The 22nd IAMUC
Batumi, 21-22 Oct 2022

Proceedings of the
International Association of Maritime Universities Conference

Knowledge representation in MET

Jana Kegalj[1,*] and Damir Zec[1]

[1] *University of Rijeka, Faculty of Maritime Studies, Croatia*
[*] *Corresponding author: jana.kegalj@pfri.uniri.hr; Tel.:* ++385-91-562-8594.

Abstract: Maritime education and training are under constant pressure to follow and adapt to the advances and trends in technology to satisfy the needs of the labour market and industry. To do that, besides revising the necessary knowledge and competencies, it is also important to properly communicate them. In that sense, the paper aims to provide insight into how maritime professional knowledge is represented and communicated in various sources. The results presented here were obtained through a corpus-driven analysis of different sources presenting maritime knowledge. Quantitative corpus data and statistical calculations from a corpus consisting of more than a million tokens, or running words, showed the main semantic domains each subcorpus focused on and provided information about word frequencies. Corpus analysis identified the underrepresentation of psychomotor and affective domains and a smaller share of higher-level competencies in the source texts, thus indicating possible areas for further improvement.

Keywords: STCW knowledge representation, GMP, semantic analyses

1. Introduction

Maritime education and training (MET) may be described as a systematic process utilised to transfer knowledge and accumulated experience to students. The scope and extent of subjects to be delivered are predominantly determined by the job description and assumed responsibilities. Consequently, there are different levels, both in scope and depth. The minimum knowledge and competencies for different levels and functions are internationally agreed upon and codified in the International Convention on Standards of Training, Watchkeeping and Certification, 1978 (STCW).

Due to recent accelerated technology development, many have raised the question of the "appropriate" level of knowledge and competencies for future seafarers. The knowledge and competencies traditionally delivered at MET institutions are confronted with changing shipboard organisation, new technologies (particularly communications and AI), accelerated business activities, etc. At the same time, the core set of competencies given in the STCW Convention has not been updated since 2010 or in several subject areas even since 1995. Consequently, the skills gap between given standards and industry needs is clearly recognised (SkillSea, 2020). The number of new or revised non-mandatory courses is increasing as well as a list of IMO approved Module Courses, with more extensive stakeholders' involvement than ever. Even the GMP initiative, promoted by the IAMU, may be understood as a response to these needs by systemising and generally improving the knowledge delivery at the global level, thus improving the efficiency of the MET as a whole.

The maritime professional knowledge, once a sole responsibility of the national authorities, today is to a large extent internationalised and represented in internationally recognised documents. Due to the numerous proficiency levels and activities, maritime knowledge is dispersed in numerous documents, with different obligatory levels, scopes, objectives, styles, structures, etc.

However, in most cases, professional knowledge is represented as statements describing the abilities the students should master during the educational process (Davis, 1993). Such an approach, in general, follows the well-known Bloom's taxonomy. It is understood as a common language to facilitate communication on learning objectives across persons, subject matter, and levels. Yet, it is not used consistently, and in different documents, the target competencies, although very similar, are represented in very different arrangements.

Therefore, the research presented here deals with the modes of knowledge representation in MET and their characteristics in different documents**.** It is based on the semantic analysis of the knowledge statements in

relevant international sources using a corpus-based approach, relying on the concept of semantic frames (cf. Fillmore & Baker, 2010).

2. Research objectives, methods and outcomes

According to one of many definitions, *professional knowledge is traditionally seen as knowledge that has undergone a formal rationalisation, knowledge that is systematic, codified and generalised, hence abstract. It thus accords with the norms of academic education, supporting the argument for professional education to be situated within the universities* (French, 2007). Maritime knowledge definitely corresponds to this definition.

To semantically analyse maritime knowledge, it was necessary to select the sources representing maritime knowledge in systematic, codified and generalised form. Unfortunately, there is no single source; in fact, maritime knowledge, as in many other areas, is dispersed across many sources in highly different forms and structures. Consequently, the authors decided to analyse a range of sources presenting maritime knowledge differently, starting with a highly formal structure (thus the most condensed) to more elaborate but less formal representations. The selected sources (corpora) include:

1 Core STCW competencies for deck and engine crew members, i.e. competencies (knowledge, understanding and proficiencies) formally described in respective STCW tables (Chapters II and III) in the form of the statements describing required capabilities at different levels.

2 Extended STCW competencies (competencies as represented in the STCW Code A & B, without Chapter I, i.e. the core competencies supported with other competencies, recommendations and guidance related to these competencies.

3 Model Courses (Deck, Engine, ETO), i.e. guidance for developing curricula of study programs required for officers at the management and operational level. Here, the competencies are described in more detail and also include guidance on delivering, references, teaching aids, etc.

4 GMP Body of Knowledge (Manuel, 2019), as a document encouraging IAMU member institutions "*to examine the learning outcomes agreed in the BoK and thereafter within the academic freedoms and requirements of their own jurisdictions, develop a curriculum (syllabi, learning activities, assessment methods etc.) that will aim at the achievement of these learning outcomes in a consistent manner.*"

5 IAMU AGA proceedings (AGA 19, 20 and 21), as a corpus representing contemporary MET teachers' perspectives and interests.

6 WÄRTSILÄ Encyclopaedia of Ship Technology (Babicz, 2015), as a comprehensive and systematic representation of the current technological state of the art.

Each corpus represents the common maritime knowledge differently, depending on goals, point of view, intended scopes, audience, etc. It is important to note that selected corpora cover all aspects of maritime knowledge: factual, conceptual, procedural, and metacognitive knowledge (as defined by Krathwohl, 2002) and domains, and therefore may be considered representative of the language register.

The corpora include 1,423,975 tokens, or 1,146,955 words. The basic analyses are accomplished using the Sketch Engine corpus analysis tool. In addition, several analytical tools, such as textual similarity measures and readability, have been implemented in R to gain insight into the semantic similarity or distance among these sources.

The linguistic analysis for each corpus includes:

- identification of the most frequent words (verbs, nouns, collocations and multi-word units),
- the usage of the action verbs identified as the preferred ones for use when developing Model Courses (IMO HTW, 2020),
- readability statistics.

Since complete analyses include numerous outcomes, only a selected set of outcomes is presented in the following paragraphs.

Corpus analysis enables the detection and description of typical patterns and frequent forms that would otherwise be unnoticed, and this shows how communication is a highly structured activity with patterns which are not random but cognitively motivated. The corpora compiled for this study serve to target specific research questions and therefore had to be representative of the area under study and balanced in terms of diverse texts that were incorporated. To meet these requirements, the texts included in the corpora were selected according to the criteria recommended by the EAGLES group, encompassing external (non-linguistic) criteria, such as style and origin, and internal (linguistic) criteria, such as genre and topic.

The basic quantitative data that corpora provide is the frequency of different linguistic units. In particular, the frequency data about action verbs ((IMO HTW, 2020) provided interesting insights into the presentation of maritime knowledge in different corpora. As Figure 1 shows, the action verbs belonging to the cognitive domain are most frequent in all corpora, while those from the psychomotor and affective domains have a smaller share. This might be attributed to the fact that the cognitive domain involves the understanding or recall of specific facts, procedures, and concepts directly related to the topics in the maritime domain. Therefore, they are more easily identified than skills or attitudes related to the psychomotor and affective domains. On the other hand, corpora compiled from STCW have the greatest frequency of psychomotor action verbs, while affective action verbs are most frequent in GMP BoK.

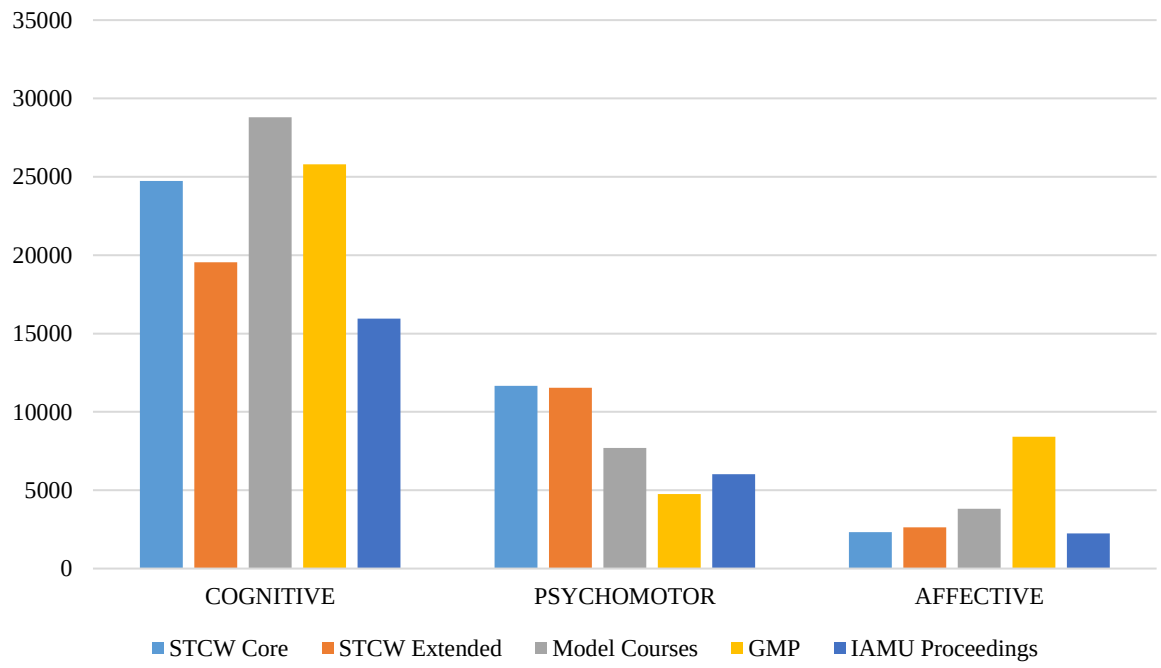

Figure 1 Distribution of the IMO-defined action verbs in corpora by domains

Taking a deeper insight into the cognitive domain, the frequency of related action verbs shows that the emphasis is on the application as the medium level of complexity, particularly in STCW, while higher complexity levels are much less represented. The exception here is the GMP BoK, in which the action verbs of all levels of complexity have similar frequencies, showing its equal dedication to all levels.

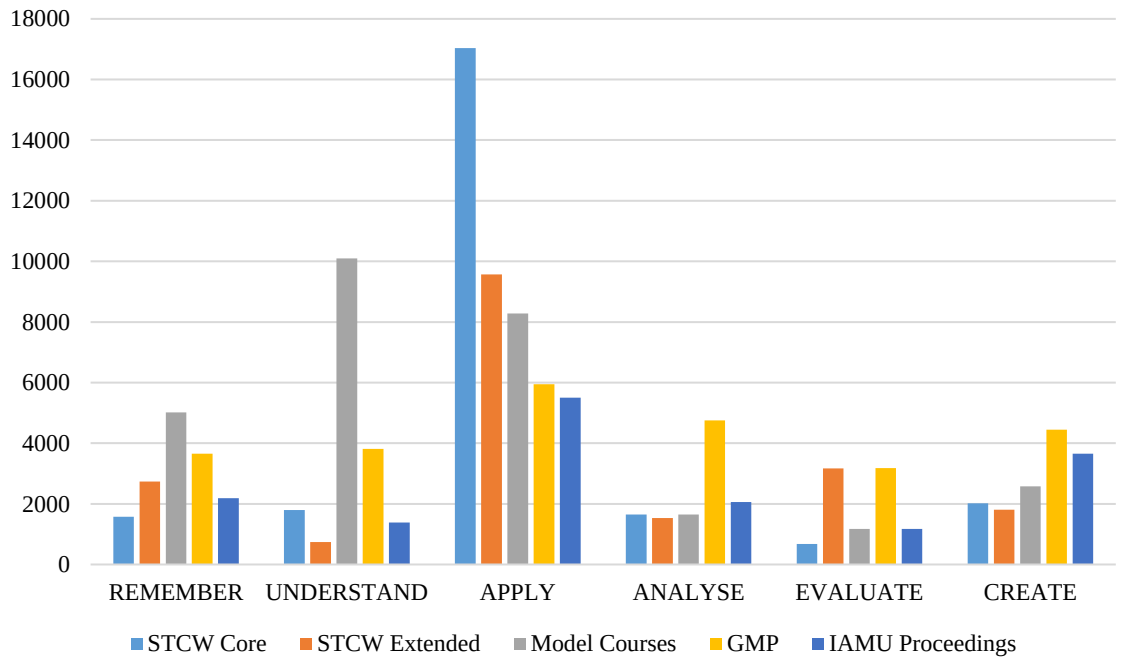

Figure 2 Distribution of the IMO-defined action verbs in corpora by knowledge domains

Furthermore, keyness score analysis was also conducted in Sketch Engine, which indicates the words or phrases typical or key in a certain context, i.e. characteristic for a specific text as opposed to another, in this case, the corpus of general English language EnTenTen20. Keyness is a textual feature providing insight into the main semantic domains of a certain text. The conducted keyness score calculation, shown in Figures 3, 4 and 5, singled out the key terms for IAMU AGA Proceedings, WÄRTSILÄ Encyclopaedia of Ship Technology and GMP BoK, which indicate the main areas the corpora are focused on. In that sense, IAMU AGA Proceedings focus most on education (maritime education, model course, learning outcome, met institution), soft skills (risk assessment, human error, risk management) and future developments (autonomous ship, cyber security, wave energy).

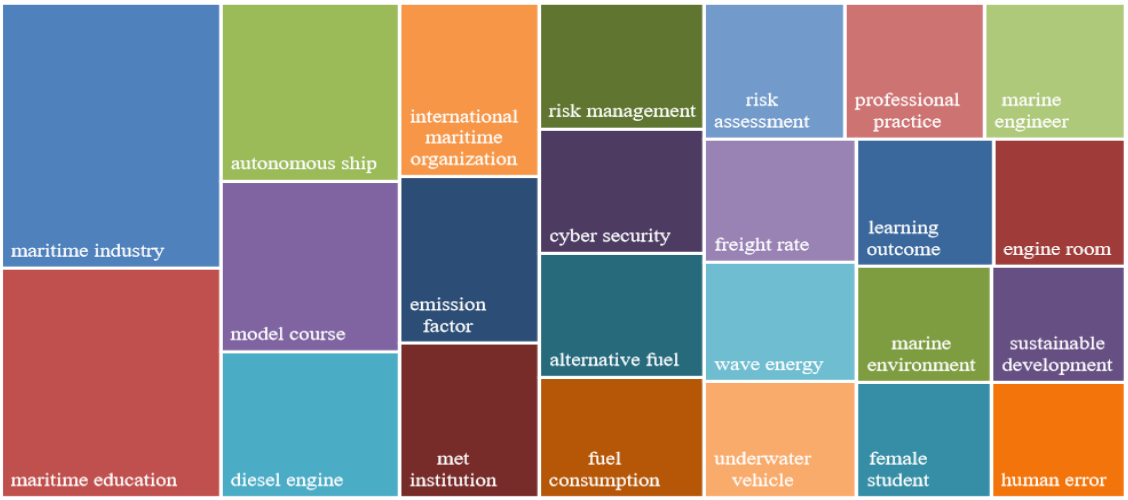

Figure 3 Key terms in IAMU AGA Proceedings

As expected, WÄRTSILÄ Encyclopaedia focuses mostly on technology, while the GMP BoK addresses mostly the topic of knowledge acquisition.

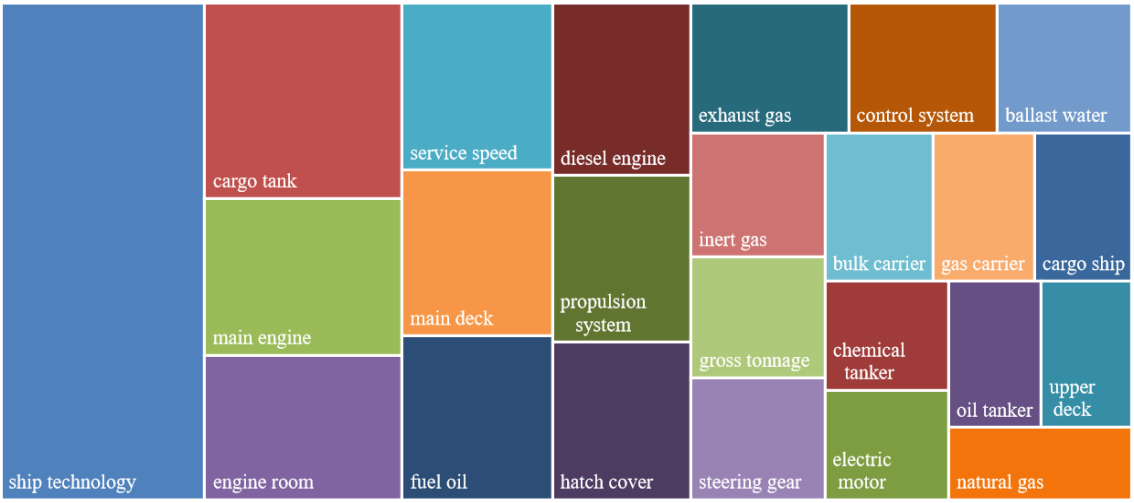

Figure 4 Key terms in WÄRTSILÄ Encyclopaedia of Ship Technology

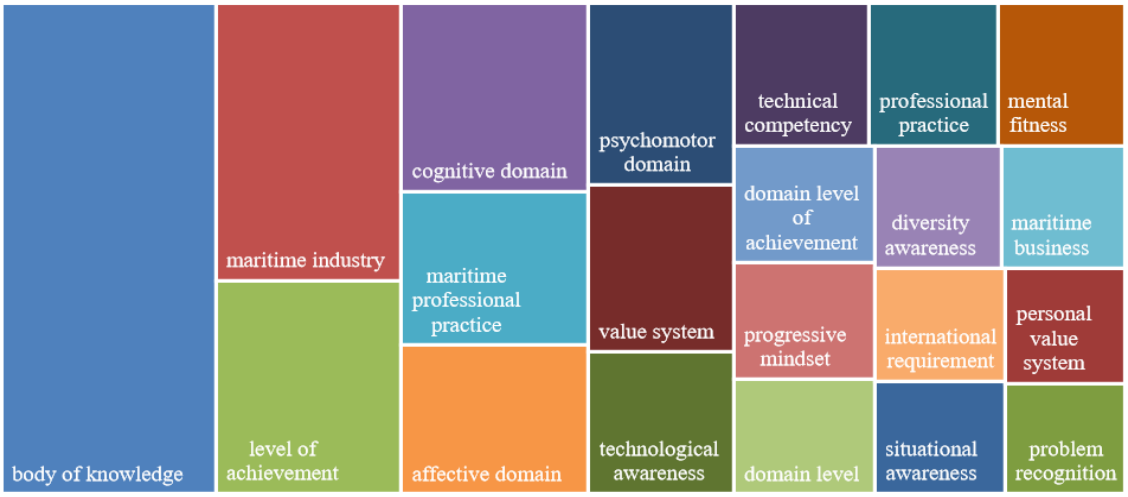

Figure 5 Key terms in GMP BoK

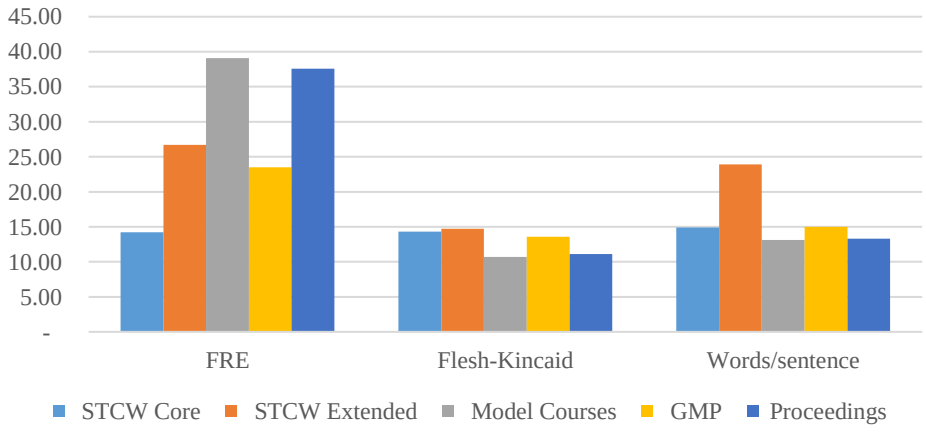

Figure 5 Readability of different corpora

The analysis also included statistical indicators of readability, which quantifies how difficult a text is to read, taking into account several factors, such as the number of syllables in words or the number of characters in a word indicating semantic difficulty, number of words in sentences indicating structural difficulty. According to the diagram in Figure 6, Model Courses are the most readable texts with 13 words per sentence

on average, while GMP BoK and STCW would qualify as the least readable with 15 and 23,9 words per sentence, respectively.

Other corpus analysis data and statistical indicators omitted due to space restrictions, such as keyword frequencies, collocation strength and semantic similarity, showed certain distinctive features of individual subcorpora. Further research could be directed towards identifying semantic frames in different maritime corpora, providing insight into the level of their compatibility.

3. Conclusions

Following the outcomes presented in the previous paragraphs, the main conclusions of the research (in part presented here) are as follows:

1. Psychomotor and affective domains are clearly underrated in all considered documents, the only notable exception being the GMP BoK. Although these domains are much more difficult to codify and formalise, their importance is significant, and further developments in these areas of maritime education are highly welcomed.
2. Most action verbs used to describe required knowledge deal with the application level. Although fully understandable and in line with predominant understanding in the shipping industry in the past, it is highly questionable whether a focus on the application is sufficient to enable future seafarers to operate highly sophisticated ships or onboard systems. Again, the only notable exception is GMP BoK containing more action verbs defining higher knowledge levels. Such a low level of dedication to higher, more creative levels of a knowledge domain is not satisfactory for maritime universities, which are assumed to pursue those higher levels in their curricula.
3. Key terms identified as the most important in the IAMU AGA Proceedings show the broad interest of contributing authors, somehow equally distributed between new technologies and demands (e.g. autonomous ships, emission factors, and cyber security) and educational subjects (e.g. model courses, MET institutions, learning outcomes).

Acknowledgements

The part of the research presented here has been initiated and carried out within the SkillSea project (a project funded by the European Union, Erasmus+ program grant – 2018-3387/001-001 Project number 601186).

References

[1] Farnese, M. L., Barbieri, B., Chirumbolo, A., & Patriotta, G. (2019). Managing knowledge in organisations: A Nonaka's SECI model operationalisation. Frontiers in psychology, 2730.

[2] Fillmore, C. J., & Baker, C. (2010). A frames approach to semantic analysis. In the Oxford handbook of linguistic analysis.

[3] French, N. (2007). Professional Knowledge, Professional Education and Journalism. IAMCR Conference. Paris, July 2007.

[4] Davis, R., Shrobe, H., & Szolovits, P. (1993). What is a knowledge representation?. AI magazine, 14(1), 17-17.

[5] Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. Theory into practice, 41(4), 212-218.

[6] Manuel, M., Ghalwash, G., Elbawab, M. E., Ahvenjarvi, S., Nakazawa, T., Farag, Y. B., ... & Rowihil, M. (2019). Global Maritime Professional Book of Knowledge (GMP-BoK).

[7] Nonaka, I. (1994). A dynamic theory of organisational knowledge creation. Organization Science, 5(1), 14-37.

[8] WÄRTSILÄ Encyclopedia of Ship Technology. (2015). Compiled by Jan Babicz

[9] IMO HTW (2020) Development of amendments to the revised guidelines for the development, review and validation of model courses (MSC-MEPC.2/Circ.15/rev.1), Report of the Correspondence Group

[10] EAGLES, Preliminary Recommendations on Text Typology, EAGLES Document EAG–TCWG–TTYP/P, Version of June 1996 (URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.28.1988&rep=rep1&type=pdf)

[11] SkillSea (2020). Current skill needs – Reality and mapping. https://www.skillsea.eu/index.php/news-events/spotlight/106-read-the-full-report-on-currents-skills-needs