational training with MET (healthcare, construction safety, and driver education). Purely technical work without an education link was excluded; one borderline study of COLREGs compliant collision avoidance was retained as a proxy for rule based training content (Sun et al., 2023). The included studies span 2017 to 2026, from an early position paper on AI in higher education (Popenici & Kerr, 2017) to a set of worked classroom examples of ChatGPT use in Maritime English lessons (Pindosova & Fedorova, 2026). Of the 44 publications, 22 address maritime education directly, 15 are reviews or position papers on AI in education, four are empirical AIEd studies, and three are transfer studies from adjacent domains.

Coding

Each included study was coded on two axes. The first axis records the application paradigm investigated, using six categories derived inductively from the maritime studies: conversational agents, intelligent tutoring and student modeling, predictive analytics, simulator augmented training, gamified mobile learning, and assessment automation (Table 1). The second axis records the methodology family, using eight families that consolidate the twelve study designs observed in the included set (Table 2). Studies combining elements of two paradigms were assigned to the paradigm matching their stated primary aim; the COLREGs chatbot study (Sharma et al., 2022), which embeds tutoring elements, was coded as a conversational agent. Position and commentary papers were not assigned to a methodology family, since they report no study design; they instead inform the interpretive sections of the chapter. The two axes were then cross-tabulated for the maritime and transfer studies only, so that the coverage matrix reflects evidence produced in or near the target domain rather than the broader AIEd background.

Boundaries of the approach

Four boundaries follow from these choices. The search was iterative rather than registered, coding was performed by a single analyst, the included set is closed rather than continuously updated, and no formal quality appraisal was applied to the included studies.

In addition, the three records excluded on availability grounds may narrow coverage of their topics.

Application typology

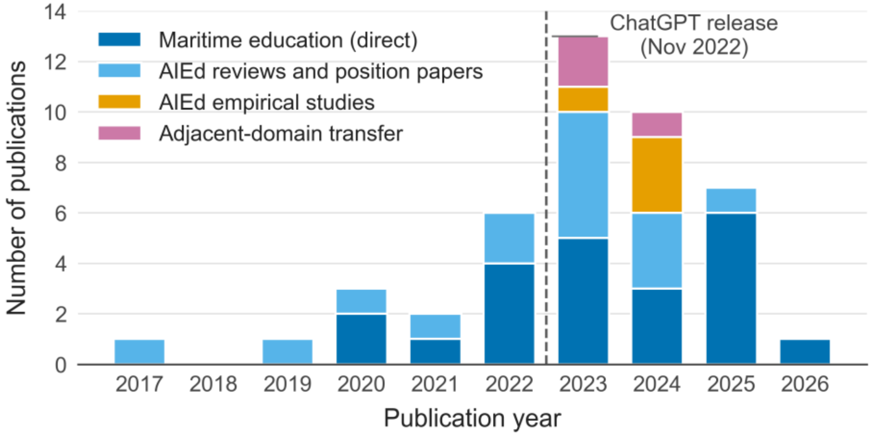
Table 1 presents the six application paradigms identified in the included studies, together with their exemplar publications and evidence types. Of the 22 maritime publications, 14 implemented or evaluated a specific application: five investigated conversational agents, two intelligent tutoring and student modeling, three predictive analytics, three simulator augmented training (including the collision avoidance proxy defined in the Methods section), and one gamified mobile learning. No maritime study was identified with assessment or feedback automation as its primary aim; the evidence for that paradigm comes from two transfer studies. The remaining eight maritime publications addressed competency requirements, curricula, or the research landscape without implementing an application. Thirteen of the 14 application studies appeared in 2022 or later (Figure 1); the exception is an intelligent tutoring study from 2021 (Yang et al., 2021).

Table 1. Application paradigms of GenAI and AI in maritime education and training, with exemplar studies and primary evaluation designs (N = maritime / transfer studies).

Paradigm	*Exemplar studies*	*N (M / T)*	*Typical GenAI/AI role*	*Primary evaluation designs*
Conversational agents	*Sharma et al. (2022) Khlebalov et al. (2022) Pindosova & Fedorova (2026) Diahyleva et al. (2025)*	*5 / 0*	*Dialogue partner for COLREGs and Maritime English practice*	*Usability evaluation; classroom interventions; two papers without evaluation*

Paradigm	Exemplar studies	N (M / T)	Typical GenAI/AI role	Primary evaluation designs
	Ratnaningsih (2025)			
Intelligent tutoring and student modeling	*Yang et al. (2021)* *Karimi et al. (2024)*	*2 / 0*	*Student modeling, adaptive instruction*	*Computational experiment; evidence synthesis*
Predictive analytics	*Munim et al. (2025)* *Tusher et al. (2023)* *Sellberg & Sharma (2024)*	*3 / 0*	*Performance prediction, at-risk identification*	*Computational experiments; design ethnography*
Simulator-augmented training	*Dewan et al. (2023)* *Hwang & Youn (2022)* *Sun et al. (2023)*	*3 / 0*	*Scenario generation, immersive training content*	*Computational experiments; evidence synthesis*
Gamified mobile learning	*Diahyleva et al. (2024)*	*1 / 0*	*Engagement mechanics in mobile course delivery*	*Classroom intervention*
Assessment and feedback automation	*Murtaza et al. (2024)* *Uddin et al. (2023)*	*0 / 2*	*Automated feedback, hazard recognition training*	*Transfer interventions only*
Total		*14 / 2*		

Figure 1. Distribution of the 44 included publications by publication year and context layer; the dashed line marks the public release of ChatGPT in November 2022.



Conversational agents

Five maritime studies investigated conversational agents, the largest group in the typology, and their designs range from scripted bots built before large language models to ChatGPT based classroom use. In the broader educational literature, more than half of 36 reviewed chatbot studies deployed the chatbot as a teaching agent, most following a predetermined conversational path (Kuhail et al., 2022).

The two earliest maritime systems were purpose built. A voice chatbot for Maritime English was implemented on a messaging platform and organised as three sub dialogues, a translator and voice cards for words and for sentences (Khlebalov et al., 2022). FLOKI, a chatbot constructed with a commercial assistant service for COLREGs training, covered Rules 11 to 18 of the 41 rules; in a usability evaluation with 18 cadets it scored 73.72 on the System Usability Scale, above the published benchmark median of 70.5, with high internal consistency (Cronbach's alpha = 0.884)

(Sharma et al., 2022). Prior navigation experience produced no significant difference in the usability ratings ($p = 0.896$), and prior chatbot experience trended higher without reaching significance ($p = 0.123$) (Sharma et al., 2022).

The studies published after the release of ChatGPT shifted from building agents to deploying a general purpose model in the classroom. One account presented five worked task templates for Maritime English lessons, including SMCP role play on the transport of dangerous goods, structured question lists referencing the IMDG Code, and chain stories built on Vessel Traffic Service vocabulary; no learner outcome data were reported (Pindosova & Fedorova, 2026). A quasi experimental study followed 153 marine engineering cadets over one semester, with 76 in the experimental group and 77 in a matched control group, and used a ChatGPT based environment to simulate professional dialogues, generate quizzes, and produce term lists that were transferred to a Moodle glossary; the experimental group made greater progress on the post tests, although no inferential significance tests were reported (Diahyleva et al., 2025). A classroom action research cycle with 69 polytechnic students found statistically significant gains ($p < 0.05$) in grammar (+12.86), vocabulary (+12.30), fluency (+12.70), and comprehension (+12.44), with the overall mean rising from 74.80 to 87.00; students rated the helpfulness of ChatGPT at 4.6 out of 5 (Ratnaningsih, 2025).

Intelligent tutoring and student modeling

Two maritime studies addressed this paradigm. The earlier one proposed a five step methodology for machine learning based student modeling, covering data preprocessing, feature engineering, algorithm selection, evaluation, and deployment, and trained eight candidate models on a navigation tutoring dataset of 1,560 samples under 4-fold cross validation (Yang et al., 2021). A Naive Bayes model with a cost matrix performed best, at 95% accuracy with 94% sensitivity and specificity, and a domain engineered risk feature received the highest permutation importance score (Yang et al., 2021).

The second study, a systematic review of AI and machine learning in adaptive instructional systems for MET, identified decision trees as the most prevalent technique (10 articles), followed by artificial neural networks (9) and support vector machines (8) (Karimi et al., 2024). The reviewed techniques predominantly supported the student model and the pedagogical model, and decision trees and k-nearest neighbours appeared most often among the techniques suited to small datasets; the cost of collecting human performance data in simulator settings was identified as a persistent constraint (Karimi et al., 2024).

Predictive analytics

Three maritime studies applied predictive modeling to learner data. In the most extensive of these, 58 algorithms were trained on simulator log data to predict student performance in Williamson Turn exercises, and an XGBoost classifier was adopted; rudder angle and propeller revolutions per minute emerged as the most influential predictors in the ballast and loaded conditions, respectively (Munim et al., 2025). The authors positioned the work as one of the first validations of an objective, machine learning based assessment approach using simulator logs (Munim et al., 2025).

An AutoML platform was used in the second study to predict final graduation grades of maritime students from their early semester exam grades; nine candidate models were screened on RMSE and R-squared, and a complementary k-means analysis yielded four clusters that flag students at risk of low final grades (Tusher et al., 2023). The third study, a design ethnography of simulator instruction, examined which parts of instructor work multimodal learning analytics could automate; it identified a recurring feedback pattern based on continuous monitoring and embedded assessment, and reported that speech recognition can supply stable records of bridge team communications (Sellberg & Sharma, 2024). It is assigned here under the primary aim rule.

Simulator augmented and immersive training

Three studies fall under this paradigm. A PRISMA guided review retrieved 27 articles on immersive and non-immersive simulators in MET over the preceding decade and found that only nine of the 27 compared simulator training against traditional training; the same review documented the falling cost of VR hardware, from roughly 50,000 USD in 2006 to 1,300 USD in 2014 (Dewan et al., 2023).

The second study generated training scenarios for remote operators of autonomous ships from actual ship trajectory data. Course altering angles extracted from South Korean AIS records followed a symmetrical normal distribution between -90 and +90 degrees, while distances between waypoints followed a right skewed inverse Gaussian distribution from 1 to 20 nautical miles; small values dominated, with 9 degree alterations accounting for 48.8% of cases and 2 nautical mile distances for 65.0% (Hwang & Youn, 2022). Permutation of the proportion weighted representative values produced 564,480 candidate navigation scenarios (Hwang & Youn, 2022).

The third entry is the collision avoidance proxy. A deep reinforcement learning agent for unmanned surface vehicles encoded COLREGs constraints in its reward signal and used a noisy network to increase exploration; across 100 runs under a fixed random seed, its success rate exceeded the compared baselines (Sun et al., 2023). It is included only as a proxy for rule based training content, as set out in the Methods section.

Gamified mobile learning

A Maritime English course on the Moodle mobile app was enriched with leaderboards, badges, points, levels, missions, maps, and scenarios, and was evaluated with first year marine engineering cadets at A2 to B1 English level, divided into a control group of 37 and an experimental group of 36 (Diahyleva et al., 2024). The experimental group improved its class attendance and subject knowledge by 43.7%, against 7.8% in the control group (Diahyleva et al., 2024).

Assessment and feedback automation

No included maritime study investigated automated assessment or feedback as its primary aim; the evidence comes from the AIEd literature and two transfer studies. A systematic scoping review of 118 empirical studies on large language models in education catalogued the educational tasks being automated, grouped into nine categories (Yan et al., 2023). A rapid review of early ChatGPT research reported that performance varied across subject domains, from outstanding in economics to unsatisfactory in mathematics, that the model is a poor judge of its own correctness, and that 40 of 50 generated essays passed plagiarism screening with similarity scores of 20% or less (Lo, 2023).

The two transfer studies tested feedback oriented uses in high stakes training. In driver education, a group trained with a ChatGPT instance restricted to the vehicle manuals reached accuracies of 80 to 100% across all five examined driving assistance functions, with significant differences from conventional training in the three functions compared statistically (p = 0.014, 0.008, and 0.034) (Murtaza et al., 2024). In construction safety education, students recognised fewer than 35% of hazards before a ChatGPT intervention and more than 60% afterwards, and over 95% agreed that the tool can support safety training (Uddin et al., 2023).

Methodology landscape

Table 2 summarises the eight methodology families observed in the included studies. Of the 25 maritime and transfer publications, 22 were assigned a primary methodology family. Evidence synthesis was the most frequent family, with seven publications, followed by intervention designs and computational experiments with five each; qualitative and ethnographic designs accounted for two, and standardised usability instruments, cross sectional surveys, and expert consensus methods for one each. No maritime or transfer study was classed as mixed methods under the primary assignment rule; mixed methods designs appear only in the AIEd comparator layer. The remaining three publications could not be assigned a family: a system design report (Khlebalov et al., 2022) and a set of worked classroom examples (Pindosova & Fedorova, 2026), both presented without any user evaluation, and an analytical paper on MASS competence requirements with no empirical design (Wang et al., 2020).

Table 2. Methodology families across the 25 maritime and transfer publications, with exemplar studies, typical instruments, and measured outcomes.

Family	N	Maritime and transfer exemplars	Typical instruments and designs	What is measured
M1 Perception and acceptance	1	Sharma et al. (2022), SUS 73.72, n = 18	System Usability Scale, subgroup comparison tests	Usability, perceptions
M2 Intervention designs	5	Diahyleva et al. (2024), n = 73 Diahyleva et al. (2025), n = 153 Ratnaningsih (2025), n = 69	Pre/post testing, classroom action research, matched-group comparison	Learning outcomes

Family	N	Maritime and transfer exemplars	Typical instruments and designs	What is measured
		Murtaza et al. (2024) Uddin et al. (2023)		
M3 Cross-sectional surveys	1	Karahalios (2025)	Digital-skills questionnaire across MSc student and ship cadet groups	Self-reported training
M4 Mixed methods	0	None in the maritime and transfer set	Survey and interview combinations (AIEd layer only)	Perceptions and experience
M5 Qualitative and ethnographic	2	Emad & Ghosh (2023), 37 stakeholders Sellberg & Sharma (2024)	Structured interviews, design ethnography	Skill requirements, instructional practice
M6 Expert consensus	1	Kim & Mallam (2020), 36 experts	Delphi with analytic hierarchy process	Competency priorities
M7 Computational experiments	5	Munim et al. (2025) Tusher et al. (2023) Yang et al. (2021) Hwang & Youn (2022) Sun et al. (2023)	AutoML screening, classifier benchmarks, scenario permutation, deep reinforcement learning	Model-level metrics
M8 Evidence synthesis	7	Karimi et al. (2024) Dewan et al. (2023), 27 articles Li et al. (2023) Munim & Haralambides (2022) Yi et al. (2025)	PRISMA guided and flow-diagram documented reviews, bibliometrics, regulatory reviews	Research landscape

Family	N	Maritime and transfer exemplars	Typical instruments and designs	What is measured
		Ahmad Fuad et al. (2025) Sallam (2023)		
No family assigned	3	Khlebalov et al. (2022) Pindosova & Fedorova (2026) Wang et al. (2020)	System description, worked examples, analytical commentary	No evaluation reported
Total	**25**			

Perception and acceptance measures

One maritime study used a standardised instrument as its primary method. The COLREGs chatbot evaluation administered the System Usability Scale to 18 cadets, reported a score of 73.72 with high internal consistency (Cronbach's alpha = 0.884), and compared subgroups with Mann-Whitney tests, none of which reached significance (Sharma et al., 2022). Perception measures also appear inside the intervention studies: the Maritime English action research collected Likert ratings alongside its tests, with helpfulness rated 4.6 and the relevance of the maritime scenarios 4.5 out of 5 (Ratnaningsih, 2025), and the construction safety study recorded agreement above 95% on items about the usefulness of the tool (Uddin et al., 2023).

The most extensive acceptance instrument in the included studies, by item count and sample, comes from the AIEd layer. A study of 300 pre-service teachers used a 41 item questionnaire based on an extended technology acceptance model with nine constructs, analysed through partial least squares structural equation modeling, and found high acceptance of ChatGPT shaped by personal competency, social influence, perceived usefulness, enjoyment, trust, perceived AI intelligence, positive attitude, and metacognitive self regulated learning (Dahri et al., 2024). Across all of these cases, the instruments captured self reported perceptions of the tool; learning gains, where reported, came from separate tests.

Intervention designs

Five studies evaluated a training intervention quantitatively: three maritime and two transfer. Sample sizes in the maritime interventions ranged from 69 to 153. The gamification study divided 73 first year cadets alphabetically by surname into a control subgroup of 37 and an experimental subgroup of 36, and reported its outcome as two aggregate percentage improvements (Diahyleva et al., 2024). The largest intervention followed 153 cadets in matched groups of 76 and 77 over one semester and reported group comparisons without inferential tests (Diahyleva et al., 2025). The classroom action research with 69 students was the only maritime intervention to report inferential statistics, using paired t tests on pre and post scores (Ratnaningsih, 2025).

Both transfer interventions quantified their outcomes. The driver education experiment compared a ChatGPT trained group against conventional training and reported p values for three of the five examined functions (Murtaza et al., 2024). The construction safety study measured hazard recognition before and after the ChatGPT intervention, with recognition rising from below 35% to above 60% (Uddin et al., 2023). None of the three maritime interventions used randomised assignment; the allocation rules were alphabetical surname order (Diahyleva et al., 2024) and proficiency based matching (Diahyleva et al., 2025).

Observational, qualitative, and consensus designs

One maritime survey was identified. It compared the prior training of MSc maritime students and ship cadets in three digital skill areas, with 37%, 13%, and 16% of the MSc students reporting training in statistics, computer programming, and cybersecurity, against 6%, 9%, and 9% of the cadets (Karahalios, 2025). In the AIEd layer, a survey of economics and business management students found that only 17.6% had received any teaching in AI and that interest in AI exceeded knowledge of it (Almaraz-López et al., 2023).

The AIEd layer holds the mixed methods designs: a study of students' AI usage in higher education combined survey and interview data (Zhou et al., 2024), a second examined the effects of AI applications on creativity and academic emotions with teachers and students (Lin & Chen, 2024), and the acceptance study above added post reflection interviews with 30 participants to its survey phase (Dahri et al., 2024).

Two maritime studies used qualitative designs as their primary method. Structured interviews with 37 maritime stakeholders from six countries, with 9 to 47 years of experience, were analysed through constant comparative analysis to derive skill requirements for autonomous ship operators (Emad & Ghosh, 2023). The design ethnography of simulator instruction observed instructor feedback practices in situ (Sellberg & Sharma, 2024). The single consensus study applied a Delphi method combined with the analytic hierarchy process to 36 experts and concluded that the current STCW framework is not fully relevant for MASS (Kim & Mallam, 2020).

Computational experiments and evidence synthesis

Five maritime studies were computational experiments. Their designs comprised AutoML screening of 58 and of nine candidate models (Munim et al., 2025; Tusher et al., 2023), a benchmark of eight classifiers on 1,560 samples under 4-fold cross validation (Yang et al., 2021), distribution fitting with permutation of scenario elements (Hwang & Youn, 2022), and deep reinforcement learning over repeated runs (Sun et al., 2023).

Evaluation in all five was confined to model level metrics such as accuracy, RMSE, or task success rate; none linked model outputs to subsequent learning outcomes, and one explicitly reported that field validation of learning gains was infeasible because the trainees had dispersed worldwide after training (Yang et al., 2021).

Seven publications were classed as evidence synthesis. One review followed the four stage PRISMA process (Dewan et al., 2023), a second documented its search and screening in a flow diagram while acknowledging deviations from a standard systematic review (Karimi et al., 2024), two were bibliometric (Li et al., 2023; Munim & Haralambides, 2022), two reviewed regulatory and curricular development around STCW and MASS (Ahmad Fuad et al., 2025; Yi et al., 2025), and one transfer review screened 60 records on ChatGPT in healthcare education, with benefits cited in 85.0% of the records and concerns in 96.7% (Sallam, 2023). The bibliometric study of Maritime English research additionally reported a low proportion of empirical studies and a predominance of qualitative analysis in that literature (Li et al., 2023).

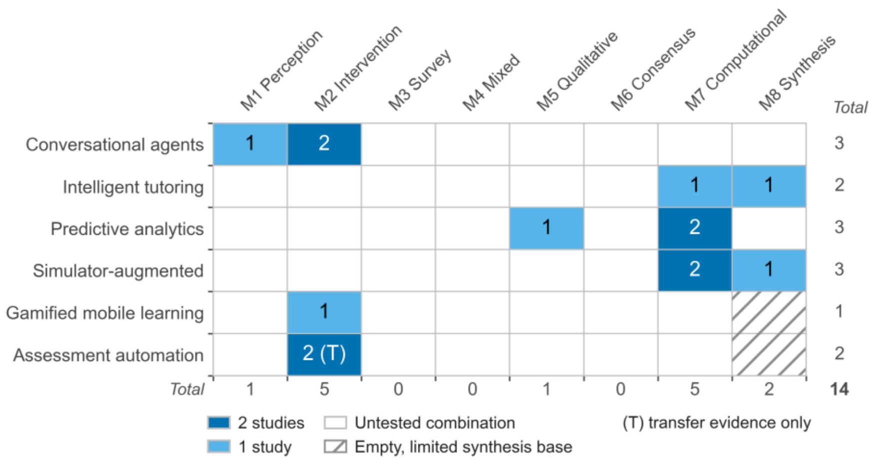
In the AIEd synthesis layer, a review of 146 articles on AI in higher education found that 73.3% applied quantitative methods, that only one study (0.7%) was qualitative, and that only two articles (1.4%) critically reflected on ethical implications (Zawacki-Richter et al., 2019). A later review of the AIEd field reported that 55.2% of articles did not prominently feature any theory and that mixed research methods remained underutilised (Wang et al., 2024).

Coverage matrix

Figure 2 cross-tabulates the six application paradigms against the eight methodology families. The matrix contains the 14 application studies that carry both a paradigm and a methodology assignment: 12 maritime studies and the two transfer interventions. The other 11 maritime and transfer publications remain outside it:

the two chatbot papers published without any evaluation, the eight maritime publications that did not implement an application, and the transfer review. The matrix column totals therefore differ from the family counts reported in the methodology inventory: the survey and consensus columns are empty because those studies examined competencies rather than an application.

Figure 2. Coverage matrix of the six application paradigms against the eight methodology families for the 14 evaluated maritime and transfer application studies; cell values give the number of studies.



Ten of the 48 cells are populated, holding the 14 studies between them. Four cells contain two or more studies: conversational agents evaluated through interventions (two studies), predictive analytics through computational experiments (two), simulator augmented training through computational experiments (two), and assessment automation through transfer interventions (two). The remaining six populated cells hold one study each: the chatbot usability evaluation, the learning analytics ethnography, the simulator review, the gamification intervention, and the two intelligent tutoring entries. Row totals of 3, 2, 3, 3, 1, and 2 match the assignments reported in the application typology for the studies that carry a methodology family. Two families dominate the matrix: intervention designs and computational experiments together account for 10 of the 14 studies. The conversational agent row, the most heavily studied paradigm, is populated in only two of its eight cells, the usability instrument and the intervention columns; no survey of chatbot users, no mixed methods study, and no qualitative study of chatbot supported learning was identified in the maritime set. For comparison, the AIEd chatbot review drew on 36 studies spanning a wider range of designs (Kuhail et al., 2022).

The empty cells form consistent row level patterns rather than scattered absences. The simulator augmented row contains no entry in any of the six learner facing families: the included simulator studies either generated training content computationally or synthesised prior work, and none evaluated learners directly. The intelligent tutoring row shows the same profile, with entries only in the computational and synthesis columns; no ITS in the maritime set was evaluated with students in an instructional setting. The assessment automation row contains no maritime entry at all, and the gamified mobile learning row holds a single cell. At column level, three families are entirely absent from the matrix: cross sectional surveys, mixed methods, and expert consensus have not been applied to any maritime or transfer application study in the included set.

Not every empty cell carries the same weight. Synthesis cells in rows with one or no primary maritime study, such as gamified mobile learning, are empty because there is little yet to synthesise, and the one consensus study addressed competency requirements rather than tool evaluation (Kim & Mallam, 2020). The remaining empty cells, in particular the learner facing columns of the simulator and intelligent tutoring rows and the entire maritime portion of the assessment automation row, mark combinations untested in the included studies. The reasons that may lie behind these patterns are taken up in the Discussion section.

Discussion

The typology and the methodology landscape together describe a young evidence base. Thirteen of the 14 application studies appeared in 2022 or later, two families hold 10 of the 14 matrix entries, no maritime or transfer study used a mixed methods design, and two application papers were published without any evaluation. The AIEd comparator literature recorded a similar profile at its own early stage: in a review of 146 articles, 73.3% applied quantitative methods, one study was qualitative, and only two articles reflected on ethical implications (Zawacki-Richter et al., 2019), while a later field review found that 55.2% of articles did not feature any theory and that mixed methods remained underutilised (Wang et al., 2024). The resemblance may indicate that GenAI research in MET is recapitulating the early AIEd pattern, in which adoption and description run ahead of evaluation practice. A further point sharpens this reading: the maritime domain is absent from the AIEd synthesis layer itself. Neither the 146 article review nor the later field review mentions maritime education at any point (Wang et al., 2024; Zawacki-Richter et al., 2019), and the chatbot review contains no maritime application among its 36 studies (Kuhail et al., 2022). MET researchers therefore appear to be working without a domain adapted evidence base on either side: the maritime literature is thin, and the general literature does not see the domain.

The matrix points to a mismatch between where applications are studied and how. The most heavily studied paradigm, conversational agents, was evaluated almost entirely through perception and short interventions, and the simulator and intelligent tutoring rows contain no learner facing evaluation at all (see the coverage matrix). Perception based evidence is fragile in this setting: student interest in chatbots tends to decline over time (Kuhail et al., 2022), ChatGPT is a poor judge of its own correctness (Lo, 2023), and single session usability ratings may not translate into durable learning value. The computational strand stops at a parallel ceiling: all five studies evaluated models against accuracy type metrics, and none connected model outputs to subsequent learning, a step one of them reported to be infeasible in practice (Yang et al., 2021). One likely reason is cost. Collecting human performance data in simulator settings was identified as a persistent constraint (Karimi et al., 2024), and instrumenting feedback in a simulator environment proved difficult even as a design exercise, with instructor judgement exceeding what the proposed analytics could capture (Sellberg & Sharma, 2024). The picture is not uniformly weak. The intervention cluster produced quantified gains, including significant pre to post improvements in Maritime English (Ratnaningsih, 2025) and in construction safety (Uddin et al., 2023) and significant advantages over conventional training in driver education (Murtaza et al., 2024), and the simulator log study validated an objective assessment approach (Munim et al., 2025). These results suggest that stronger designs are feasible within ordinary course structures; they have simply been rare.

The empty cells cluster where the competency literature places its demands. The eight maritime publications without an application argue, from consensus, interview, and regulatory analysis, that MASS era operations will require redefined competencies and reorganised training (Emad & Ghosh, 2023; Kim & Mallam, 2020; Munim & Haralambides, 2022; Wang et al., 2020; Yi et al., 2025). Yet the consensus study proposes no training solution for the competencies it ranks (Kim & Mallam, 2020), the interview study concludes that MET institutes are not yet ready to deliver the required skills (Emad & Ghosh, 2023), and the matrix shows no learner level evidence for exactly the paradigms, simulator augmented and assessment oriented, that remote operation would lean on. The field is specifying what future seafarers must learn faster than it is testing how GenAI could help them learn it. If this reading is correct, the most consequential research investments may lie in the simulator and assessment rows of the matrix rather than in further chatbot demonstrations.

A second cross cutting pattern concerns oversight. The classroom studies that worked with ChatGPT positioned the teacher as an unremovable component, whether as verifier of technically specific content (Diahyleva et al., 2025) or as facilitator supervising student interaction with a system that the authors caution is not an authority on any topic (Pindosova & Fedorova, 2026). The wider literature gives this stance an evidential basis: ChatGPT was observed to fabricate confident statements and to vary its answers across identical queries (Liu et al., 2023), and it produced a fully referenced but nonexistent article in one documented case (Lo, 2023). The healthcare review reached the same position in another high stakes training domain, recommending extreme caution and human oversight (Sallam, 2023), and the education position literature ties responsible adoption to continuous human monitoring (Kasneci et al., 2023). In MET the stakes are operational: errors absorbed during training can surface later as competence deficits at sea. The included evidence is therefore most consistent with a teacher mediated model of GenAI use, in which the tool extends rather than replaces instruction (Grassini, 2023).

Four limitations qualify these interpretations. First, the search was iterative rather than registered, and three records were excluded on availability grounds, so relevant studies may be missing (see the Methods section). Second, coding was performed by a single analyst and the primary assignment rule flattens multi method studies; the largest maritime intervention contained survey and interview elements that the matrix does not show (Diahyleva et al., 2025). Third, the included set is closed while the field moves quickly: much of the early ChatGPT literature rested on preprints of uncertain quality, as one included review itself reported (Lo, 2023), and the AIEd comparator layer largely predates GenAI, a boundary its authors acknowledge (Wang et al., 2024). Fourth, with 14 studies across 48 cells, the matrix patterns are descriptive; they locate absences but cannot estimate effects. These limitations soften the strength of the claims made here, not their direction: the absences in the matrix would remain visible under any reasonable recoding, and the research directions they imply are taken up in the Conclusion.

Conclusion

Across 44 publications spanning 2017 to 2026, six paradigms organise the maritime evidence: conversational agents, intelligent tutoring and student modeling, predictive analytics, simulator augmented training, gamified mobile learning, and assessment automation, the last supported by transfer studies alone. Of the 16 application studies, two were published without any evaluation; intervention designs and computational experiments carry 10 of the 14 evaluated studies, and no maritime or transfer study used a mixed methods design. Cross-tabulating the axes leaves 10 of 48 combinations populated, and the empty cells form patterns rather than scattered absences: simulator and intelligent tutoring applications were never evaluated directly with learners in the included studies, and no maritime study has examined assessment automation.

The agenda that follows from these patterns, read within the limits set out in the Discussion, is concrete. Controlled chatbot interventions that pair the existing pre and post designs with inferential testing and matched outcome measures would consolidate the best populated row of the matrix. Simulator and tutoring research needs learner facing studies that connect model outputs to instructional events and to learning outcomes, the link the computational work itself reports as missing. A maritime pilot of assessment automation would test whether the positive transfer evidence from construction safety and driver education survives contact with maritime content such as COLREGs scenarios and Maritime English communication. The two untouched methodology columns also have work to do: consensus methods could establish which GenAI competencies instructors and cadets will need as autonomous shipping matures, and mixed methods designs could capture the classroom dynamics that test scores alone miss. The included evidence consistently favours arrangements that keep instructors in the loop: in a safety critical training domain, GenAI is better treated as a supervised instrument than as an autonomous teacher.

References

Ahmad Fuad, A. F., Choong, D. P. A., Suhrab, M. I. R., & Zainol, I. (2025). Maritime education in the era of Society 5.0 and autonomous ships: a concise literature review. *Australian Journal of Maritime & Ocean Affairs, 17*(2), 251-260. https://doi.org/10.1080/18366503.2024.2336740

Almaraz-López, C., Almaraz-Menéndez, F., & López-Esteban, C. (2023). Comparative Study of the Attitudes and Perceptions of University Students in Business Administration and Management and in Education toward Artificial Intelligence. *Education Sciences, 13*(6), 609. https://doi.org/10.3390/educsci13060609

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access, 8*, 75264-75278. https://doi.org/10.1109/ACCESS.2020.2988510

Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., Shutaleva, A., & Soomro, R. B. (2024). Extended TAM based acceptance of AI-Powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study. *Heliyon, 10*(8), e29317. https://doi.org/10.1016/j.heliyon.2024.e29317

Dewan, M. H., Godina, R., Chowdhury, M. R. K., Noor, C. W. M., Wan Nik, W. M. N., & Man, M. (2023). Immersive and Non-Immersive Simulators for the Education and Training in Maritime Domain—A Review. *Journal of Marine Science and Engineering, 11*(1), 147. https://doi.org/10.3390/jmse11010147

Diahyleva, O., Yurzhenko, A., & Kononova, O. (2024). Gamification techniques in Maritime English online courses: Motivating learners in virtual environments. *Advances in Mobile*

Learning Educational Research, 4(1), 965-972. https://doi.org/10.25082/AMLER.2024.01.008

Diahyleva, O., Kononova, O., & Yurzhenko, A. (2025). Maritime English Instruction for Future Ship Engineers: A Pedagogical Study Using ChatGPT and Digital Learning Platforms. *International Journal of Multilingualism and Languages for Specific purposes, 7*(02), 20-32. https://doi.org/10.52919/ijmlsp.v7i02.119

Emad, G. R., & Ghosh, S. (2023). Identifying essential skills and competencies towards building a training framework for future operators of autonomous ships: a qualitative study. *WMU Journal of Maritime Affairs, 22*(4), 427-445. https://doi.org/10.1007/s13437-023-00310-9

Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2025). The promise and challenges of generative AI in education. *Behaviour & Information Technology, 44*(11), 2518-2544. https://doi.org/10.1080/0144929X.2024.2394886

Grassini, S. (2023). Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. *Education Sciences, 13*(7), 692. https://doi.org/10.3390/educsci13070692

Hwang, T., & Youn, I. H. (2022). Navigation Scenario Permutation Model for Training of Maritime Autonomous Surface Ship Remote Operators. *Applied Sciences, 12*(3), 1651. https://doi.org/10.3390/app12031651

Karahalios, H. (2025). Enhancing maritime education for digital sustainability. *International Journal of Innovation Science*. https://doi.org/10.1108/IJIS-06-2024-0156

Karimi, E., Smith, J., Billard, R., & Veitch, B. (2024). AI-based adaptive instructional systems for maritime safety training: a systematic literature review. *Discover Artificial Intelligence, 4*(1). https://doi.org/10.1007/s44163-024-00153-0

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Fedor, K., Strinyuk, S., & Lanin, V. (2022). Developing Voice Chatbot for Learning Maritime English. *2022 IEEE 16th International Conference on Application of Information and Communication Technologies (AICT)*, 1-5. https://doi.org/10.1109/AICT55583.2022.10013615

Kim, T. E., & Mallam, S. (2020). A Delphi-AHP study on STCW leadership competence in the age of autonomous maritime operations. *WMU Journal of Maritime Affairs, 19*(2), 163-181. https://doi.org/10.1007/s13437-020-00203-1

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2022). Interacting with educational chatbots: A systematic review. *Education and Information Technologies, 28*(1), 973-1018. https://doi.org/10.1007/s10639-022-11177-3

Li, T., Chen, Q., Xi, Y., & Lau, Y. Y. (2023). A 40-Year Bibliometric Analysis of Maritime English Research: Insights and Implications. *Sustainability, 15*(5), 4348. https://doi.org/10.3390/su15054348

Lin, H., & Chen, Q. (2024). Artificial intelligence (AI)-integrated educational applications and college students' creativity and academic emotions: students and teachers' perceptions and attitudes. *BMC Psychology, 12*(1). https://doi.org/10.1186/s40359-024-01979-0

Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., & Ge, B. (2023). Summary of ChatGPT-Related Research and Perspective Towards the Future of Large Language Models. *Meta-Radiology, 1*(2), 100017. https://doi.org/10.1016/j.metrad.2023.100017

Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

Munim, Z. H., Kjeldsberg, F., Bustgaard, M., Bhagat, S., Haavardtun, P., Kim, T. E., Lindroos, E., Thorvaldsen, H., Nyairo, F., & Lampiola, J. (2025). Predictive Performance Assessment in Simulation Training using Machine Learning. *International Journal of Artificial Intelligence in Education, 35*(4), 2154-2184. https://doi.org/10.1007/s40593-025-00464-y

Munim, Z. H., & Haralambides, H. (2022). Advances in maritime autonomous surface ships (MASS) in merchant shipping. *Maritime Economics & Logistics, 24*(2), 181-188. https://doi.org/10.1057/s41278-022-00232-y

Murtaza, M., Cheng, C. T., Fard, M., & Zeleznikow, J. (2024). Transforming Driver Education: A Comparative Analysis of LLM-Augmented Training and Conventional Instruction for Autonomous Vehicle Technologies. *International Journal of Artificial Intelligence in Education, 35*(2), 736-773. https://doi.org/10.1007/s40593-024-00407-z

Pindosova, T., & Fedorova, O. (2026). Use of ChatGPT in Maritime English Lessons. *Перспективи та інновації* *науки*(1(59)). https://doi.org/10.52058/2786-4952-2026-1(59)-153-165

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12*(1). https://doi.org/10.1186/s41039-017-0062-8

Ratnaningsih, D. (2025). Enhancing students' speaking skills using ChatGPT application in ESP maritime English learning. *JEES (Journal of English Educators Society), 10*(2), 186-197. https://doi.org/10.21070/jees.v10i2.1964

Sallam, M. (2023). ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare, 11*(6), 887. https://doi.org/10.3390/healthcare11060887

Sellberg, C., & Sharma, A. (2024). Toward multimodal learning analytics in simulation-based collaborative learning: A design ethnography of maritime training. *International Journal of Computer-Supported Collaborative Learning, 20*(2), 201-221. https://doi.org/10.1007/s11412-024-09435-2

Sharma, A., Undheim, P. E., & Nazir, S. (2022). Design and implementation of AI chatbot for COLREGs training. *WMU Journal of Maritime Affairs, 22*(1), 107-123. https://doi.org/10.1007/s13437-022-00284-0

Sun, Z., Fan, Y., & Wang, G. (2023). An Intelligent Algorithm for USVs Collision Avoidance Based on Deep Reinforcement Learning Approach with Navigation Characteristics. *Journal of Marine Science and Engineering, 11*(4), 812. https://doi.org/10.3390/jmse11040812

Mahbub Tusher, H., Haque Munim, Z., Hussain, S., & Nazir, S. (2023). An Automated Machine Learning Approach for Early

Identification of At-Risk Maritime Students. *AHFE International, 109*. https://doi.org/10.54941/ahfe1003150

Uddin, S. M. J., Albert, A., Ovid, A., & Alsharef, A. (2023). Leveraging ChatGPT to Aid Construction Hazard Recognition and Support Safety Education and Training. *Sustainability, 15*(9), 7121. https://doi.org/10.3390/su15097121

Deling, W., Dongkui, W., Changhai, H., & Changyue, W. (2020). Marine Autonomous Surface Ship - A Great Challenge to Maritime Education and Training. *American Journal of Water Science and Engineering, 6*(1), 10. https://doi.org/10.11648/j.ajwse.20200601.12

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications, 252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology, 55*(1), 90-112. https://doi.org/10.1111/bjet.13370

Yang, C., Chiang, F. K., Cheng, Q., & Ji, J. (2021). Machine Learning-Based Student Modeling Methodology for Intelligent Tutoring Systems. *Journal of Educational Computing Research, 59*(6), 1015-1035. https://doi.org/10.1177/0735633120986256

Yi, S. Y., Jung, M., & Lee, S. I. (2025). Adapting to change: International maritime education and training for future seafarers – focusing on the comprehensive review of the STCW convention and code of the 10th session of the IMO HTW. *Journal of International Maritime Safety, Environmental Affairs, and Shipping, 9*(2). https://doi.org/10.1080/25725084.2025.2464486

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators?. *International Journal of Educational Technology in Higher Education, 16*(1). https://doi.org/10.1186/s41239-019-0171-0

Zhou, X., Zhang, J., & Chan, C. (2024). Unveiling students' experiences and perceptions of Artificial Intelligence usage in higher education. *Journal of University Teaching and Learning Practice, 21*(06). https://doi.org/10.53761/xzjprb23

GENERATIVE AI AND THE TRANSFORMATION OF SEAFARER COMPETENCIES IN THE ERA OF MARITIME AUTONOMOUS SURFACE SHIPS

Hüseyin Fatih NACAR[1]

Introduction

In May 2026 the International Maritime Organization adopted the first International Code of Safety for Maritime Autonomous Surface Ships, a non-mandatory instrument that took effect on 1 July 2026, with a mandatory version targeted for entry into force in 2032 (IMO, 2026). The Code completes a regulatory arc that began when the IMO approved a four degree autonomy taxonomy, ranging from ships with automated decision support to vessels whose operating systems determine actions by themselves (IMO, 2021). Autonomy in commercial shipping is no longer a forecasting exercise; it is a regulated category with dates attached.

The people who will operate these ships are still trained and certified under a framework written for a different vessel. The competency regime of the STCW Convention was last revised comprehensively in Manila in 2010, and it requires watchkeeping

[1] Research Assistant, Istanbul Technical University, Department of Marine Engineering, ORCID: 0009-0001-7178-141X

officers to be physically present on the navigating bridge or in a directly associated location throughout their watch (IMO, 2010). Analyses in the maritime education literature anticipated the strain early: a curriculum gap study published when the autonomy taxonomy was new found artificial intelligence, remote control and fault diagnosis entirely absent from existing programmes, and cited industry forecasts of around 1,000 autonomous and 2,000 semi autonomous ships by 2025 (Wang et al., 2020). A research thread has since formed around what the coming period demands of seafarers. Expert panels have rescored the STCW leadership requirements for remote operations (Kim & Mallam, 2020), stakeholder interviews have mapped the skills expected of shore based operators (Emad & Ghosh, 2023), and the IMO itself has opened a comprehensive review of the Convention with adoption targeted for 2027 (Yi et al., 2025).

What this thread does not ask is how the new competencies will be taught. None of the competency studies considers generative AI, or any AI based tool, as a training delivery mechanism. The application literature runs in the opposite direction: the studies that build and test AI supported training in maritime education, surveyed in the preceding chapter, train conventional competencies such as Maritime English and collision regulations, and only one engages the competency demands that autonomy creates. Two bodies of work that concern the same students at the same historical moment have developed without touching.

This chapter connects them. Taking the 2010 Manila framework as the baseline, it asks two questions. First, which STCW defined competencies remain stable, which are transformed, and which new demands emerge across the IMO degrees of autonomy? Second, how can the application paradigms of generative AI in maritime education be mapped onto those emerging demands, and which mappings are evidence backed, speculative, or unaddressed?

The next section describes the approach. The section that follows answers the first question and derives six emerging competency categories (Figure 1). The mapping section crosses them with six application paradigms (Table 1), the discussion takes up the resulting asymmetry, and the conclusion sets out a research agenda.

Approach

This chapter draws on the 44 studies identified in the scoping review reported in the preceding chapter of this volume, supplemented by three regulatory documents of the International Maritime Organization (IMO): the 2010 Manila Amendments to the STCW Convention (IMO, 2010), the outcome of the regulatory scoping exercise for Maritime Autonomous Surface Ships (MASS) (IMO, 2021), and the IMO announcement of the non-mandatory MASS Code adopted at the 111th session of the Maritime Safety Committee in May 2026 (IMO, 2026). Eleven of the included studies examine seafarer competencies under autonomous or remote operations, and these form the evidence base for the competency transformation section. The six application paradigms identified in the preceding chapter (conversational agents, intelligent tutoring and student modeling, predictive analytics, simulator augmented and immersive training, gamified mobile learning, and assessment and feedback automation) supply the second axis of the analysis. Findings were synthesised narratively; no new database search was conducted, and no formal systematic review protocol was applied.

Two rules governed the mapping that follows. First, an emerging competency category was retained only when at least two independent sources supported it, a threshold intended to filter out single-study terminology. Second, each cell of the matrix in Table 1 was classified into one of three states: evidenced, meaning that at least one application study was conducted in a maritime education setting; speculative, meaning that the available evidence comes from another domain, from general educational research, or from maritime studies of the same paradigm that target conventional competencies; and unaddressed, meaning that neither body of literature provides evidence. Speculative cells are reported as transfer possibilities rather than recommendations.

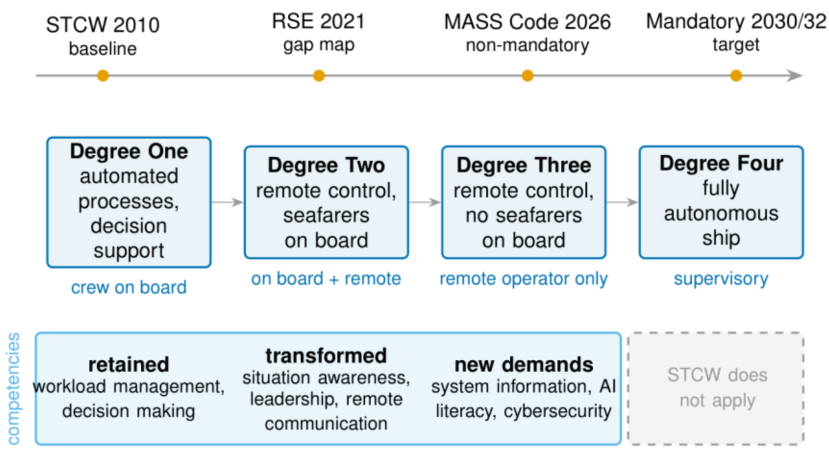
Generative AI is the focal technology of this chapter, but the paradigm set deliberately spans the wider family of AI based training tools, including tutoring and predictive systems that predate large language models, consistent with the preceding chapter. Emerging competency demands do not distinguish between generations of the underlying technology, and excluding earlier systems would understate the available evidence.

Competency transformation in the MASS era

The 2010 Manila Amendments replaced the entire STCW Annex with text that entered into force on 1 January 2012 and required full compliance by 1 January 2017 (IMO, 2010). The framework certifies people for shipboard roles: every officer in charge of a navigational watch on a ship of 500 gross tonnage or more holds a certificate of competency, the detailed knowledge, understanding and proficiency for each role resides in the tables of STCW Code Part A, and continued competence is demonstrated at intervals not exceeding five years under Regulation I/11 (IMO, 2010). Two provisions of the same revision matter for what follows. Regulation I/12 grants formal standing to simulator based training and assessment, and Regulation III/6 created the electro technical officer, the closest precedent for a digitally oriented certification category absorbed into the Convention (IMO, 2010).

The Convention also fixes where competence is exercised. Officers of the navigational watch "shall be physically present on the navigating bridge or in a directly associated location" at all times during their watch (IMO, 2010). The autonomy taxonomy set out in the IMO regulatory scoping exercise strains this assumption step by step. Degree Two and Degree Three both describe a ship "controlled and operated from another location", with seafarers available on board in the former and absent in the latter, while in Degree Four the operating system of the ship makes decisions and determines actions by itself (IMO, 2021). The degrees are not a hierarchy, and a single ship may operate at more than one degree during one voyage (IMO, 2021). For Degree Four the exercise recorded that no trained and qualified seafarers serve on board at all, which leaves the Convention without a subject to certify (Figure 1).

Figure 1. Regulatory timeline, IMO autonomy degrees, and the corresponding transformation of seafarer competencies.



Source: Author's own elaboration based on IMO (2010, 2021, 2026).

The scoping exercise also produced the official map of what is missing. Across the safety instruments it identified eleven common potential gaps and themes, three of which were designated high priority: the meaning of the terms master, crew or responsible person; the functional requirements of the remote control station or centre; and the possible designation of the remote operator as a seafarer (IMO, 2021). For STCW itself, the assessment concluded that Degrees Two and Three require equivalences, amendments or a new instrument, and the gap list for SOLAS chapter V under Degree Three names training and drilling explicitly (IMO, 2021). The analysis of the Casualty Investigation Code went one step further and stated that the definition of a seafarer needs to be amended to include personnel engaged in remote operation of the vessel (IMO, 2021). No competency content, training hours or assessment criteria accompany any of these findings.

Against this background, the empirical literature first records what survives. In a Delphi study with 36 maritime experts, knowledge of task and workload management kept 100 per cent agreement under both the manned and the remotely controlled scenario, decision making techniques kept 100 and 92 per cent, and the application of standard operating procedures remained above the consensus threshold in both cases (Kim & Mallam, 2020). A separate interview study with 37 stakeholders from six countries reached the same conclusion: every respondent group regarded a traditional deck or engineering licence as the prerequisite for any shore based operator role, and participants proposed a "smart ship operator's licence" layered above the conventional one rather than replacing it (Emad & Ghosh, 2023).

What rises is more specific. For remote operators of unmanned ships, situation awareness ranked first among leadership requirements with an importance weight of 30.9 per cent, and second place went to a competence that exists nowhere in the current STCW tables, the knowledge and ability to acquire, handle and comprehend large amounts of system information, at 23.9 per cent (Kim & Mallam, 2020). The same new competence ranked second even for officers serving on board highly automated ships. Agreement on shipboard personnel management fell from 92 per cent under the manned scenario to 63 per cent under the remote one, below the 80 per cent consensus threshold, and the experts described the leadership style required in its place with the vocabulary of horizontal management, remote collaboration and delegation (Kim & Mallam, 2020). The stakeholder interviews condensed the demand side into five clusters, cognitive, operational, leadership and teamwork, decision making, and communicative skills, and added explicitly technical expectations including artificial intelligence, satellite communication and integrated systems linking shore and ship (Emad & Ghosh, 2023).

Curricula have not followed. A gap analysis at Shanghai Maritime University compared the knowledge required to operate MASS with existing maritime curricula and coded artificial intelligence, autonomous navigation, fault diagnosis, remote control, data mining, environmental information perception and the internet of things as absent from current programmes (Wang et al., 2020). The IMO response is in motion but slow: the tenth session of the Human Element, Training and Watchkeeping Sub-Committee in February 2024 agreed 22 key review items for the comprehensive STCW revision, including cybersecurity awareness and digitalisation of certification, with adoption targeted for 2027, and recommended that Member States develop separate competency frameworks for future MASS operators covering remote situation awareness and IT proficiency (Yi et al., 2025). Survey data show how far practice lags: in one institutional study, 16 per cent of MSc students and 9 per cent of ship cadets had received any cybersecurity training, and fewer than 29 per cent applied more than two protective practices (Karahalios, 2025). Parallel curriculum work recommends revising IMO Model Course 7.01 to add cybersecurity, the internet of things, cloud operations and drone technology, and frames adaptive "learning to learn" capacity as a core requirement of the coming period (Ahmad Fuad et al., 2025).

Six emerging competency categories were derived from these sources, each supported by at least two of them: remote situation awareness and system supervision; system information comprehension; human-automation teaming and horizontal leadership; digital and AI literacy; cybersecurity awareness; and communication for remote and mixed operations. These six form the rows of Table 1. Competencies that the evidence marks as retained rather than transformed, such as workload management and decision making, were excluded from the matrix, since the question addressed in the next section concerns the new demands.

Mapping AI based training tools to the emerging competencies

Table 1 crosses the six emerging competency categories with the six application paradigms and classifies each of the 36 cells using the three states defined in the Approach section. The result is an uneven map. Three cells are evidenced by maritime application studies, four are speculative, and 29 are unaddressed. Two competency rows, digital and AI literacy and cybersecurity awareness, contain no entry of any kind.

Table 1. Mapping of the six emerging competency categories to the six AI application paradigms; ✓ marks cells evidenced by maritime studies (number of studies in parentheses), ~ marks cells with transfer or general educational evidence only, and empty cells are unaddressed.

Emerging competency	**Conversational agents**	**Intelligent tutoring**	**Predictive analytics**	**Simulator augmented**	**Gamified mobile**	**Assessment automation**
Remote situation awareness and system supervision			~	✓ (1)		
System information comprehension		~				
Human-automation teaming and horizontal leadership	~					
Digital and AI literacy						
Cybersecurity awareness						
Communication for remote and mixed operations	✓ (5)				✓ (1)	~

The pairing of communication with conversational agents carries the most detailed evidence in the map. The earliest system, a purpose built voice chatbot for Maritime English, organised instruction into three sub dialogues with a translator and voice cards and logged right and wrong answers (Khlebalov et al., 2022). FLOKI, a chatbot for collision regulations training, scored 73.72 on the System Usability Scale with 18 cadets, above the published benchmark median of 70.5 (Sharma et al., 2022). After the release of ChatGPT the emphasis moved into the classroom. One account presented five worked task templates for Maritime English lessons but reported no learner outcomes (Pindosova & Fedorova, 2026). A semester long quasi experiment followed 153 marine engineering cadets, 76 of them using a ChatGPT based environment for professional dialogues and quiz generation, and recorded greater progress in the experimental group without inferential testing (Diahyleva et al., 2025). A classroom action research cycle with 69 polytechnic students reported statistically significant gains in grammar, vocabulary, fluency and comprehension, lifting the overall mean from 74.80 to 87.00 (Ratnaningsih, 2025). The same competency row holds the single gamification study, in which a Maritime English course enriched with leaderboards, badges and missions improved attendance and subject knowledge by 43.7 per cent in the experimental group against 7.8 per cent in the control group (Diahyleva et al., 2024). One qualification applies to the entire row: all six studies trained conventional maritime content in conventional settings, and none addressed the ship to shore coordination that remote operations will demand.

The second evidenced cell pairs remote situation awareness with simulator augmented training, and it is the only cell in the map whose evidence was produced explicitly for autonomous shipping. Working from South Korean AIS records, Hwang and colleagues extracted the statistical structure of real voyages, in which course altering angles of 9 degrees accounted for 48.8 per cent of cases and waypoint distances of 2 nautical miles for 65.0 per cent, and permuted the representative values into 564,480 candidate navigation scenarios for training remote operators of autonomous ships (Hwang & Youn, 2022). The surrounding paradigm is mature on the hardware side but thin on comparative evidence: a decade spanning review of simulators in maritime education retrieved 27 articles and found that only nine compared simulator training against traditional instruction (Dewan et al., 2023). A reinforcement learning agent that encodes collision regulations in its reward signal marks the algorithmic edge of the same paradigm, although it trains a controller rather than a learner (Sun et al., 2023).

The two speculative cells in the tutoring and analytics columns share one feature: maritime evidence exists for the paradigm, but it targets conventional competencies. Fifty eight algorithms were trained on simulator logs to predict student performance in Williamson Turn manoeuvres, with an XGBoost classifier adopted and rudder angle and propeller revolutions identified as the dominant predictors (Munim et al., 2025). A design ethnography of simulator instruction showed that speech recognition can supply stable records of bridge team communication for automated feedback (Sellberg & Sharma, 2024). On the tutoring side, a Naive Bayes student model reached 95 per cent accuracy on a navigation tutoring dataset of 1,560 samples (Yang et al., 2021), and a systematic review of adaptive instructional systems in maritime education counted decision trees in ten articles, neural networks in nine and support vector machines in eight (Karimi et al., 2024). Every one of these systems models learners performing tasks that the STCW tables already describe. Whether the same log based assessment and student modeling machinery transfers to remote operation consoles, or to the system information comprehension competence that ranked second in the expert hierarchy reported above, has not been examined in any included study.

The remaining speculative cells rest on imported evidence. In driver education, a ChatGPT instance restricted to vehicle manuals brought trainees to accuracies of 80 to 100 per cent across five driving assistance functions, with significant differences from conventional training in the three functions tested statistically (Murtaza et al., 2024). The manual grounded design is the transferable element, since marine equipment manuals and safety management system documentation occupy the same position in shipboard work that vehicle manuals occupy in driving. In construction safety education, hazard recognition rose from below 35 per cent to above 60 per cent after a ChatGPT intervention, and over 95 per cent of the students endorsed the tool for safety training (Uddin et al., 2023); the study sits outside Table 1 because hazard recognition maps to no single competency row, but demonstrates the same scenario generation mechanism. For human-automation teaming, the only available anchor is a position paper arguing that large language models require teachers and learners to develop the competencies and literacies needed to understand the technology and its brittleness (Kasneci et al., 2023); routine interaction with such systems is a plausible but untested vehicle for practising collaboration with an automated agent. None of these three transfers has been replicated with seafarers, cadets or maritime instructors.

The two empty rows complete the map. For digital and AI literacy, the demand side is documented from three directions, the curriculum gap table that codes artificial intelligence and data mining as absent from maritime programmes (Wang et al., 2020), the model course revisions that would add AI and related technologies (Ahmad Fuad et al., 2025), and the position paper that frames AI literacy as a prerequisite competency for learners and teachers alike (Kasneci et al., 2023). The supply side is silent: no included study, maritime or otherwise, evaluates a generative AI tool that teaches students about AI systems. Cybersecurity awareness shows the same two sided void, listed among the 22 items of the STCW comprehensive review (Yi et al., 2025) and measured at training rates of 16 per cent or less in the student survey cited above (Karahalios, 2025), yet absent from every application study in the map.

Read along the expert hierarchy of the preceding section, the map inverts the priority order. The two competencies that ranked first and second for remote operators, situation awareness and system information comprehension, hold one evidenced cell and two speculative cells between them. The densest row, communication, draws six maritime studies, yet it entered the emerging set through the ship to shore coordination requirement rather than through the expert rankings. The interpretation of this asymmetry is taken up in the discussion.

Discussion

The central pattern of Table 1 is a mismatch between where the evidence sits and where the expert rankings point. Three of 36 cells are evidenced, and five of the seven studies behind them train conventional maritime communication. This distribution is consistent with the institutional unreadiness reported in the stakeholder interviews, in which educators attributed the standstill to the infancy of the technology itself (Emad & Ghosh, 2023). One plausible explanation is practical rather than pedagogical: a general purpose chatbot costs a language instructor an afternoon to adopt, whereas studying remote situation awareness requires a remote operation console that few institutions possess. Research appears to have followed tool availability rather than competency priority, and if that persists, the competencies most specific to autonomous operations may remain the least supported while demand for them grows.

The regulatory calendar suggests this is a closing window rather than an open ended question. The non-mandatory MASS Code took effect on 1 July 2026, with a mandatory version targeted for adoption by 2030 and entry into force in 2032 (IMO, 2026), and the comprehensive STCW review is aimed at adoption in 2027 (Yi et al., 2025). The current Convention already offers institutional entry points for this period: Regulation I/12 gives simulator based training formal standing, and the Regulation I/11 revalidation cycle carries new requirements to serving officers (IMO, 2010). Tools that mature before the mandatory Code arrives could enter maritime education through doors that already exist.

Any such development faces a constraint that the Convention itself articulates. Regulation VII/3 prohibits the use of alternative certification "to lower the integrity of the profession or 'de-skill' seafarers" (IMO, 2010), and the stakeholder consensus that shore based operators should hold a traditional licence first points in the same direction (Emad & Ghosh, 2023). Generative AI applications in this space will likely be judged by whether they extend seafarer competence toward the new demands or substitute for it; designs in which the tool performs the reasoning that the trainee is meant to acquire would conflict with both the regulation and the expert expectation.

Four limitations qualify these conclusions. First, the synthesis is bounded by the study set of the preceding chapter plus three regulatory documents; no new search was conducted, and tools reported after that search fall outside the map. Second, the competency evidence is largely two poled: the Delphi study covers only the first and third autonomy degrees, so the mapping cannot resolve differences between adjacent degrees (Kim & Mallam, 2020). Third, the speculative cells rest on transfer evidence from driver and construction safety education, and none of those interventions has been tested with maritime learners. Fourth, the full text of the MASS Code adopted in May 2026 had not been published at the time of writing, so its treatment here relies on the IMO summary announcement (IMO, 2026).

Conclusion

This chapter asked which seafarer competencies the transition to autonomous shipping reshapes, and how the application paradigms of generative AI in maritime education align with the new demands. On the first question, the regulatory baseline and the empirical literature converge: decision making and workload management survive the transition, shipboard personnel management loses relevance once crews leave the ship, and six categories rise that the current STCW tables do not cover in their present form, led by remote situation awareness and the comprehension of large volumes of system information. On the second question, the alignment is thin. Of the 36 pairings in Table 1, three carry maritime evidence, four rest on indirect evidence, and 29 are unaddressed; the two highest ranked competencies for remote operators hold one evidenced cell between them, while the best supported row, communication, belongs to the conventional curriculum.

Four lines of work follow directly from the empty cells. First, AI literacy needs its own tools, not just its place in revised model courses; no study yet evaluates a generative AI application that teaches students how such systems behave and fail. Second, cybersecurity awareness training is an open field for generative scenario production, against a documented baseline in which fewer than one in six maritime students has received any cybersecurity training. Third, the scenario generation approach already demonstrated for remote operators should be extended with generative debriefing and assessment, moving the best evidenced autonomy specific cell from scenario supply toward full training loops. Fourth, the transfer designs that worked in driver and construction safety education, above all the manual grounded chatbot, need replication with cadets and serving officers before the comprehensive STCW review and the mandatory MASS Code arrive. The window between 2026 and 2032 is when maritime education can prepare deliberately rather than react; the empty cells of Table 1 are a concrete way to spend it.

References

Ahmad Fuad, A. F., Choong, D. P. A., Suhrab, M. I. R., & Zainol, I. (2025). Maritime education in the era of Society 5.0 and autonomous ships: A concise literature review. *Australian Journal of Maritime & Ocean Affairs*. https://doi.org/10.1080/18366503.2024.2336740

Dewan, M. H., Godina, R., Chowdhury, M. R. K., Noor, C. W. M., Wan Nik, W. M. N., & Man, M. (2023). Immersive and non-immersive simulators for the education and training in maritime domain: A review. *Journal of Marine Science and Engineering*, *11*(1), 147. https://doi.org/10.3390/jmse11010147

Diahyleva, O., Kononova, O., & Yurzhenko, A. (2025). Maritime English instruction for future ship engineers: A pedagogical study using ChatGPT and digital learning platforms. *International Journal of Multilingualism and Languages for Specific Purposes*. https://doi.org/10.52919/ijmlsp.v7i02.119

Diahyleva, O., Yurzhenko, A., & Kononova, O. (2024). Gamification techniques in Maritime English online courses: Motivating learners in virtual environments. *Advances in Mobile Learning Educational Research*. https://doi.org/10.25082/AMLER.2024.01.008

Emad, G. R., & Ghosh, S. (2023). Identifying essential skills and competencies towards building a training framework for future operators of autonomous ships: A qualitative study. *WMU Journal of Maritime Affairs*. https://doi.org/10.1007/s13437-023-00310-9

Hwang, T., & Youn, I. H. (2022). Navigation scenario permutation model for training of maritime autonomous surface ship remote operators. *Applied Sciences*, *12*(3), 1651. https://doi.org/10.3390/app12031651

International Maritime Organization. (2010). *Adoption of amendments to the annex to the International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW), 1978* (Conference Resolution 1, STCW/CONF.2/DC/1). 2010 Manila Conference.

International Maritime Organization. (2021). *Outcome of the regulatory scoping exercise for the use of maritime autonomous surface ships (MASS)* (MSC.1/Circ.1638).

International Maritime Organization. (2026, May 22). *IMO adopts first global code for autonomous ships* [Press briefing]. https://www.imo.org/en/mediacentre/pressbriefings/pages/imo-adopts-mass-code.aspx

Karahalios, H. (2025). Enhancing maritime education for digital sustainability. *International Journal of Innovation Science*. https://doi.org/10.1108/IJIS-06-2024-0156

Karimi, E., Smith, J., Billard, R., & Veitch, B. (2024). AI-based adaptive instructional systems for maritime safety training: A systematic literature review. *Discover Artificial Intelligence*. https://doi.org/10.1007/s44163-024-00153-0

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Khlebalov, F., Strinyuk, S., & Lanin, V. (2022). Developing voice chatbot for learning Maritime English. *2022 IEEE 16th International Conference on Application of Information and*

Communication Technologies (AICT). https://doi.org/10.1109/AICT55583.2022.10013615

Kim, T. E., & Mallam, S. (2020). A Delphi-AHP study on STCW leadership competence in the age of autonomous maritime operations. *WMU Journal of Maritime Affairs*. https://doi.org/10.1007/s13437-020-00203-1

Munim, Z. H., Kjeldsberg, F., Bustgaard, M., Bhagat, S., Haavardtun, P., Kim, T. E., Lindroos, E., Thorvaldsen, H., Nyairo, F., & Lampiola, J. (2025). Predictive performance assessment in simulation training using machine learning. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-025-00464-y

Murtaza, M., Cheng, C. T., Fard, M., & Zeleznikow, J. (2024). Transforming driver education: A comparative analysis of LLM-augmented training and conventional instruction for autonomous vehicle technologies. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-024-00407-z

Pindosova, T., & Fedorova, O. (2026). Use of ChatGPT in Maritime English lessons. *Naukovi Innovatsii ta Peredovi Tekhnolohii*, *1*(59), 153-165. https://doi.org/10.52058/2786-4952-2026-1(59)-153-165

Ratnaningsih, D. (2025). Enhancing students' speaking skills using ChatGPT application in ESP maritime English learning. *JEES (Journal of English Educators Society)*, *10*(2). https://doi.org/10.21070/jees.v10i2.1964

Sellberg, C., & Sharma, A. (2024). Toward multimodal learning analytics in simulation-based collaborative learning: A design ethnography of maritime training. *International Journal of*

Computer-Supported Collaborative Learning. https://doi.org/10.1007/s11412-024-09435-2

Sharma, A., Undheim, P. E., & Nazir, S. (2022). Design and implementation of AI chatbot for COLREGs training. *WMU Journal of Maritime Affairs*. https://doi.org/10.1007/s13437-022-00284-0

Sun, Z., Fan, Y., & Wang, G. (2023). An intelligent algorithm for USVs collision avoidance based on deep reinforcement learning approach with navigation characteristics. *Journal of Marine Science and Engineering*, *11*(4), 812. https://doi.org/10.3390/jmse11040812

Uddin, S. M. J., Albert, A., Ovid, A., & Alsharef, A. (2023). Leveraging ChatGPT to aid construction hazard recognition and support safety education and training. *Sustainability*, *15*(9), 7121. https://doi.org/10.3390/su15097121

Wang, D., Wu, D., Huang, C., & Wu, C. (2020). Marine autonomous surface ship: A great challenge to maritime education and training. *American Journal of Water Science and Engineering*, *6*(1). https://doi.org/10.11648/j.ajwse.20200601.12

Yang, C., Chiang, F. K., Cheng, Q., & Ji, J. (2021). Machine learning-based student modeling methodology for intelligent tutoring systems. *Journal of Educational Computing Research*, *59*(6). https://doi.org/10.1177/0735633120986256

Yi, S., Jung, M., & Lee, S. (2025). Adapting to change: International maritime education and training for future seafarers. Focusing on the comprehensive review of the STCW convention and code of the 10th session of the IMO HTW. *Journal of International Maritime Safety, Environmental Affairs, and Shipping*. https://doi.org/10.1080/25725084.2025.2464486

BİDGE