Applying Large Language Models to Maritime Near-Miss Safety Data Analysis

Safe Maritime Transportation System

Overview

The goal of this project was to evaluate the effectiveness of artificial intelligence (AI)-enabled analysis within a secure environment to combine and interpret raw near-miss data submitted by SafeMTS (Safe Maritime Transportation System) participants. The team trained and applied large language models (LLMs), guided by maritime subject matter experts, to extract discrete data, such as causal factors, from over 19,000 records. The models also assessed critical risk areas like high-potential consequence events. Integrating these models into data workflows reduced manual review efforts, allowing experts to focus on the most important data while enhancing the efficiency and consistency of large-scale analysis. The project has improved SafeMTS' ability to identify hazards and analyze trends.

About SafeMTS

The SafeMTS program, sponsored by the U.S. Department of Transportation's Maritime Administration and operated by the Bureau of Transportation Statistics (BTS) in collaboration with the maritime industry, enables the commercial maritime industry to voluntarily and securely share near-miss safety information to identify early warnings, develop preventive measures, and reduce the risk of serious incidents. Participating companies share data confidentially with BTS, which serves as the program's data steward. Incident records often contain critical safety information in narrative and unstructured formats. BTS tested secure applications of LLMs to improve analysis speed and accuracy without increasing reporting burden. These technologies help amplify human expertise, shifting safety management toward more predictive, adaptive, and continuously learning systems.

Implications for Federal Statistics

This project demonstrates the ability of AI-enabled analysis to:

- Efficiently process large, diverse, and sensitive safety datasets within a secure environment
- Enhance the completeness of statistical variables
- Enable faster and more reliable dissemination of actionable safety insights
- Serve as a potential model for other precursor safety data programs across industries

Improvements to Data Quality With Application of LLMs

Dataset Completeness by Company (C)—Pilot Phase

Data Field	C1	C2	C3	C4	C5	C6	C7	C8	Total Values Submitted or Extracted	Percent of 19,161 Events
Vessel Type	11254		5	189	316	576	10	10	12360	64.5%
Near Miss Classification	10729	4491	40	455	316	51	10	10	16102	84.0%
Location on Vessel (High Level)	9106	1859	87	452	316	29	9	10	11868	61.9%
Location on Vessel (Detailed Lvl)		1412		0		53			1465	7.6%
Operations|Activity Ongoing		1		189	316	54	10	10	580	3.0%
Task Performed	8958	1494	39	448					10939	57.1%
System|Equipment Involved	6989		37	187	316	53	2	9	7593	39.6%
Observing Personnel Type			38	188	314	52	10	10	612	3.2%
Factor Preventing Further/Worse incident	244		38	189	316	52	10	10	859	4.5%
Potential Consequence			39	189	316	52	10	10	616	3.2%
Actual Consequence				2	34	12	3	1	52	0.3%
Potential Severity	2073			450					2523	13.2%
Actual Severity				452	316				768	4.0%
Causal | Contributing Factors		10		4	313	44		10	381	2.0%
Root Cause				0		28		8	36	0.2%

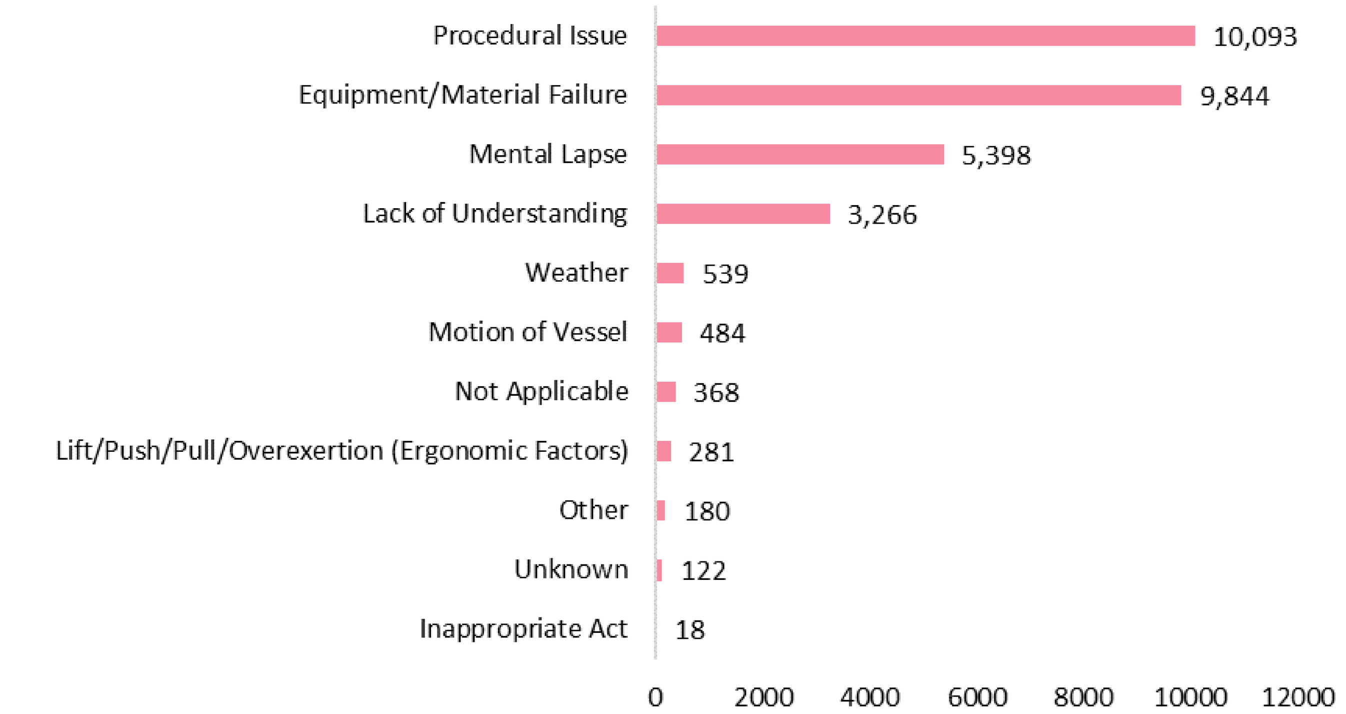
Dataset Completeness by Company (C)—Post LLM Processing

Data Field	C1	C2	C3	C4	C5	C6	C7	C8	Total Values Standardized	Percent Change
Vessel Type	12358		5	185	316	197	10	9	13080	3.8%
Near Miss Classification	8754	175	21	311	292	43	10	8	9614	-33.9%
Location on Vessel (High Level)	1102	703	11	139	36	37	4	1	2033	-51.3%
Location on Vessel (Detailed Lvl)	8089	1752	81	338	280	46	6	10	10602	47.7%
Operations|Activity Ongoing	7847	720	39	454	310	54	10	10	9444	46.3%
System|Equipment Involved	6305		37	187	284	50	10	10	6883	-3.7%
Observing Personnel Type			38	155	314	52	7	7	573	-0.2%
Factor Preventing Further/Worse incident	244		38	186	316	52	10	10	856	0.0%
Potential Consequence	2673		39	436	316	52	10	10	3536	15.2%
Actual Consequence				4	33	11	3	1	52	0.0%
Potential Severity				450					450	-10.8%
Actual Severity				452	316				768	0.0%
Causal | Contributing Factors	11331	4491	2026	455	315	522	9	10	19159	98.0%
Root Cause				0		3	1	7	11	-0.1%

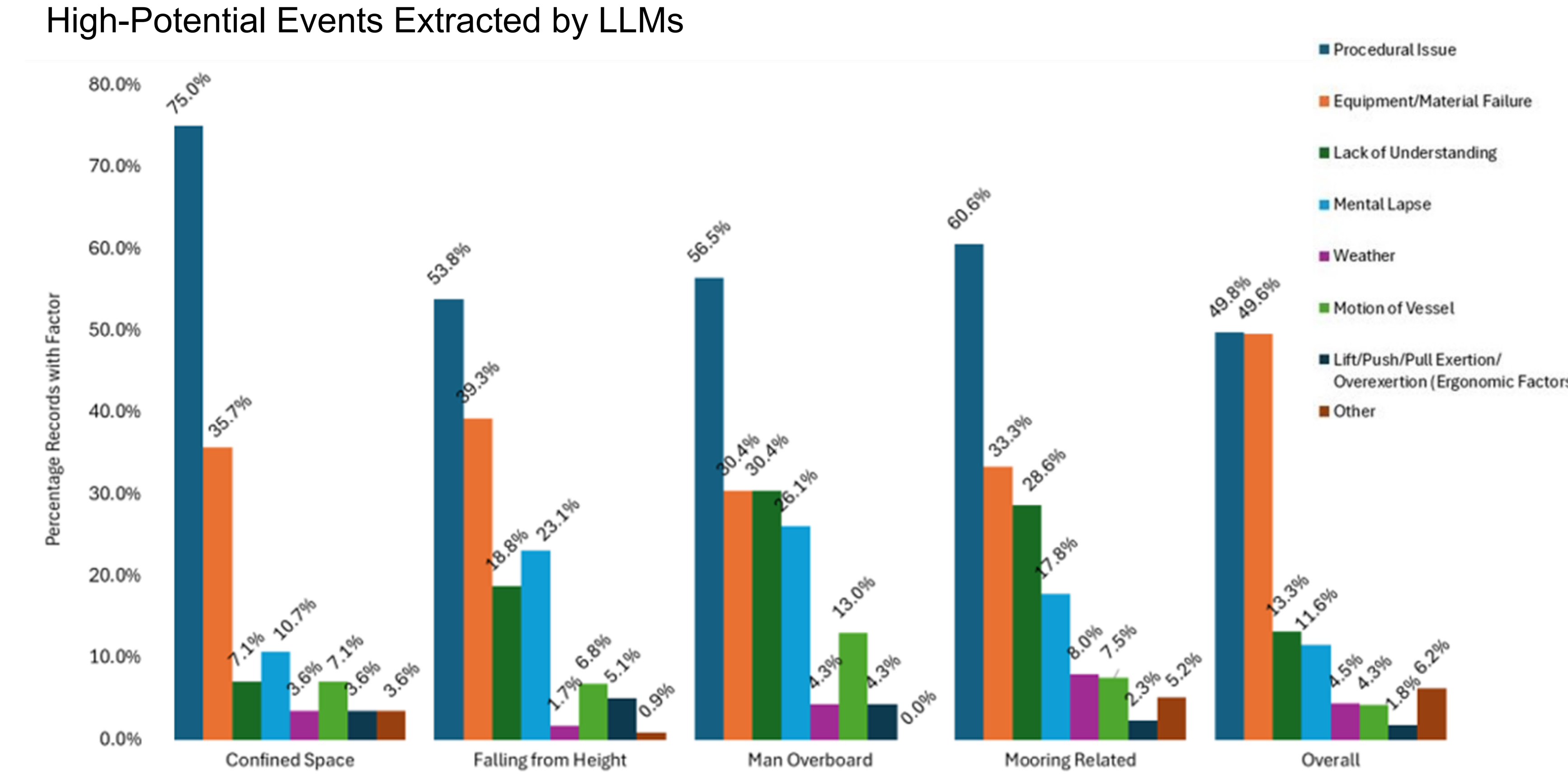
Sample Near-Miss Event Description	Extracted Causal Factors
"On [the] Articulated Tug Barge, the activation/reset switch for **the watertight door system on the bridge was accidentally bumped** into the closed position while a chair was being moved, causing the three watertight sliding doors in the engineering spaces to auto-close. The local alarm and 20-second delay notified personnel in that space of a closing. ***Two electrical extension cords being run through two of the watertight doors***	The watertight door system switch was accidentally bumped → **Carelessness**
"***Persons on the bridge were unaware until notified by personnel in affected spaces due to there being no alarm on the bridge panel.*** The only indication was the small light at each door symbol on the panel. Upon realization of the closure, personnel were notified of the unintended closing of the watertight doors. The activation switch was reset on the bridge and the watertight doors were electrically opened at the local operation switch."	Bridge personnel were unaware because no alarm existed on the bridge panel → **Poor/Insufficient Communications and Poor/Bad Design**
"The previous day, welding leads had been run through the doors and could have caused a worse incident. A temporary cover was immediately placed over the spring-loaded activation switch as an added security measure to prevent accidental operation by being bumped or brushed up against. A more permanent cover will be fabricated. I recommend we add a new procedure to lock-out the watertight doors from being operated remotely whenever ***cables are routed temporarily through these doors.*** All cables should always be removed whenever work is not being carried out."	Temporary routing of cables through watertight doors without safeguards → **Poor Planning and Failure to Follow Procedure**

Surfacing Hidden Risks

High-Level Causal Factors Extracted and Standardized by LLMs



High-Potential Events Extracted by LLMs



Key Findings

Causal factor extraction **increased from less than 2% coverage to nearly 100%**, with over **85% accuracy**. LLM-assisted processing:

- Significantly reduced manual review time
- Enabled actionable safety analysis on datasets that were previously incomplete
- Effectively identified high-potential safety events and emerging industry-wide trends

Next Steps

- Develop a proof-of-concept dashboard with publicly available trends and a secure participant benchmarking portal.
- Extend AI-enabled extraction to additional safety data fields
- Recruit more industry participants to strengthen dataset representativeness
- Iteratively refine prompts and models to handle new data types and improve accuracy

U.S. Department of Transportation
Office of the Secretary of Transportation
Bureau of Transportation Statistics

Project Manager: Amanda Lemons
Contributors: Michael Bohlman, Honglei Dai, Curtis Doucette, Allison Fischman, Bahadir Inozu, Benjamin Irwin, Will Nabach, Peter Schaedel, Brian Sumner

Acknowledgements: Special thanks to SafeMTS member companies within the maritime industry for sharing safety data used in this project and providing periodic feedback on results. BTS would also like to thank Todd Ripley, Kevin Kohlmann, and Thane Gilman for sharing their valuable input and perspective.

Learn More: https://www.c3rs.bts.gov/safemts-home/
Questions? SafeMTS@dot.gov