

# Demon Hunt - The Role of Endsley's Demons of Situation Awareness in Maritime Accidents

Tim Claudius Stratmann<sup>(✉)</sup> and Susanne Boll

Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany  
{tim.claudius.stratmann,susanne.boll}@uni-oldenburg.de

**Abstract.** Human Error is the cause of most maritime accidents. In a majority of the cases the source of the Human Error is a lack of Situation Awareness. Endsley et al. have identified eight causes that corrupt the Situation Awareness of human operators, the so-called Demons of Situation Awareness (SA Demons). We analyzed over five-hundred maritime accident reports for each of the eight SA Demons to provide a ranking of the causes and to identify the most prominent ones. Addressing these SA Demons enables maritime system designers to enhance the Situation Awareness of maritime operators and thereby improves the safety at sea.

**Keywords:** Situation awareness · Accident analysis · Maritime · Human Error

## 1 Introduction

Human Error accounts for 80–85% of maritime accidents [2]. The type of Human Errors are diverse, however they can be classified. Prior work of Grech et al. showed that a lack of Situation Awareness causes about 71% of all Human Error related accidents [9]. Similar findings have been made in other transportation domains. The problem of a lack in Situation Awareness was extensively investigated in the aviation domain.

Mica R. Endsley has defined Situation Awareness as “the *perception* (level 1) of the elements of the environment within a volume of time and space, the *comprehension* (level 2) of their meaning, and the *projection* (level 3) of their status in the near future” [6]. In 1999 Endsley introduced a taxonomy of Situation Awareness errors [7]. This taxonomy later led to the definition of the so-called Demons of Situation Awareness (SA Demons) [8]. The SA Demons stand for eight common causes for a lack of Situation Awareness. They address all three levels of Situation Awareness.

The aim of our work is to classify maritime accidents by Endsley's SA Demons. Moreover we provide our corpus of 1376 maritime accident reports and perform our analysis on a subset of it. We investigate the occurrences and

distribution of the eight SA Demons in a corpus of 535 maritime accident reports. Information about occurrences and distribution of the SA Demons enables maritime system designers to adjust their systems to the SA related needs of the operator and to build assistance systems that focus on mitigating the specific cause of SA errors.

## 2 Related Work

Most scientific accident analyses in the maritime domain focus on the statistical classification of error causes, whereas the investigations of safety authorities focus on deriving guidelines to prevent the same accidents from happening again. Human Error as cause of maritime accidents is no new phenomenon. As long ago as 1987 Wagenaar et al. analyzed 100 maritime accidents, of which 96 were caused by Human Error [13].

In 2005 Baker et al. published their three year enduring analysis of maritime accidents from the United States, Australia, Canada, Norway and the United Kingdom. Their results show that the frequency of accidents is declining, but that Human Error continues to be the dominant factor in approximately 80 to 85 % of maritime accidents [2,3]. According to their findings failures in Situation Awareness are a causal factor in the majority of accidents attributed to Human Error. They identified some significant factors associated with Situation Awareness failures which include: Cognitive and decision errors, Knowledge-Skill-Ability errors, task omissions and risk taking. They expect most of them to be artifacts of fatigue.

Endsley et al. determined eight types of causes for failures in Situation Awareness, the so-called Demons of Situation Awareness (SA Demons) [8]. The following list<sup>1</sup> describes each SA Demon and states the Situation Awareness levels it affects.

### SAD1 **Attention Tunneling** (*SA level 1*)

Good Situation Awareness is dependent on switching attention among multiple data streams. Locking in on certain data sources and excluding others is attention tunneling.

### SAD2 **Requisite Memory Trap** (*SA level 2*)

The working memory processes and holds chunks of data to support Situation Awareness level 2. The working memory is a limited resource. Systems that rely on robust memory do not support the user.

### SAD3 **Workload, Anxiety, Fatigue, and other Stressors** (*SA level 1 and 2*)

Stress and anxiety are likely issues in the warning environment. *WAFOS* taxes attention and working memory.

### SAD4 **Data Overload** (*SA level 1*)

There is more data available than can be processed by the human “bandwidth”.

<sup>1</sup> Based on: [http://www.au.af.mil/au/awc/awcgate/noaa/anti\\_situation\\_awareness.pdf](http://www.au.af.mil/au/awc/awcgate/noaa/anti_situation_awareness.pdf).

**SAD5 Misplaced Salience** (*SA level 1*)

Salience is the “compellingness” of a piece of data, often dependent on how it is presented.

**SAD6 Complexity Creep** (*SA level 1, 2 and 3*)

Complexity slows down the perception of information and it undermines the understanding and the projection of information.

**SAD7 Errant Mental Models** (*SA level 2 and 3*)

Wrong mental model may result in incorrect interpretation of data.

**SAD8 Out-of-the-loop syndrome** (*SA level 1*)

For example: Automated systems that do not involve the human until there is a problem.

Antão et al. used BNN models to analyse maritime accidents [1]. Other work dealing with human error in maritime accidents is [4, 5, 10, 12].

### 3 Corpus

Performing an analysis on specific SA error causes requires a data source with a high level of detail. Therefore we retrieved full-text reports of maritime accident investigations. The full corpus consists of 1376 maritime accident reports from five transportation safety authorities between the years 1987 and 2015. The reports were gathered from the British Marine Accident Investigation Branch (MAIB)<sup>2</sup>, the American National Transportation safety Board (NTSB)<sup>3</sup>, the United States Coast Guard (USCG)<sup>4</sup>, the Australian Transportation safety Board (ATSB)<sup>5</sup> and the Transportation safety Board of Canada (TSBC)<sup>6</sup> (Table 1).

**Table 1.** This table gives an overview of the retrieved corpus. The table shows the total number of available full-text reports, the time period covered by the reports and the country of origin. (*last update: April 15, 2016*)

Authority	MAIB	NTSB	USCG	ATSB	TSBC
# Reports	535	152	202	80	407
Period	1989–2015	1994–2015	2005–2015	1987–2015	1990–2015
Country	Great Britain	USA	USA	Australia	Canada

We specifically chose authorities from these countries, because they have English as their first language and a high number of available full-text reports. We share

<sup>2</sup> <https://www.gov.uk/government/organisations/marine-accident-investigation-branch>.

<sup>3</sup> <http://www.nts.gov/investigations/accidentreports/pages/accidentreports.aspx>.

<sup>4</sup> <http://www.uscg.mil/>.

<sup>5</sup> <https://www.atsb.gov.au/marine/>.

<sup>6</sup> <http://www.tsb.gc.ca/eng/rapports-reports/marine/index.asp>.

our full corpus on request to support further research in this area. For our following analysis we focused on the MAIB sub-corpus, but we intend to apply the same method of analysis to the whole corpus in the future (Fig. 1).

### 4 Analysis

The MAIB corpus consists of over five hundred accident reports. In order to classify these accident reports we applied a request-oriented classification approach. We used boolean queries to perform a full-text search on all documents in the corpus. Beforehand the corpus had to be prepared and a list of keywords had to be created in order to build meaningful queries. The preparation of the corpus is described in the following. Thereafter we describe the generation of keywords and our classification method.

#### 4.1 Preparation

In order to perform an analysis on the gathered corpus of accident reports some pre-processing of the corpus is necessary. We gathered full-text maritime accident reports in English and PDF-format from five transportation safety authorities. As preparation for the analysis we extracted the plain-text from the reports in PDF-format and performed simple cleaning of the reports by removing the front page and fixing character encoding issues. Further we converted the reports to lowercase in order to simplify case-insensitive search. As some reports consisted of several files we merged these documents into one. After this preparation one document represents one accident.

For the exploration of the corpus we created a document-term-matrix

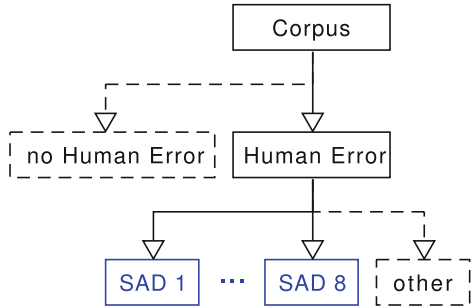


Fig. 1. This chart shows a segmentation of the corpus in relevant subsets. The SA Demons are colored in blue. (Color figure online)



Fig. 2. This is a word cloud of the one hundred most frequent terms in the MAIB Corpus. The terms have been reduced to their stem using the SnowballC stemming algorithm.

of the corpus and checked the most frequent terms. Furthermore this enabled us to check term correlations in the corpus. We used R version 3.2.4 with a number of text-mining packages to perform this. The creation of the document-term-matrix requires some further pre-processing in R. The following pre-processing steps were conducted on the data-frame in R, only. We removed URLs, punctuation, numbers, standard English stop-words and some custom stop-words from the corpus. From the pre-processed corpus we created a tf-idf-weighted document-term-matrix. Figure 2 shows a word cloud of the one hundred most frequent terms in the corpus.

## 4.2 Generation of Keywords

A list of keywords for each SA Demon was created. The selected keywords were derived from the description of the SA Demons and examples from [8] and keywords identified during the exploration of the corpus. The keyword list was complemented with fitting synonyms of the keywords using wordnet [11].

A meaningful choice of keywords and the proper construction of the boolean queries is critical to the success of the retrieval. Adjusting the query based on the exploration of the corpus always bears the danger of overfitting the queries to the specific corpus. Furthermore the composition of the queries directly influences the precision and recall of the retrieval. Although we did not measure the recall, we tried to achieve a good balance between precision and recall.

The iterative exploration of the corpus showed unexpected usage of keywords, such as 'fatigue' in 'fatigue wear' of machine parts. The identification of these negative keyword combinations helped us to increase the precision of our search queries. We increased the recall of the query by reducing the keywords in the queries to their stems. The aim was to create a list of queries that can be applied to any set of maritime accident reports in English.

The generation of fitting keywords is a semi-automated process that highly relies on the judgment of a human analyst. In summary, the process consists of the following steps:

1. derive keywords from definitions and examples
2. find fitting synonyms for keywords using wordnet
3. explore corpus, find term-correlations for the keywords, and add new keywords
4. reduce some keywords to their stems to increase the recall based on human judgment

Furthermore we added keywords for Situation Awareness and Human Error to be able to search on a meta level if none of the keywords for a SA Demon returns any results. Table 2 shows the resulting list of generated keywords for each of the SA Demons.

## 4.3 Retrieval Method

We classified the documents into the SA Demon categories by using boolean queries constructed from the SA Demon keywords. We performed test queries to

**Table 2.** This table shows the generated keywords for each SA Demon. Some of them are reduced to their stem to increase their recall.

SA Demon	Keywords
SAD1	Preoccupied (himself/herself), pre-occupied (himself/herself), focus, not attent, no attent, fixate, concentrate, not note
SAD2	Forget, forgot, not remember, no memory
SAD3	Fatigue, workload, stress; <i>Negative keyword combinations:</i> Fatigue crack, fatigue wear, unlikely, distress, rope stress
SAD4	Occupied, data overload
SAD5	Caught attention, attracted attention, draw attention, distract, mislead attention
SAD6	Misinterpret, not understand, complex
SAD7	Misinterpret, misunderstand, misunderstood, not correct
SAD8	Not aware, unaware
Human Error	Human error, human element, human factor
Situation Awareness	Situation, aware

**Table 3.** This table shows the composition of the boolean search queries for each SA Demon.

SA Demon	Query
SAD1	$(preoccupied \wedge (himself \vee herself)) \vee (pre-occupied \wedge (himself \vee herself)) \vee (attent \wedge not) \vee focus$
SAD2	$(human \wedge error) \vee (human \wedge element) \vee (human \wedge factor) \vee (situation \wedge aware)$
SAD3	$fatigue \wedge \neg crack \wedge \neg wear \wedge \neg not \wedge \neg unlikely \vee workload$
SAD4	$occupied$
SAD5	$(human \wedge error) \vee (human \wedge element) \vee (human \wedge factor) \vee (situation \wedge aware)$
SAD6	$(human \wedge error) \vee (human \wedge element) \vee (human \wedge factor) \vee (situation \wedge aware)$
SAD7	$misinterpret$
SAD8	$not \wedge aware$

identify the keyword combinations with the best balance of precision and recall. Table 3 shows the final queries we used to retrieve the accidents caused by the SA Demons.

For SAD2, SAD5, and SAD6 we could not find a fitting query that specifically retrieves them. Queries constructed from SAD6 keywords always delivered SAD7

problems, as the keywords are too similar. The keywords for SAD5 are often used by the authors of the accident reports to emphasize their findings and recommendations, e.g. “[...] draw the attention of Owners, Skippers, Mates and crews to [...]”. We therefore used a more general query for these SA Demons.

We applied the *pipeline and filters* design pattern to implement the boolean queries as a unix pipeline combining the unix programs *find* and *grep*. The advantage of this approach over using an indexing search engine is the support of full-text search. To remove false positives, the query results were inspected manually in the context of the sentences containing the positive keywords. If we were unsure, the sentence before and after the finding was also inspected.

## 5 Results

The SA Demons with the highest proportion in the investigated sample were *WOFAS* with fatigue as main cause, *errant mental models*, and *attention tunneling*. We were not able to find accidents caused by *requisite memory trap*, *misplaced salience* or *complexity creep*. Not all Situation Awareness related accidents can be explained by the SA Demons. During analysis we found some accidents caused by a lack of Situation Awareness that could not be derived directly from a SA Demon, such as insufficient trip planning.

**Table 4.** Retrieved occurrences of SA Demons in MAIB Corpus. The Percentage behind the absolute counts indicates the count relative to the corpus size.

SA demon	Description	True positives	Query precision
SAD1	attention tunneling	36 (7%)	0.14
SAD2	requisite memory trap	not found	-
SAD3	WOFAS	216 (40%)	0.87
SAD4	data overload	34 (6%)	0.17
SAD5	misplaced salience	not found	-
SAD6	complexity creep	not found	-
SAD7	errant mental models	40 (7%)	0.93
SAD8	out-of-the-loop syndrome	7 (1%)	0.02
Total		<b>333 (62%)</b>	
Expected		<i>300–321 (56–60%)</i>	

Table 4 shows the results of our request-oriented classification. The table lists the absolute count of retrieved accidents and the precision of the query for each SA Demon. The query precision is the positive predictive value of the query-request on the corpus.

It is calculated as follows:

$$query\ precision = \frac{number\ of\ true\ positives}{number\ of\ true\ positives + number\ of\ false\ positives}$$

Moreover we calculated an estimate of the total number of results based on statistics from related work. This provides us with a weak sanity check of the total number of retrieved accidents related to SA problems. For the MAIB corpus this estimate amounts to 300–321 accidents. We based the estimate on the results of Baker et al. stating that in about 80–85% of maritime accidents Human Error is the dominant factor [2]. Further Grech et al. have stated that about 71 % of maritime accidents caused by Human Error are caused by a lack of Situation Awareness [9]. We combined these two statistics in the following calculations:

$$estimate_{lower} = 535 \times (71 \% \times 80 \%) = 535 \times 56 \% = 300$$

$$estimate_{higher} = 535 \times (71 \% \times 85 \%) = 535 \times 60 \% = 321$$

## 6 Discussion

The relationship between documents and SA Demons is a many-to-many relationship. That means more than one SA Demon could have lead to the accident. Our previously introduced weak sanity check suggests that our retrieval was quite successful. However, without knowing the number of false negatives and the recall of our retrieval this is just an educated guess.

We did not find any accidents caused by *requisite memory trap*, *misplaced salience* or *complexity creep*. However, the fact that we were not able to find them does not mean they do not exist. We expected these demons to be hard to find. Endsley has already stated in her definition of the SA Demons that “complexity is a subtle SA Demon” [8]. *Requisite memory trap* is an internal processing problem that is hard to observe. The same applies to *misplaced salience*. As our source of data are accident reports, the completeness of the data regarding SA failures depends on the ability of the respective inspector conducting the investigation to identify human errors and their causes. The completeness therefore varies from inspector to inspector depending on their interpretation abilities and work experience.

The most frequent found SA Demons all affect Situation Awareness Level 1. This confirms prior work by Grech et al. identifying Situation Awareness Level 1 as most prominent cause for a lack of Situation Awareness [9].

## 7 Conclusion

Our results confirm that Situation Awareness Level 1 is the most prominent source of Human Error in maritime accidents. Likewise it is the most prominent one of the three levels of Situation Awareness. All in all, the SA Demons with the highest proportion in the investigated sample were *WOFAS* with fatigue as main cause, *errant mental models*, and *attention tunneling*. We were not able to find accidents caused by *requisite memory trap*, *misplaced salience* or *complexity*



*creep*, however we still think that these exist. Unfortunately detecting these SA Demons in maritime accident investigation reports will remain difficult, unless it becomes part of the investigation itself. We advise taking special care of the SA Demons *WOFAS*, *errant mental models*, *data overload* and *attention tunneling* when designing new interfaces for ship bridges. Our findings might also be beneficial for the design of user interfaces for Vessel Traffic Management (VTM) centers. We intend to improve our retrieval method by labeling ship personnel in the corpus using a custom-built tool for named-entity-recognition of maritime personnel such as the 'master' or the 'OOW'<sup>7</sup>. This will enable us to use higher level queries such as *PERSON*  $\wedge$  *fatigue*. Also, we intend to apply our method of analysis to our full corpus of retrieved reports with refined keywords and queries in the future. We expect to get similar results in relation to the corpus size and to find further examples of SA failures caused by SA Demons.

**Acknowledgments.** We thank the Ministry of Science and Culture of Lower Saxony for supporting us with the graduate school *Safe Automation of Maritime Systems (SAMS)*.

## References

**All links were last followed on April 15, 2016.**

1. Antão, P., Guedes Soares, C., Grande, O., Trucco, P.: Analysis of maritime accident data with bbn models. In: Martorell et al. (eds.) *Safety, Reliability and Risk Analysis: Theory, Methods and Applications*. Taylor & Francis Group, London (2009)
2. Baker, C., McCafferty, D.: Accident database review of human element concerns: what do the results mean for classification?. In: *Proceedings of the International Conference Human Factors in Ship Design and Operation*, RINA Feb. Citeseer (2005)
3. Baker, C.C., Seah, A.K.: Maritime accidents and human performance: the statistical trail. In: *MarTech Conference*, Singapore (2004)
4. de la Campa Portela, R.: Maritime casualties analysis as a tool to improve research about human factors on maritime environment. *J. Marit. Res.* **2**(2), 3–18 (2005)
5. Chauvin, C., Lardjane, S., Morel, G., Clostermann, J.P., Langard, B.: Human and organisational factors in maritime accidents: analysis of collisions at sea using the hfacs. *Accid. Anal. Prev.* **59**, 26–37 (2013)
6. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. *Hum. Factors J. Hum. Factors Ergon. Soc.* **37**(1), 32–64 (1995)
7. Endsley, M.R.: Situation awareness and human error: designing to support human performance. In: *Proceedings of the High Consequence Systems Surety Conference*, pp. 2–9. Lawrence Erlbaum Associates (1999)
8. Endsley, M.R., Jones, D.G.: SA demons: the enemies of situation awareness. In: *Designing for Situation Awareness: An Approach to User-Centered Design*, chap. 3, pp. 31–41. CRC Press (2011)

---

<sup>7</sup> Officer of the watch.

9. Grech, M.R., Horberry, T., Smith, A.: Human error in maritime operations: analyses of accident reports using the leximancer tool. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. vol. 46, pp. 1718–1721. SAGE Publications (2002)
10. Koester, T.: Human error in the maritime work domain. In: Proceedings of 20th European Annual Conference on Human Decision Making and Manual Control, pp. 149–158 (2001)
11. Miller, G.A.: Wordnet: a lexical database for english. *Commun. ACM* **38**(11), 39–41 (1995)
12. Rothblum, A.M.: Human error and marine safety. In: National Safety Council Congress and Expo, Orlando, FL (2000)
13. Wagenaar, W.A., Groeneweg, J.: Accidents at sea: Multiple causes and impossible consequences. *Int. J. Man Mach. Stud.* **27**(5), 587–598 (1987)