User-centered Data Driven Approach to Enhance Information Exploration, Communication and Traceability in a Complex Systems Engineering Environment

Fahim A. Salim
Department of Science and Industry Systems
University of South-Eastern Norway
email: fahim.a.salim@usn.no

Haytham B. Ali
Department of Science and Industry Systems
University of South-Eastern Norway
email: haytham.ali@usn.no

Gerrit Muller
Department of Science and Industry Systems
University of South-Eastern Norway
email: gerrit.muller@gmail.com

Kristin Falk
Department of Science and Industry Systems
University of South-Eastern Norway
email: kristin.falk@usn.no

***Abstract*—Organizations working on complex engineering projects have data scattered across many different systems. The data is often disconnected, and its potential remain largely untapped. Enterprises large and small find it difficult to explore the information cluttered around different systems. A major factor in this difficulty is a lack of the user perspective in complex engineering environments. The presented research focused on a case study of information exploration needs of engineers testing sub-sea equipment. The case study observed that enterprise software tools in complex systems engineering environment are often designed for the content producer and not the consumer. Which makes these tools difficult, and time consuming to use and discourage their adoption. By utilizing user-centered design and co-creation, the experience and needs of users is identified to design, a data driven approach for enhance information exploration. The proposed design has the potential to make modifications to existing information systems, that would create a large impact in information exploration, data utilization and would provide a better experience for engineers, management and other stakeholders and enhance the productivity of teams, equipment testing and design.**

***Index Terms*—User-centred Design; Participatory Design; Big data; Early Phase Decisions;**

I. Introduction

Most organizations working on complex engineering projects, generate a large amount of document centred data during the project's life cycle [1] [2]. The data is usually scattered around many different often disconnected systems where it is not easily explorable [1] [3]. Organizations understand the need of getting more value out of their data to improve their routine operations. However, its potential remains largely untapped [4].

There are many factors identified by researchers, which limits the utilization of available data in a complex systems engineering environment to enhance operations. These includes among others, a lack of common language among engineers and interdisciplinary teams, ineffective knowledge sharing, finding system information and ineffective communication among stakeholders [1] [5]–[8].

The presented work focuses on the issue of data dispersed around different systems and is not easily searchable by users. As a case study, we focused on test engineers testing sub-sea oil drilling equipment at our partner organization, which is a multinational corporation that provides complete project life cycle services for the energy industry.

The problem is not exclusive for our case study. Organization such as enterprises large and small find it difficult to explore the information cluttered around different system which may or may not be connected to each other [4] [7] [9]. One solution to this problem is to use a model base system engineering approaches and store data in central repository [2], [10]. However, it is not very likely as complex engineering projects involve different companies such as vendors, other contractors, separate subsidiaries, and different departments geographically dispersed with their own workflows and Standard Operating Procedures (SOP). Therefore, the data is bound to end up in silos and would require exploration through multiple systems.

A major factor in this difficulty is a lack of the user perspective in complex systems engineering environments [1] [3] [11]. In a recent survey, Qlik and Accenture [4] found that only 21% engineers reported to have access to dashboard style exploration system suitable for their job roles while 18% said that the systems were appropriate for their skill level.

Researchers have found that while employees are regarded as crucial in utilizing big data systems, they are very often neglected in the design process [12]. The needs of user control is important in the data exploration process [13] [14] and also

 ISBN: 978-1-68558-021-6

there is a need to take in multiple perspectives in a soft systems context [11].

Building upon previous work [3], It is our contention that designing a data driven methodology to enable end users to explore the content in an efficient manner in a complex systems engineering environment would help users in utilizing the available data in efficient manner to enhance operations and make data driven early phase decisions.

The remainder of the paper is as follows: Section II describes the related work. The case study and proposed approach is described in Section III. Which is followed by a discussion in Section IV and finally Section V contains the concluding remarks and some plans for the future.

II. Related Work

Dashboard style systems are used extensively to explore data [15]. They often use semantic and video data [16]–[18]. Dashboard style systems often follow a familiar pattern of providing detail on demand access to multimodal information based on an issued query. Research is also done on novel visualization to give quick insights for further exploration [19] [20], and on new ways of querying large data sets [21].

In their recent survey regarding the utilization of big data in enterprise environments, Qlik and Accenture [4] found that about 21% of engineers have access to dashboard style exploration systems which are designed to be suitable for their needs and 18% believe that the tools are designed in accordance to their skill level. Disparate data silos across multiple systems and technical capabilities are regarded as major obstacles [9].

Researchers have long identified the need for more user control in exploration of data [13] and need for taking multiple perspective when designing data exploration in a soft systems context [11]. General purpose commercially available data exploration and document retrieval tools have limitations [2] [7] [22].

Co-creation [23] and user centred design (UCD) [24]–[26] are often utilized to understand the user techniques to design systems which are suitable for evolving user needs. Haesen et al. use a user centred design approach and semi-structured interviews to perform a qualitative analysis of the information exploration needs of professional video searchers [25]. Similar approach is utilized by Gravier et al. [24] in which authors utilized mockups and non-functional prototypes to understand the information exploration needs of different users to come up with design solutions. Their analysis concluded that different users may have different needs in terms of desired features in exploration systems. Salim et al. [26] extend this approach and use semi-functional prototypes, screen capturing, think out loud and questionnaires to understand the user journeys in engaging exploration of content. Similar techniques with additional semi-structured interviews is used by Koesten et al. [27] to understand users sense making with data-sets for recommending design of tools and documentation practices to allow data reuse.

To allow engineers to explore content beyond simple document retrieval, researchers have proposed solutions like social network systems specifically designed for engineering design communication and project management [22] model based system engineering (MBSE) approaches to model central repositories to allow engineers and other stakeholders access to relevant information [2].

Currently, there is a lack of taking in the user perspective in designing exploration systems [12]. There exist a need for considering the multiple user perspectives and user needs while designing exploration systems [3] [11] [28].

III. Case Study and Proposed Design

The presented research is utilizing Industry as a Laboratory [29] , user-centered design [28] and Human Media Interaction [14] [26] methodologies to focus on the issue of data dispersed around different systems and is not easily searchable by users.

As a case study we focused on test engineers at our partner organization. The partner organization is a multinational corporation that provides life cycle services for the energy industry. We conducted multiple workshops, observation sessions and semi-structure interviews to understand the real-life context of the company(team) of interest in a complex engineering project. While the sessions overall focused on the holistic view of the early phase activities of company of interest in acquiring and developing/executing projects (subject of future publications). The current paper is focused on the information needs of test engineers.

A. Information Needs of Test Engineers

Test engineers need to access information during their routine operations. The current case study focuses on test engineers running different test on sub-sea oil drilling equipment at their workshop. They require access to information about the equipment they are testing. The information they require about different parts may include:

- The project specification.
- The design specification of parts.
- The provider or designer of the part (i.e., the supplier).
- Product responsible department or person at that department or supplier organization.
- The maintenance history of the part.
- The information about any modification done to the part and its history.

During the different tests they run, e.g., System Integration Test (SIT) where a combination of different part is tested to see if the parts are performing as per the specifications. In case of an unexpected event such as equipment not performing as expected or has a fault or have a part missing or have any other problem. It is logged in an event log. The event contains information such as data about the equipment and its serial number, etc. along with the description of the problem. The log also contains information about the project such as work package and product responsible department or supplier and other project specific information. Currently those logs are kept in excel sheets files which are project specific, and

 ISBN: 978-1-68558-021-6

the company does not do any analysis on the logs. They have logs going back more than 10 years.

The test engineers at SIT, do know the value of data analysis in terms adding value to their operations. They require the following:

- Ability to log information centrally and follow up (can be achieved by any online system).
- Traceability (of the equipment and reports):
  - Keep track of reported errors.
  - Keep track of equipment and its maintenance history.
- Trends: see what patterns or lack of them appears in the event logs. Natural Language Processing (NLP) analysis of the event logs shared with us has similarities with other partners case studies [5]).

B. Proposed Approach Design

Figure 1 shows the component diagrams of the current environment with our proposed components. In the first step a prototype system is developed for them to log the unexpected events and related data in a convenient manner i.e., enhance their ability to log information centrally.

In terms of trends, basic NLP analysis of the event logs is performed, and the functionality is added to the prototype.

In terms of traceability, the testing engineers need to be able to keep track of the progress on the reported problems by product responsible departments or suppliers. For this, the researchers suggest a semantic ontology driven equipment register based on the Digital Design Basis initiative to quickly search the product related information and to update the event log with relevant data, instead of current manual look up at different systems, which is time consuming and cumbersome.

Testing engineers and other departments require traceability of their equipment. The problem is that the equipment may be at different departments or at different subsidiaries which are dispersed all over the globe. Apart from that, the equipment may have gone through maintenance-based changes and may have alter in specification as the years have gone by. The ontology base register has the potential to streamline the process of crawling through different systems and find a match of equipment based on the ontology properties [30]–[32]. Connected to the event log system it could also help test engineers in quickly finding the specification and functionality of the different equipment and see product relevant information in a dynamic manner, i.e., due to the machine-readable nature of the data it can be used to transform the information to be curated in any UI design, e.g., an excel style sheet or a graphical representation of the equipment with overlapping information i.e. (digital twins of various fidelities).

C. Description of prototype and interviews with test engineers

To facilitate the understanding of the information needs of maintenance engineers. We developed a semi-functional prototype and use this to conduct interviews with SIT test engineering team. During our sessions we identified three basic requirements of SIT engineers namely logging, traceability and identifying trends (see Section III-A for details).

For event logging we propose an online system which shows the event logs in a grid (see Figure 2). The reason behind is that it is similar to an excel file so that, the engineers are comfortable using it as they are very used to, using excel files. In addition to centrally log all the events the users have functionalities like sorting and filtering. They can also have additional controls on updating the values of certain fields so, for example, only a test manager could edit certain values (see Figure 2).

For identifying trends, the prototype allows users to perform analytics on the event logs, e.g., identify what equipment or tag number is appearing most in the logs indicating a trend of equipment which causes errors. Additionally, it also lets user do NLP analysis of the event logs, e.g., to identify repeatedly appearing phrases such as “need to be calibrated” to identify trends and to communicate them with product designers and other responsible parties (see Figure 3). Details of the NLP analysis of the event logs are the subject of another publication.

For traceability, the test engineers require information from different systems. The information they usually require are product specification, and equipment maintenance history and location etc. currently engineers have to search for this information in several different systems and it is a very time consuming and a tedious task. In the prototype we suggested connections with those systems and to allow the engineers to get that information from within the prototype.

The prototype also showed some alternative ways of representing information to engineers, e.g., graphical representation of the equipment with equipment specific information overlayed in a detailed-on demand manner. The idea is that user could switch between excel like list and other representations to better suite their exploration and communication needs. E.g., communicating an error with equipment image and overlayed information to provide a richer experience when communicating events with different departments and stakeholders (see Figure 4).

The prototype was well received. Participants really liked the concept. During our sessions and discussions, users wanted to centrally view the events and the required information. They really liked the simplicity of the concept. Interview participants also liked the analytics features and saw a lot of potential in the NLP analysis of the logs and the ability to identify the trends and patterns occurring during different equipment testing.

Participants reiterated our observations from our co-creation sessions and observation of the team meetings, that looking for information in different systems is indeed a very time-consuming task and a drag to their productivity and user experience. They really liked the idea of linking the different information and presenting it in one interface. They also saw the potential of using digital twin representation of different fidelity and product specific information overlayed, instead of simply linking to the product documentation.

When asked about what is missing in the prototype and what functionality would further enhance their experience: the reply

 ISBN: 978-1-68558-021-6

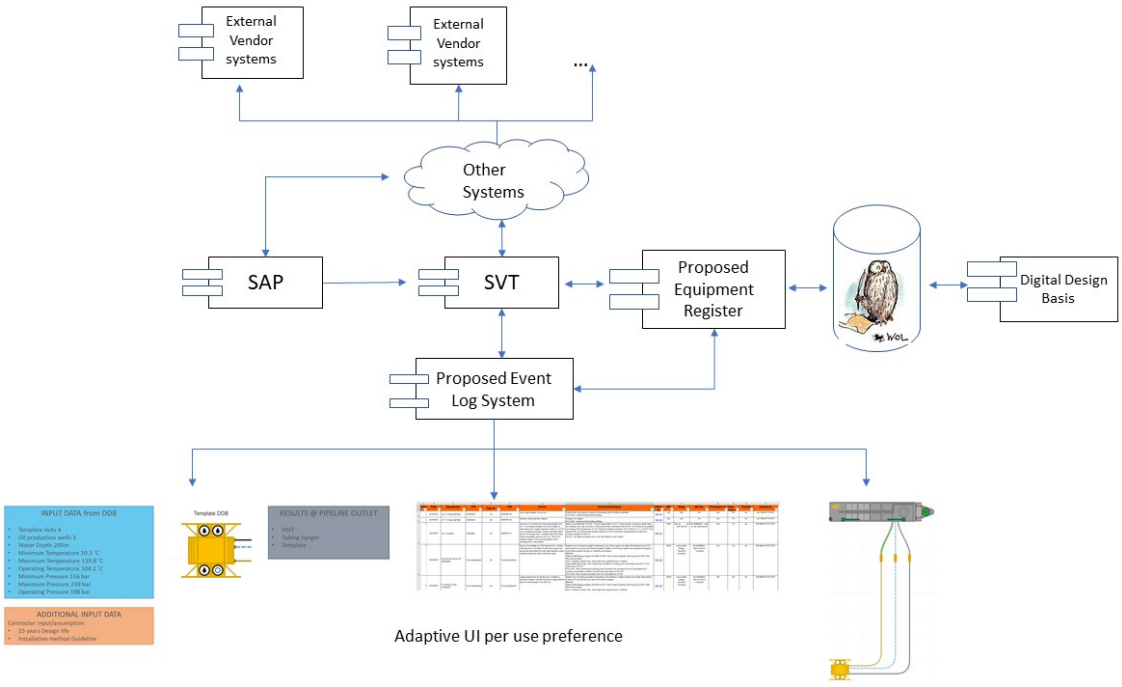

Figure. 1: Components Diagram overall approach (proposed).

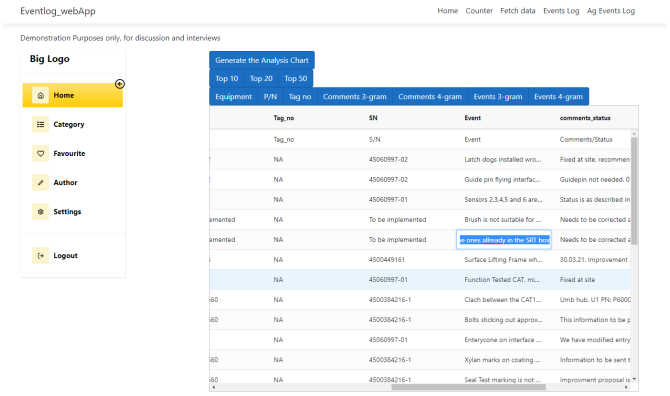

Figure. 2: Viewing and updating event logs prototype.

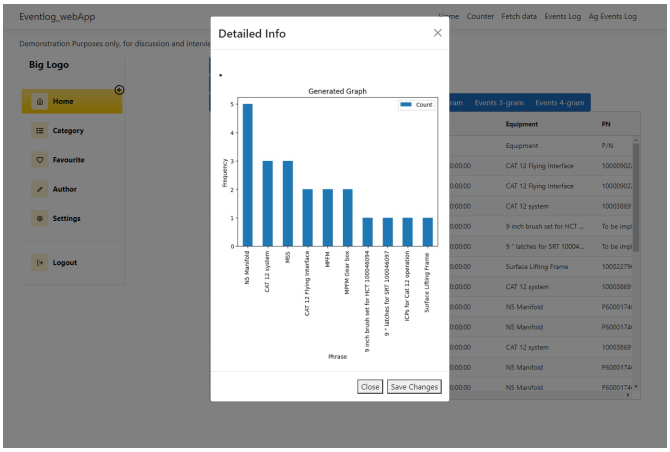

Figure. 3: Trend identification by NLP analysis.

was that they would prefer to have the information in different layers i.e., they wanted better control on the amount of detail they get about the equipment. It was also mentioned that different stakeholder may need access to different information and the need to control who can edit and view the information. Finally, the need to have easy to use system was reiterated and a need to standardized way to enter and edit the logs consistently across different teams and projects.

IV. Discussion

The co-creation sessions and interviews demonstrated the need for considering the user perspective to allow users to easily explore the information in a complex systems engineering environments. The case study shows that in large organizations despite having most of the infrastructure in place, users in

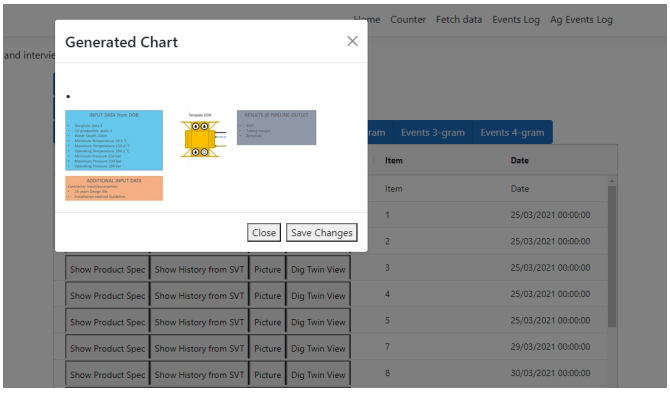

Figure. 4: alternative representation of equipment information.

certain functions still end up using sub-optimal exploration tools during their routine operations. This creates a big drag on their productivity. It is the sentiment expressed by the test engineers during our sessions.

Test engineers require a central place to search for the required information and the ability to make changes in a user-friendly manner. They also desire more control in viewing and editing of certain data and the amount of detail they get when looking at information i.e., detail on demand representation. Linking information from different systems and representing to them in easy to view manner is helpful for them. They also require standardized way to log information that can be achieve by implementing controlled vocabularies in the input interfaces.

The key takeaway from our sessions is that in complex systems engineering environments; the software tools are often designed for the content producers and not for the different users (consumers) of that information. There are also not easy to user for non-daily users to explore the available information. It makes it difficult for them to use, as it becomes time consuming for them to perform their routine tasks. It leads users to stick to old tools like excel files and not being able to utilize the benefits of technological developments.

V. Conclusion

The presented case study demonstrated that while large enterprises have many different systems to collect information and generate documentation and they may also have the necessary infrastructure to allow efficient use of the data. Users still end up not having a good experience with such tools and it creates hurdles in accessing the required information in an efficient manner. It discourages the adaptation of new systems and causes lost opportunities in terms of learning insights from the available information which could enhance daily operation or allow design decisions to avoid problems occurring in the future.

The case study demonstrates that by understanding the user needs, minor modifications to existing infrastructure can create a large impact in information exploration and utilization and a better experience for engineers, management and other stakeholders and enhance the productivity of teams, equipment testing and design.

In the future, we aim to extend the study to other teams at the partner organization to understand the similarities in information needs of different users in a complex systems engineering environment. We also aim to extend this to other partner organizations of the H-SEIF2[1]. consortium to come up with a design specification which could help in creating better interactive systems for enterprise users.

Acknowledgement

This research is part of the H-SEIF 2 project and is funded by the Research Council of Norway through the innovation project (IPN project no.: 317862).

[1]https://www.usn.no/hseif last accessed on 20-06-2022

References

[1] S. Engen, K. Falk, and G. Muller, "The need for systems awareness to support early-phase decision-making—a study from the norwegian energy industry," *Systems*, vol. 9, no. 3, 2021.

[2] H. Kannan, "Formal reasoning of knowledge in systems engineering through epistemic modal logic," *Systems Engineering*, vol. 24, no. 1, pp. 3–16, 2021.

[3] F. A. Salim, H. B. Ali, and K. Falk, "Towards a Data Processing Framework for Enhanced User Centric System Engineering ," in *PROCEEDINGS OF The Society for Design and Process Science, Workshop on Smart Pervasive Computing*, 2021, pp. 15–18.

[4] Qlik and Accenture, "The Human Impact of Data Literacy," Tech. Rep., 2020. [Online]. Available: https://thedataliteracyproject.org/humanimpact

[5] H. B. Ali and F. A. Salim, "Transferring Tacit Knowledge into Explicit: A Case Study in a Fully (Semi) Automated Parking Garage," in *PROCEEDINGS OF The Society for Design and Process Science*, 2021, pp. 11–14.

[6] B. A. Delicado, A. Salado, and R. Mompó, "Conceptualization of a T-Shaped engineering competency model in collaborative organizational settings: Problem and status in the Spanish aircraft industry," *Systems Engineering*, vol. 21, no. 6, pp. 534–554, 2018.

[7] K. Falk, A. K. Kamara, E. P. Brathen, K. Helle, P. T. Moe, and S. Kokkula, "Digitizing the Maintenance Documentation; A System of Systems in Oil and Gas Industry," *SOSE 2020 - IEEE 15th International Conference of System of Systems Engineering, Proceedings*, pp. 493–499, 2020.

[8] Juzgado P.D. Borches, "A3 Architecture overviews. A tool for effective communication in product evolution," Ph.D. dissertation, University of Twente, 2010.

[9] R. Technology, "AI / ML Annual Research Report 2022," 2022.

[10] J. Bickford, D. L. Van Bossuyt, P. Beery, and A. Pollman, "Operationalizing digital twins through model-based systems engineering methods," *Systems Engineering*, vol. 23, no. 6, pp. 724–750, 2020.

[11] M. Kjørstad, K. Falk, and G. Muller, "Exploring a Co-Creative Problem Solving Toolbox in the Context of Norwegian High-Tech Industry," *IEEE Systems*, pp. 1–11, 2020.

[12] D. H. Shin, "Demystifying big data: Anatomy of big data developmental process," *Telecommunications Policy*, vol. 40, no. 9, pp. 837–854, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.telpol.2015.03.007

[13] C. Cobârzan, K. Schoeffmann, W. Bailer, W. Hürst, A. Blažek, J. Lokoč, S. Vrochidis, K. U. Barthel, and L. Rossetto, "Interactive video search tools: a detailed analysis of the video browser showdown 2015," *Multimedia Tools and Applications*, vol. 76, no. 4, pp. 5539–5571, 2017.

[14] F. A. Salim, F. Haider, S. Luz, and O. Conlan, "Automatic transformation of a video using multimodal information for an engaging exploration experience," *Applied Sciences (Switzerland)*, vol. 10, no. 9, 2020.

[15] A. Jackson, J. Lin, I. Milligan, and N. Ruest, "Desiderata for exploratory search interfaces to Web archives in support of scholarly activities," *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, vol. 2016-Septe, pp. 103–106, 2016.

[16] J. Moreno-Schneider, A. Srivastava, P. Bourgonje, D. Wabnitz, and G. Rehm, "Semantic Storytelling, Cross-lingual Event Detection and other Semantic Services for a Newsroom Content Curation Dashboard," *Association for Computational Linguistics (ACL).*, pp. 68–73, 2018.

[17] F. A. Salim, F. Haider, D. Postma, R. van Delden, D. Reidsma, S. Luz, and B.-J. van Beijnum, "Towards Automatic Modeling of Volleyball Players' Behavior for Analysis, Feedback, and Hybrid Training," *Journal for the Measurement of Physical Behaviour*, vol. 3, no. 4, pp. 323–330, 2020.

[18] B. Yousuf and O. Conlan, "VisEN: Motivating learner engagement through explorable visual narratives," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9307, pp. 367–380, 2015.

[19] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, and S. Carpendale, "State of the Art of Sports Data Visualization," *Computer Graphics Forum*, vol. 37, no. 3, pp. 663–686, 2018.

[20] T. Polk, J. Yang, Y. Hu, and Y. Zhao, "TenniVis: Visualization for tennis match analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2339–2348, 2014.

[21] L. Shao, D. Sacha, B. Neldner, M. Stein, and T. Schreck, "Visual-interactive search for soccer trajectories to identify interesting game

situations," *IS and T International Symposium on Electronic Imaging Science and Technology*, vol. 0, pp. 1–10, 2016.

[22] J. A. Gopsill, H. C. McAlpine, and B. J. Hicks, "Supporting engineering design communication using a custom-built social media tool - PartBook," *Advanced Engineering Informatics*, vol. 29, no. 3, pp. 523–548, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.aei.2015.04.008

[23] M. Guntveit, M. Kjrstad, and B. Sevaldson, "Early Validation of Stakeholder Needs by Applying Co-creation Sessions," *INCOSE International Symposium*, vol. 30, no. 1, pp. 1398–1415, 2020.

[24] G. Gravier, M. Ragot, A. Laurent, R. Bois, G. Jadi, E. Jamet, and L. Monceaux, "Shaping-Up Multimedia Analytics: Needs and Expectations of Media Professionals," in *The 22nd International Conference on Multimedia Modelling*, vol. 9516, Miami, 2016, pp. 303–314. [Online]. Available: http://link.springer.com/10.1007/978-3-319-27671-7

[25] M. Haesen, J. Meskens, K. Luyten, K. Coninx, J. H. Becker, T. Tuytelaars, G. J. Poulisse, P. T. Pham, and M. F. Moens, "Finding a needle in a haystack: An interactive video archive explorer for professional video searchers," *Multimedia Tools and Applications*, vol. 63, no. 2, pp. 331–356, 2013.

[26] F. A. Salim, F. Haider, O. Conlan, and S. Luz, "An approach for exploring a video via multimodal feature extraction and user interactions," *Journal on Multimodal User Interfaces*, no. October 2017, 2018. [Online]. Available: http://link.springer.com/10.1007/s12193-018-0268-0

[27] L. Koesten, K. Gregory, P. Groth, and E. Simperl, "Talking datasets – Understanding data sensemaking behaviours," *International Journal of Human Computer Studies*, vol. 146, no. October 2020, p. 102562, 2021. [Online]. Available: https://doi.org/10.1016/j.ijhcs.2020.102562

[28] J. Pinto, K. Falk, and M. Kjørstad, "Inclusion of human values in the specification of systems: bridging design and systems engineering," *INCOSE International Symposium*, vol. 29, no. 1, pp. 284–300, 2019.

[29] C. Potts, "Software-engineering research revisited," *IEEE Software*, vol. 10, pp. 19–28, 1993.

[30] Y. A. Akter and M. A. Rahman, "Extracting RDF Triples from Raw Text," *1st International Conference on Advances in Science, Engineering and Robotics Technology 2019, ICASERT 2019*, vol. 2019, no. Icasert, pp. 31–34, 2019.

[31] W. Baes, F. Portet, H. Mirisaee, and C. Labbé, "RDF triples extraction from company web pages: Comparison of state-of-the-art deep models," *CEUR Workshop Proceedings*, vol. 2708, 2020.

[32] P. Exner and P. Nugues, "Entity extraction: From unstructured text to dbpedia rdf triples," *CEUR Workshop Proceedings*, vol. 906, pp. 58–69, 2012.

 ISBN: 978-1-68558-021-6