What is a Control Room? In Search of a Definition using Word Embeddings

Jonas Pöhler
University of Siegen
Siegen, Germany
jonas.poehler@uni-siegen.de

Nadine Flegel
Trier University of Applied Sciences
Trier, Germany
N.Flegel@hochschule-trier.de

Tilo Mentler
Trier University of Applied Sciences
Trier, Germany
mentler@hochschule-trier.de

Kristof Van Laerhoven
University of Siegen
Siegen, Germany
kvl@eti.uni-siegen.de

Abstract

Definitions are hard. But they provide valuable insight into the subject at hand. This paper presents a comprehensive analysis of the term "control room" using word embeddings to provide a nuanced definition. The study leverages the ACM Digital Library to compile a corpus of texts related to control rooms, generating word embeddings to capture semantic relationships. Key aspects identified include the physical space, ergonomic design, multimodal information presentation, warning systems, contextual awareness, and command functions of control rooms. The analysis also explores adjacent clusters, highlighting tactile interfaces, interactive surfaces, and the use of large displays to enhance operator performance and situational awareness. The findings underscore the technological, ergonomic, and operational dimensions critical to the definition and functionality of control rooms.

CCS Concepts

• **Human-centered computing → Ubiquitous computing**; *Information visualization.*

Keywords

definition of control room, word embeddings

1 Introduction

Control rooms play a pivotal role in ensuring the security and well-being of humans across various domains, from emergency medical services to utility management. A comprehensive understanding of control rooms begins with their definition. According to Hollnagel and Woods, control rooms are "location[s] designed for an entity to be in control of a process." [9]. This definition emphasizes the centrality of control rooms in managing critical processes, highlighting their importance in maintaining order and responding effectively to emergencies. Further elaborating on this concept, Santos et al. describe control rooms as "place[s] with an associated physical structure, where the operators carry out [...] responsibilities." [18]. This definition introduces the idea of a control room as a physical space where operators perform their duties. It underscores the physicality of control rooms, which are often characterized by a well-defined structure, frequently enclosed by walls and sometimes windowless, as noted by Schoen [20]. The physicality is also underscored by the definition of a control room in DIN EN ISO 11064: "core functional entity, and its associated physical structure, where control room operators are stationed to carry out centralized control, monitoring and administrative responsibilities" [1]. But finding a functional definition which encompasses all aspects of a control room is difficult, due to the complexities of natural language that pose a significant challenge when attempting to define abstract concepts or domain-specific terminology. In this regard, the use of word embeddings has emerged as a promising approach, as these computational representations can capture the semantic relationships between words, potentially offering a more nuanced understanding of their meanings [5].

When constructing word embeddings, the underlying assumption is that words with similar contexts or usages tend to have similar meanings [6, 17]. This is known as the "distributional hypothesis" [6], which has been explored in various fields, including speech recognition [2, 10]. By leveraging this principle, word embeddings can be employed to investigate the definition of the term "control room" and uncover its semantic associations.

One approach to defining "control room" using word embeddings could involve training a model on a large corpus of text, such as academic or technical literature, that is likely to contain relevant usage of the term [6, 10, 17]. The resulting vector representations could then be analyzed to identify words that are semantically similar to "control room", potentially providing insights into its core characteristics and functionalities. For this project, we downloaded all the titles for every paper available on the ACM Digital Library related to "control room". The analysis provides insight into the aspects of a control room that have been studied in literature, as well as the terms commonly associated with research on control rooms. Research in control rooms especially is highly diverse and covers a large set of control room environments, making a standard survey analysis of existing research work in this area very difficult. his work is an attempt to define the concept of a control room through a deeper analysis of research publications in computer science

Mensch und Computer 2024 – Workshopband, Gesellschaft für Informatik e.V., 01.-04. September 2024, Karlsruhe, Germany

https://doi.org/10.18420/muc2024-mci-ws13-208

literature." "Although we specifically target literature for control rooms, these methodologies might also apply to other targets of research, where more manual approaches such as PRISMA [22] tend to result in time-intensive undertakings. The remainder of this paper is structured as follows: In the next section, the related work is presented. Then we elaborate on how the corpus for the data analysis was selected. After that the process how the word groups where refined is presented, followed by a discussion about the results in section 5.

2 Related Work

Extensive research has been conducted on the application of word embeddings to various natural language processing tasks [8]. These studies have demonstrated the utility of these techniques in capturing semantic relationships and facilitating tasks such as speech recognition, text classification, and knowledge representation.

For instance, [17] provides a comprehensive overview of the fundamental principles underlying word embeddings, highlighting their ability to encode word meaning and relationships in a densely populated vector space. The authors emphasize the distributional hypothesis as a key concept in the development of these computational representations, which aligns with the proposed approach for defining "control room".

Similarly, it explores the role of visualization in working with word embeddings and emphasizes the potential for interactive analysis to reveal insights about the semantic properties of words and their associations.

Furthermore, the introduction of word embedding techniques, such as GloVe and word2vec, has significantly advanced the field of natural language processing, as these methods have demonstrated the ability to capture contextual and semantic information in a more nuance manner compared to traditional bag-of-words approaches [12, 16, 23].

Word embeddings can also be used to explore the relative significance and meaning of words within a given context [3, 7]. For instance, one study utilized word embeddings to identify features related to ischemic stroke, acuity, and level of consciousness in medical text data, highlighting the potential for these techniques to support domain-specific applications [15].

There is also considerable discussion in the field of computer-supported cooperative work about what the concept of a control room entails. CSCW researchers have examined the collaborative nature of control room operations, exploring how control room personnel work together to monitor, coordinate, and respond to complex industrial processes. This includes investigating the design of control room interfaces and technologies to support effective teamwork and information sharing among control room operators and other stakeholders involved in the control of industrial systems [14, 21].

3 Corpus Selection and Model Training

To investigate the definition of "control room" using word embeddings, we first need to select an appropriate corpus of text. As mentioned, the use of academic and technical literature is likely to yield relevant usage of the term, as control rooms are commonly encountered in various domains, such as industrial, transportation, and emergency management settings. In this study, we utilized the ACM Digital Library, a comprehensive repository of computer science and related literature, to gather a corpus of text related to control rooms. We downloaded the titles for all available papers that mentioned the term "control room", resulting in a dataset of 151,573 titles. The search was a fulltext search so that the term "control room" could have been appearing anywhere in the paper. We specifically chose ACM Digital Library due to its accessibility and because of several high-impact venues that focus on interaction (such as CHI or CSCW). For this corpus word embeddings were generated using the OpenAI word embedding model. This

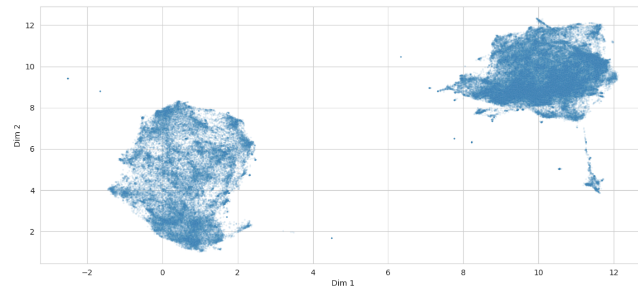

Figure 1: Scatter plot of the first two dimensions of the reduced vector space.

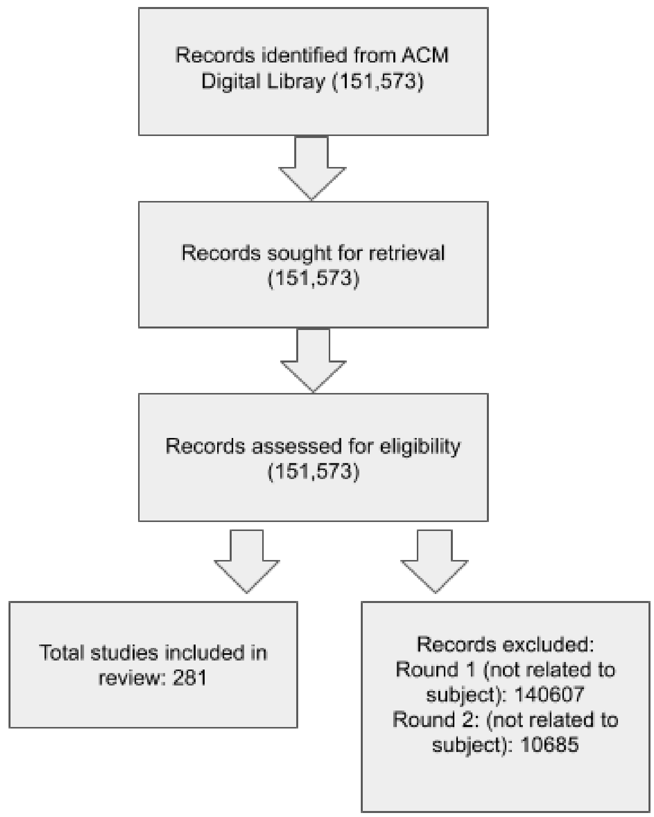

Figure 2: Flow chart for the corpus selection process.

resulted in a vector representation of length 3,072 for each word in the corpus, capturing the semantic relationships between the terms.

4 Refinement Process

Due to the substantial size of the vector space, directly comparing and searching for relationships between the entries is computationally intensive. Therefore, we reduced the vector space in the following step using Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) to its four most significant dimensions [11]. The scatter plot in Figure 1 demonstrates a distribution with two primary clusters when mapping the first two dimensions. Next, we mapped the entries to a hierarchical cluster using Ward variance minimization algorithm to compute the linkage matrix [13].

We then transformed this hierarchical cluster into a flat cluster by determining an optimal number of clusters based on silhouette scores and employing the kneedle algorithm for decision-making [19]. The 78 unique clusters correspond to various papers on control rooms. Each cluster was analyzed using a word cloud to emphasize the key words. Term Frequency Inverse Document Frequency was used for vectorizing the titles of the papers in order to compute these word clouds [4]. The analysis revealed that many of them do not relate to safety-critical control rooms, but rather concentrate on wider topics such as computer systems or software engineering. Subsequently, clusters directly related to safety-critical control rooms were isolated, reducing the number of relevant papers to 7,214. The word clouds of those relevant clusters are shown in Figure 3.

Due to the vague nature of these word clouds, a second round of refinement was conducted. The distribution of the reduced vector space for those 7,214 papers is depicted in Figure 4, but no clear clusters are evident at first glance.

After this second round, an optimal clustering of 160 clusters was achieved and the most significant word clouds were selected. Figure 5 shows the resulting word clouds from prominent clusters. Those clusters were selected by identifying papers featuring "control room" prominently in the title and the clusters these papers belong to.

Following this step, the number of papers was further reduced to 281 and an optimal amount of 22 distinct clusters (displayed in Figure 6) were computed, with the core control room cluster identified by word cloud shown in Figure 7 and highlighted in Figure 6.

5 Results

The final analysis of the core control room cluster reveals the common terms associated with the concept of a control room. Further examination of the hierarchical word clusters reveals several specific aspects and functionalities of industrial control rooms. The most prevalent terms identified are "industrial," "control," "room," "physical," "multimodal,", "environments" and "display."

The analysis of the literature emphasizes the central role of the industrial control room as a physical facility designed to oversee and manage complex industrial processes. These control rooms are integral to ensuring the efficient and safe operation of various industrial activities.

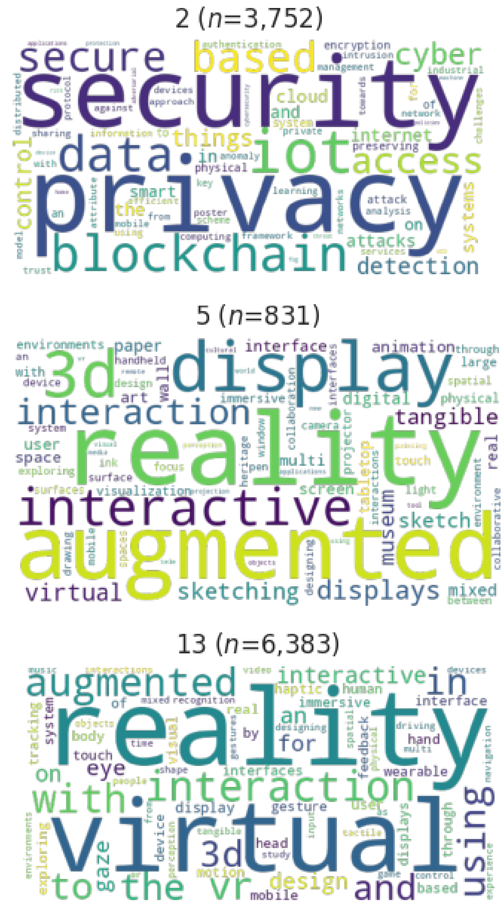

Figure 3: Word cloud of the selected clusters after round 1.

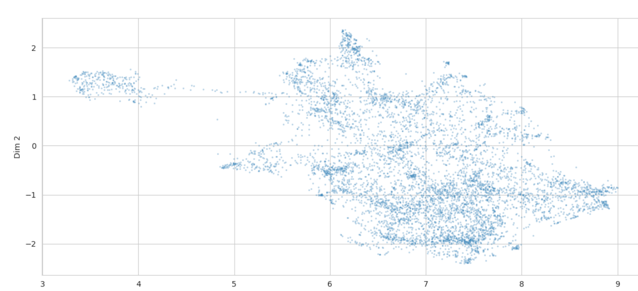

Figure 4: Scatter plot of the first two dimensions of the reduced vector space for the second round.

The concept of "multimodal" information presentation and interaction within control rooms is highlighted, encompassing visual, auditory, and potentially haptic feedback mechanisms. This multifaceted approach aims to enhance operators' situational awareness

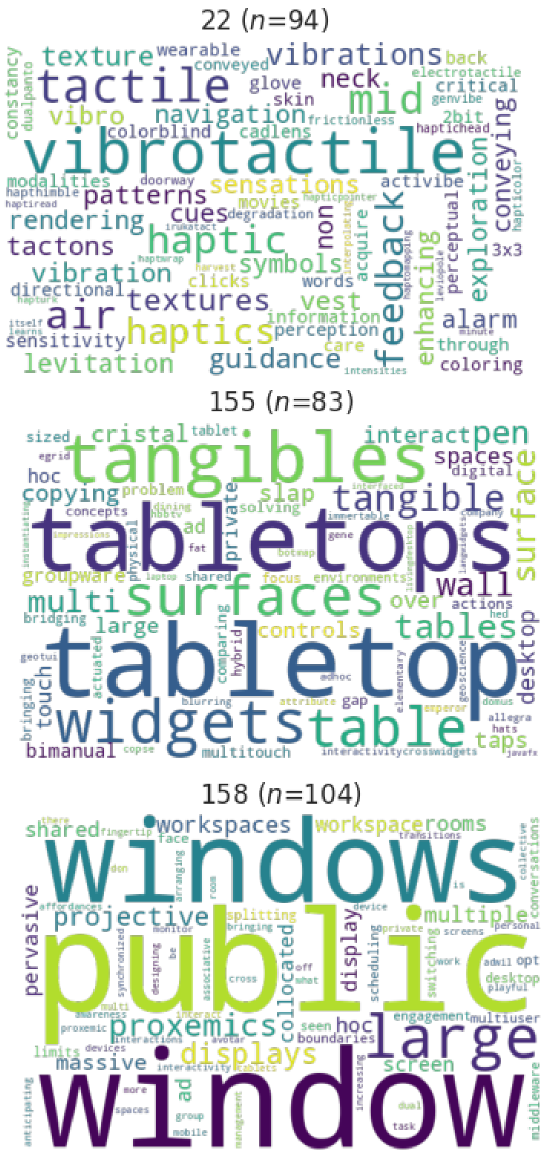

Figure 5: Word cloud of the selected clusters after round 2.

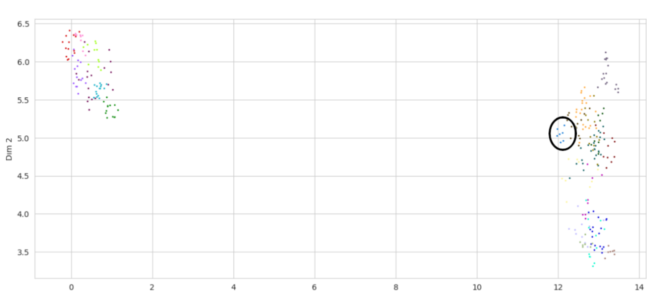

Figure 6: Scatter plot of the final distribution. The core control room group of papers is encircled on the right.

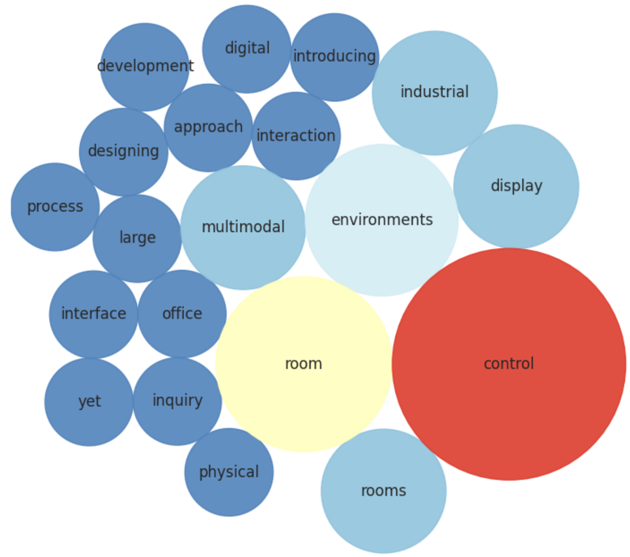

Figure 7: Hierarchical word cluster of the core control room group.

and decision-making capabilities by providing information through multiple sensory channels.

The prominent focus on large and pervasive displays in control rooms underscores their vital role in presenting comprehensive overviews of system states and processes. These visual information dissemination systems enable operators to effectively monitor, analyze, and respond to the dynamics of their operations, facilitating continuous oversight and quick access to relevant data.

Furthermore, the analysis reveals a strong emphasis on the control room as a tangible environment, where operators physically interact with systems and technologies to ensure the smooth functioning of processes.

This emphasis on the physical space and infrastructure of control rooms emphasizes their significance as the central hub for command, control, and coordination of their activities.

In conclusion, the analysis of the word embeddings reveals that the core concept of a control room encompasses several key aspects: the physical space and infrastructure, ergonomic design considerations, multimodal information presentation, warning systems, contextual awareness, and the command and control of industrial processes.

Our analysis also identified several adjacent hierarchical word clusters (see Figure 8) that provide additional insights into the nuanced functionalities and design considerations of control rooms.

First, a cluster of terms such as "tangibles," "tabletop," and "surfaces" points to the use of tangible user interfaces and interactive surfaces in control rooms. Tangible interfaces involve physical objects that can manipulate digital information, often on large,

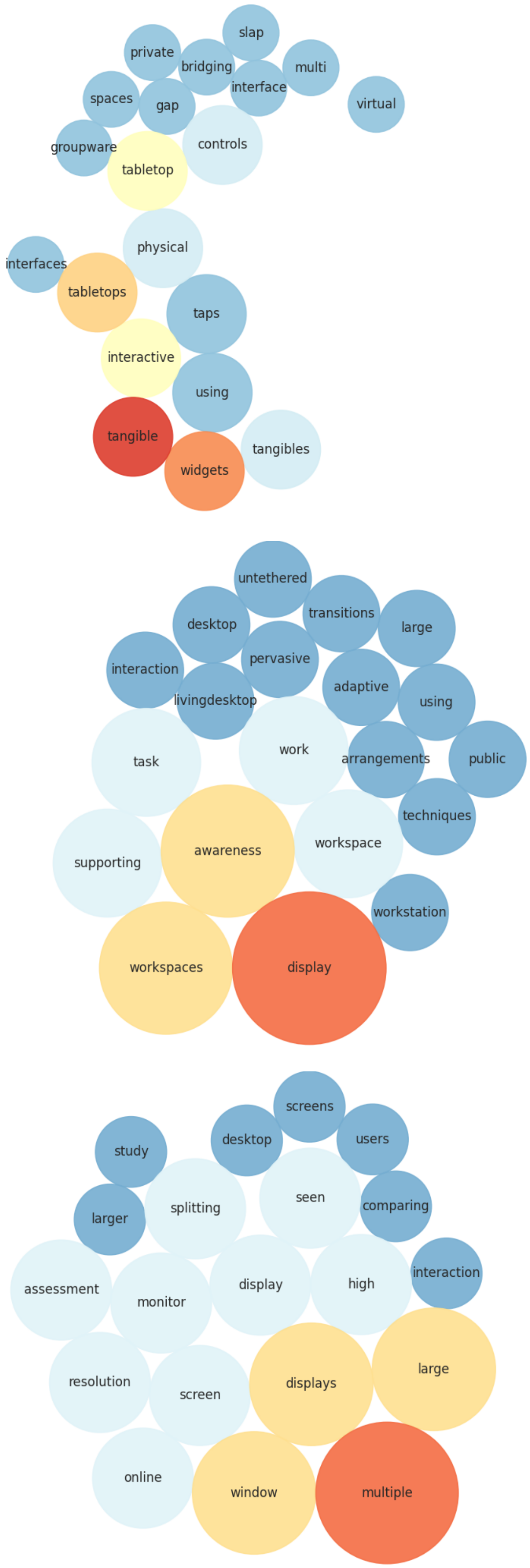

Figure 8: Hierarchical word clusters of the groups directly adjacent to the core control room group.

collaborative tabletop surfaces. These interfaces enable more intuitive and direct data/control manipulation, improving the operator experience and facilitating teamwork.

Second, a cluster of words like "workspace," "awareness," and "displays" indicates a focus on creating an optimal, intuitive workspace for control room operators. This includes large, comprehensive display systems that provide situational awareness and facilitate decision-making.

Third, a cluster with "large displays," "public displays," and "pervasive workspaces" emphasizes the importance of comprehensive visual information dissemination. Large and public displays are crucial for presenting system states and processes, allowing operators to effectively monitor and analyze vast amounts of data. The "pervasive" nature of these display systems ensures critical information is always accessible, supporting continuous monitoring and operational responsiveness.

6 Conclusions

This comprehensive analysis of the literature on control rooms, using word embeddings and hierarchical word clusters, provides a detailed understanding of the various themes and functions associated with these critical environments. The focus on industrial control rooms highlights their primary role in overseeing and managing complex industrial processes. These control rooms are characterized by their emphasis on monitoring, control, automation, safety, communication, and optimization.

Tangible user interfaces and interactive tabletop surfaces emphasize the need for intuitive and collaborative control mechanisms, enabling more direct manipulation of digital information through physical objects and shared workspaces. Large and pervasive displays underscore the necessity for comprehensive and readily accessible visual information to support continuous monitoring and decision-making, ensuring that critical data is always available to operators.

Ultimately, the design of modern control rooms reflects a concerted effort to create technologically advanced, ergonomically optimized, and user-centric environments that enhance operator performance, situational awareness, and overall efficiency in managing processes.

Drawing on these insights, we can define a control room as a centralized environment designed to monitor, control, and optimize complex processes and systems, particularly within industrial settings. It integrates advanced technological solutions, including automation, multimodal feedback, and large-scale displays, to enhance situational awareness and operational efficiency. Ergonomically optimized for operator comfort and effectiveness, control rooms facilitate continuous communication and coordination among operators and stakeholders. They employ tactile and tangible interfaces to provide intuitive control mechanisms, ensuring a responsive and safe operational environment. Through the pervasive dissemination of critical information and real-time data analysis, control rooms play a crucial role in maintaining the stability, safety, and productivity of processes.

This definition captures the multifaceted nature of control rooms, highlighting their technological, ergonomic, and operational dimensions. It underscores the centrality of these spaces in managing activities and responding to the dynamic challenges of their environments.

This study, while comprehensive, is not without its limitations. The analysis is based on predominantly computer science research articles, and a more extensive literature review could potentially uncover additional insights. Furthermore, this research has shown that a lot of the literature is about classic control rooms in industrial settings, and it is important to note that the findings may not directly translate to non-traditional control rooms in other domains. Additionally, the use of word embeddings and word cloud analysis, while powerful, has inherent limitations in capturing the nuances and context-specific meanings of terms. To fully understand the complex nature of control rooms, a more in-depth qualitative analysis, including interviews with subject matter experts, may be necessary.

References

[1] [n. d.]. DIN EN ISO 11064-1:2001-08, Ergonomische Gestaltung von Leitzentralen_- Teil_1: Grundsätze für die Gestaltung von Leitzentralen (ISO_11064-1:2000); Deutsche Fassung EN_ISO_11064-1:2000. https://doi.org/10.31030/9129488

[2] Samy Bengio and Georg Heigold. 2014. Word embeddings for speech recognition. https://doi.org/10.21437/interspeech.2014-273

[3] José Camacho-Collados and Mohammad Taher Pilehvar. 2018. From Word To Sense Embeddings: A Survey on Vector Representations of Meaning. https://doi.org/10.1613/jair.1.11259

[4] Hans Christian, Mikhael Agus, and Derwin Suhartono. 2016. Single Document Automatic Text Summarization using Term Frequency-Inverse Document Frequency (TF-IDF). *ComTech: Computer, Mathematics and Engineering Applications* 7 (12 2016), 285. https://doi.org/10.21512/comtech.v7i4.3746

[5] Kenneth Church. 2016. Word2Vec. , 155-162 pages. https://doi.org/10.1017/s1351324916000334

[6] Felipe Luan Souza de Almeida and Geraldo Xexéo. 2019. Word Embeddings: A Survey. https://doi.org/10.48550/arxiv.1901.09069

[7] Katrin Erk, Diana McCarthy, and Nicholas Gaylord. 2013. Measuring Word Meaning in Context. https://doi.org/10.1162/coli_a_00142

[8] Florian Heimerl and M. Gleicher. 2018. Interactive Analysis of Word Vector Embeddings. , 253-265 pages. https://doi.org/10.1111/cgf.13417

[9] Erik Hollnagel and David D Woods. 2005. *Joint cognitive systems: Foundations of cognitive systems engineering*. CRC press.

[10] Amit Mandelbaum and Adi Shalev. 2016. Word Embeddings and Their Use In Sentence Classification Tasks. https://doi.org/10.48550/arxiv.1610.08229

[11] Leland McInnes and John J. Healy. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. https://arxiv.org/abs/1802.03426

[12] Tomáš Mikolov, Kai Chen, Greg S. Corrado, and J. Michael Dean. 2013. Efficient Estimation of Word Representations in Vector Space. https://doi.org/10.48550/arxiv.1301.3781

[13] Fionn Murtagh and Pierre Legendre. 2011. Ward's Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm. http://arxiv.org/abs/1111.6285

[14] Maria Normark, Petter G. Almklov, Paul Luff, and Ilaria Redaelli. 2024. Introducing Moving Back to the Control Room—Revisiting Centres of Coordination. *Computer Supported Cooperative Work (CSCW)* 33, 1 (March 2024), 1–6. https://doi.org/10.1007/s10606-023-09477-4

[15] Charlene Ong, Agni Orfanoudaki, Rebecca Zhang, Francois Pierre M. Caprasse, Meghan R. Hutch, Liyuan Ma, Darian Fard, Oluwafemi Balogun, Matthew Miller, Margaret A. Minnig, Hanife Saglam, Brenton Prescott, David M. Greer, Stelios M. Smirnakis, and Dimitris Bertsimas. 2020. Machine learning and natural language processing methods to identify ischemic stroke, acuity and location from radiology reports. , e0234908-e0234908 pages. https://doi.org/10.1371/journal.pone.0234908

[16] Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global Vectors for Word Representation. https://doi.org/10.3115/v1/d14-1162

[17] Navin Sabharwal and Amit Agrawal. 2021. Introduction to Word Embeddings. , 41-63 pages. https://doi.org/10.1007/978-1-4842-6664-9_3

[18] Isaac JA Santos, Nelson Fernandes, and Luana F Marques Santos. 2019. Evaluation of performance mode of the operators of nuclear control rooms: an approach based on human factors. In *Proceedings of the INAC 2019: international nuclear atlantic conference. Nuclear new horizons: fueling our future.*

[19] Ville Satop¨a¨a, Jeannie Albrecht, David Irwin, and Barath Raghavan. [n. d.]. Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior.

[20] Donald A. Schon. 1992. *11. Educating for Reflection-in-Action.* University of Pennsylvania Press, Philadelphia, 142–161. https://doi.org/doi:10.9783/9781512801576-013

[21] Antti Silvast, Mikko J. Virtanen, and Simone Abram. 2022. Habits Over Routines: Remarks on Control Room Practices and Control Room Studies. *Computer Supported Cooperative Work (CSCW)* (Dec. 2022). https://doi.org/10.1007/s10606-022-09460-5

[22] Peter Tugwell and David Tovey. 2021. PRISMA 2020. *Journal of Clinical Epidemiology* 134 (2021), A5–A6. https://doi.org/10.1016/j.jclinepi.2021.04.008

[23] Peter D. Turney and Patrick Pantel. 2010. From Frequency to Meaning: Vector Space Models of Semantics. https://doi.org/10.1613/jair.2934