Supporting Data Analytics in Manufacturing with a Digital Assistant

Stefan Wellsandt[1], Mina Foosherian[1], Katerina Lepenioti[2], Mattheos Fikardos[2], Gregoris Mentzas[2], Klaus-Dieter Thoben[3]

[1] BIBA – Bremer Institut für Produktion und Logistik GmbH, Hochschulring 20, 28359 Bremen, Germany
{wel, fos}@biba.uni-bremen.de
[2] Information Management Unit (IMU), Institute of Communication and Computer Systems (ICCS), National Technical University of Athens (NTUA), Athens, Greece
{albous, gmentzas}@mail.ntua.gr
[3] University of Bremen, Faculty of Production Engineering, Badgasteinerstr. 1, 28359 Bremen, Germany
tho@biba.uni-bremen.de

Abstract. The shortage of skilled workers is a barrier to applying data analytics. Augmented analytics is an approach to lower it by using machine learning to automate related activities and natural language applications to assist less-skilled employees. Public information about augmented analytics case studies in manufacturing is hardly available. Therefore, this article presents a related case study from the white goods industry. It focuses on a quality test lab in a production line where workers use a digital assistant prototype to interact with descriptive and predictive data analytics. This article derives a framework from this case study to organize how an assistant could augment analytics. The framework has five areas: training data modification, model training, starting an analysis, retrieval of results, and decision support. The latter is relevant to the other four areas and includes, for instance, suggesting options to customize analytics. Four scenarios of different complexity concretize the framework's areas. Finally, this article outlines four questions for future research.

Keywords: Augmented Analytics, Digital Intelligent Assistant, Voice Assistant, Natural Language Understanding, Assistance Technology.

1 Introduction

When producers adopt and practice data analytics, they need skilled employees familiar with its approaches, activities, and tools. Often, this includes machine-learning (ML) expertise. The employees also need a deep understanding of the data's context, including the target audience that will use the results at work. A critical barrier that slows down the adoption of data analytics in production is the shortage of skilled employees [1]. Drivers for this situation are that the education of data analysts takes a long time, and more companies want to employ these experts, which increases the demand.

Technology providers aim to reduce data analytics' skill floor and the education time by automating activities and supporting employees who work with related tools [2–4]. **Augmented analytics** a) uses technologies such as ML to automate data analysis activities and b) assists employees with data preparation, processing, and understanding results [5]. The latter includes the application of artificial intelligence (AI) and natural language processing (NLP) in chatbots and voice assistants. Such tools augment human cognitive capabilities, so less-skilled employees can use data analytics tools.

Public information about augmented analytics in manufacturing is sparse – especially case studies with sufficient depth. Consequently, advancing research in this field is challenging because researchers cannot build on previous work.

This article provides a *case study* and a related *framework* for augmented analytics in manufacturing. It includes the architecture for a voice-enabled digital intelligent assistant (DIA) and dialogs outlining how users interact with data analytics through it. The remainder of this article has four sections. Section 2 summarizes aspects of augmented analytics to scope this paper, while section 3 presents the case study and example conversations. Section 4 introduces a framework to augment different analytics activities. Finally, section 5 concludes the paper and outlines research directions.

2 Related Work

Augmented analytics is a term brought up by Gartner to summarize the use of AI in analytics [1]. It is a broad concept proposing ML and conversational user interfaces (CUI) to get insight from data. CUIs allow users to interact with software through natural language and are an essential characteristic of a DIA [6].

2.1 Automated machine learning in manufacturing analytics

Automated machine learning (AutoML) is a sub-topic of augmented analytics. It focuses on the progressive automation of manual ML tasks [7] and aims at minimizing human intervention to save time and make analytics accessible to non-experts.

The application of AutoML in manufacturing concerns, for instance, the prediction of lead times and process quality. Bender, Trat, and Ovtcharova [8, 9] applied AutoML to predict lead times in two small and medium-sized enterprises. They found that it created superior prediction models in one case. The authors reported mixed results for the other case because the error rates were too high for some production steps. Besides, the applied AutoML approach could not support highly labor-intensive tasks, such as data understanding, transformation, filtering, pre-processing, and feature engineering. Denkena et al. [10] used AutoML to optimize shape error prediction in milling processes. They could significantly decrease prediction errors with AutoML, which is a substantial advantage in their targeted application area. However, the authors reported that applying AutoML requires much expertise in ML compared to manually making all decisions. The two example applications above do not provide clear answers on how automation lowers the skill floor of analytics tasks.

2.2 Digital assistants in manufacturing analytics

The application of DIAs and similar systems in manufacturing is a comparably new research area with few articles on the topic and even less related to analytics. The following two articles cover assistants in combination with analytics.

Abner et al. [11] describe a software robot that provides descriptions, diagnostic information, predictions, and prescriptions. It interacts with users via chatting and providing URLs to dashboards with multi-dimensional information, such as time series. Listl, Fischer, and Weyrich [12] applied a chatbot as an alternative user interface for a plant simulation tool. Their prototype lets users adjust simulation parameters, model topology, and schedules via chatting.

Besides academic case studies, some commercial assistants support analytics in manufacturing. Software, such as Oracle Digital Assistant [13], SAP Conversational AI [14], and SPIX [15], ground their features on connectors to existing business software, including analytics services. We did not identify specific case studies about assistants and analytics among commercial tool providers, but feature descriptions contain cues. For instance, SAP has augmented analytics and conversational AI (digital assistant) services. The former offers a conversational user interface to query descriptive analytics results [16]. A demo video from 2020 [17] presents the following query text: "*show order value by product by customer segment for product […] for previous year*". Such natural language queries are helpful if analytics results are complex and navigating through a dashboard overloads the user.

The case studies above demonstrate that users can query results and parameterize analyses through natural language. However, they remain rather unspecific about how they augment the analytics process. They do not provide deeper insight into the type of support they provide to the users. Without this information, it is hard to understand how and in which areas augmentation can reduce the skill floor of analytics.

This paper contributes preliminary findings from a DIA case study in manufacturing and proposes an augmentation framework to address the issue above.

3 Case study

This article bases on a case study in the white goods industry where an assistant supports employees during quality tests in a laboratory at the end of the production line. Lab operators perform various manual tests to identify deviations from the expected product behavior or characteristics (e.g., correctly printed and placed labels). Lab supervisors assess deviations and use analysis results to decide how to manage them.

3.1 Context and process

Product and process quality concerns most manufacturing firms because negative consequences only show in production or worse when the customer returns with a complaint [18]. Producers must prevent adverse effects of quality flaws before they become evident in the product's use [18, 19] or even before they cause inefficiencies in the production process.

One approach to minimizing production quality issues is applying predictive quality analytics. It extracts valuable insights from various data sources by determining patterns, revealing correlations between products and defects, and predicting future outcomes (e.g., product defects and fault localization) [18–21]. Data sources include measurements performed by technical systems and information that employees create, for instance, through reports.

3.2 Assistant architecture

Access to predictive quality analytics typically requires expertise, as outlined in the introduction. Therefore, we developed a DIA to support, amongst other tasks, this analytics process. It uses open-source technology to transcribe the human voice, understand the resulting natural language texts, decide how to respond, and return responses in synthetic speech. Our augmented manufacturing analytics (AMA) assistant combines the digital assistant framework Mycroft and the chatbot framework Rasa. A custom Mycroft skill manages the exchange of messages between these two frameworks. Users interact with the assistant through a tablet with a custom App providing different conversation modes and rich-media contents (e.g., buttons, images, and videos). The Rasa chatbot contains a custom dialog model and actions to transform user utterances into formalized queries for the quality analytics component. Besides, it builds natural language responses out of the computer-readable analysis results.

3.3 User story and augmentation examples

Test laboratory supervisors inspect defects and search for root causes. They need quality analytics to better understand which product types and how many items had a specific defect. Besides, they need predictions for the quality to become aware of emerging issues that require inspection or sophisticated root cause analysis.

The assistant can recognize and manage several intents related to augmented analytics. Most involve a specific product with a unique identifier, such as a serial number or stock-keeping unit (SKU) code. The user scans the item's matrix or barcode with the tablet to add this identifier in a conversation. Scanning proved to be the most viable option because typing is slow, and speech transcription is error-prone. However, the assistant supports the latter two to help in specific situations (e.g., scanning issues and hands-free interaction).

Once the assistant knows which product the user is interested in, the user can request follow-up information. For example, the assistant could provide statistics and performance indicators as a text response or link to an online dashboard. Each time, the assistant triggers an analysis process according to the user's request. Besides, the assistant can make predictions on request. In some situations, the assistant takes the initiative after a trigger event. Measurements or automated software processes can create these events. For example, when a new product arrives in the test lab, the assistant communicates which defect groups the user can expect based on a prediction. **Fig. 1** summarizes augmented analytics conversations with our voice assistant. Some responses contain encoded identifiers that future versions will decode.

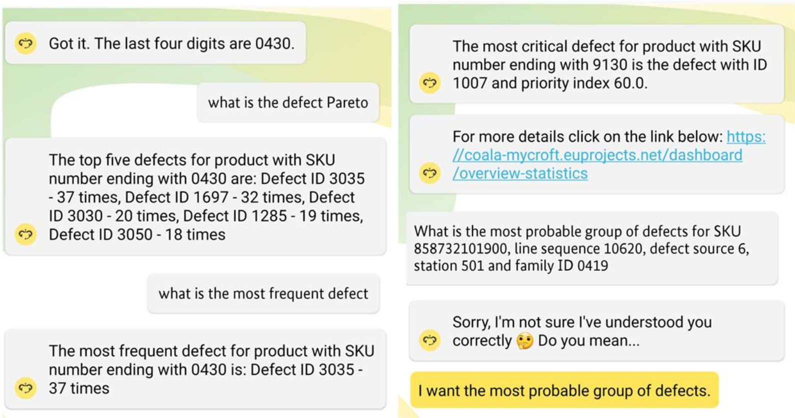

Fig. 1. Product registration, requesting descriptive analytics for an item, and predictions

4 Analytics support framework

The case study above outlines how a DIA could augment employees that interact with data analytics. We summarize our preliminary findings and the planned features in the framework illustrated in **Fig. 2**.

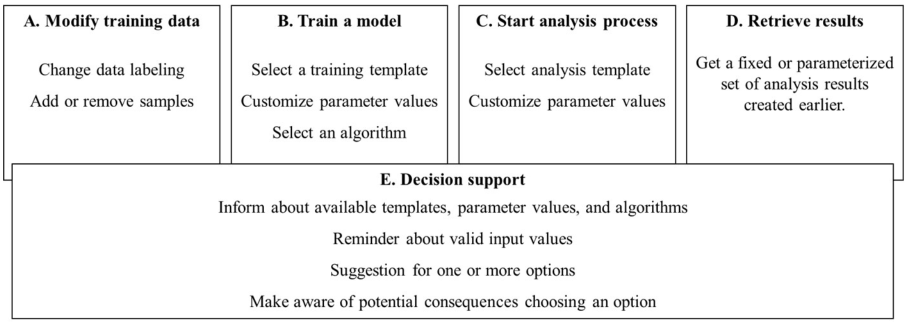

Fig. 2. Augmented analytics framework for Digital Intelligent Assistants in manufacturing

The framework above has five areas with one or more augmentable activities. Areas A, B, and C represent typical steps in data analytics. Area D concerns retrieving analysis results and, therefore, is not a typical analytics task but necessary to access information. Areas A to D represent the logical sequence of a data analytics process using machine learning. Finally, area E focuses on decision support for the augmentable activities in all other areas – representing augmentation in the narrow sense. Its activities help users understand decision options, valid input values, and potential consequences of choosing options. The areas A-D refer to augmentation in the broader sense. Here, users may interact with analytics via natural language, granting intuitive access to information.

In the framework, the smallest possible augmented analytics scenario is the retrieval of existing results. Many users in manufacturing may only retrieve existing analytics information to perform a task. Users can request pre-defined information (i.e., a named view), such as a shortlist of ranked defects, or formulate a parametrized request. The latter uses pre-defined parameters like dates, times, locations, and other product, process, and environmental characteristics. An example scenario is requesting a Pareto diagram (e.g., top three defects) to support a decision.

More complex scenarios include augmentation of the analysis process where the user can trigger pre-defined analysis templates or analyses with customized parameters. Machine learning is not part of these scenarios. An example of an analysis template is a system diagnosis to identify all faulty components. Users may customize a system diagnosis by mentioning sub-systems, locations, and time frames in their natural language request.

Even more complex scenarios include machine learning. In these cases, augmentation can cover training data modification, model training, or both. Changes to the training data are relevant for hybrid-augmented intelligence systems [22]. In such systems, humans and AI support each other's learning and improvement. For example, users could inspect new quality issue reports through a DIA and instruct it to add specific ones as training data. The next model training would include these changes.

The *most complex augmentation scenarios* include decision support for one or more activities in areas A to D. The assistant could **inform** users about available options (e.g., parameters, templates, and algorithms) at the beginning of a task. Second, it could **remind** users about valid input values after providing invalid ones. Valid means that values are allowed because they relate to existing information such as components, defect groups, and locations. This support could be a response to faulty transcriptions or when users accidentally provide invalid input. Third, the DIA could **suggest** adding or removing a sample, selecting a template, or using a specific parameter value. The specific suggestion is the result of a deterministic or probabilistic model. An example is a user who requests a prediction for quality incidents by a given date, such as tomorrow, next week, or the month's end. If the prediction accuracy for the next week is below a threshold, the assistant could suggest using the timeframe just above the minimum allowed accuracy. Fourth, the assistant could **make the user aware** of the potential consequences of choosing a template, algorithm, or parameter value. For example, if the user removes training data samples via the assistant, the model training will likely produce an overfitting model. A second example is an assistant that tests training data changes for bias and warns the user if it finds them.

5 Conclusion

The case study and the framework above provide concrete examples of how a DIA could help employees perform descriptive and predictive analytics. We did not realize all scenarios in the framework with an assistant yet. Our current prototype focuses on areas C, D, and E, and we expect it will provide additional results for more profound theoretical and applied research. Future research could cover the following questions:

- *How deep should the assistance be to be most efficient*? Sophisticated assistance requires more development effort, which is costly. Besides, it leads to more complex software that is more difficult to maintain and error-prone because the assistant can be confused easier.

- *How could the assistant explain its predictions to users*? The capability to explain predictions (and other results based on machine learning) is an essential aspect of trustworthy AI. Explanations can be quite technical and thus difficult to understand by non-experts.
- *How can developers transfer dialog models for augmented analytics from one case to another*? There are many training data formats, machine learning algorithms, and parameters. Transferable dialog models could significantly reduce the effort to build DIAs for augmented analytics.
- *How does the augmentation affect the skill floor of using data analytics in manufacturing?* Literature suggests that augmented analytics helps less-skilled employees perform analytics, but it seldomly backs this claim with evidence. Since augmented analytics solutions tend to become more available, researchers could use them for systematic evaluations.

Our future work will focus on the assistant's evaluation in factory environments and the extension of its features to cover more areas in the framework above.

Acknowledgments

This work is funded by the European Union's Horizon 2020 research and innovation program via the project COALA, "COgnitive Assisted agile manufacturing for a LAbor force supported by trustworthy Artificial Intelligence" (Grant agreement No 957296).

References

1. Prat N (2019) Augmented Analytics. Bus Inf Syst Eng 61:375–380. https://doi.org/10.1007/s12599-019-00589-0
2. SAP (2022) Augmented and predictive analytics features. https://www.sap.com/germany/products/cloud-analytics/features/augmented-analytics.html. Accessed 27 Jun 2022
3. Avidon E, Kronz A (2022) Gartner: Augmented analytics, ecosystem for BI now key. https://www.techtarget.com/searchbusinessanalytics/news/252515720/Gartner-Augmented-analytics-ecosystem-for-BI-now-key. Accessed 27 Jun 2022
4. Tableau Software (2022) Augmented analytics explained: definition, use cases, benefits, features, and more. https://www.tableau.com/learn/articles/augmented-analytics. Accessed 27 Jun 2022
5. Gartner (2022) Gartner Glossary: Augmented Analytics. https://www.gartner.com/en/information-technology/glossary/augmented-analytics. Accessed 12 Apr 2022
6. Wellsandt S, Hribernik K, Thoben K-D (2021) Anatomy of a Digital Assistant. In: Dolgui A, Bernard A, Lemoine D et al. (eds) Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems, vol 633. Springer International Publishing, Cham, pp 321–330
7. Hutter F (2022) What is AutoML? https://www.automl.org/automl/. Accessed 13 Apr 2022

8. Bender J, Trat M, Ovtcharova J (2022) Benchmarking AutoML-Supported Lead Time Prediction. Procedia Computer Science 200:482–494. https://doi.org/10.1016/j.procs.2022.01.246
9. Bender J, Ovtcharova J (2021) Prototyping Machine-Learning-Supported Lead Time Prediction Using AutoML. Procedia Computer Science 180:649–655. https://doi.org/10.1016/j.procs.2021.01.287
10. Denkena B, Dittrich M-A, Lindauer M et al. (2020) Using AutoML to Optimize Shape Error Prediction in Milling Processes. SSRN Journal. https://doi.org/10.2139/ssrn.3724234
11. Abner B, Rabelo RJ, Zambiasi SP et al. (2020) Production Management as-a-Service: A Softbot Approach. In: Lalic B, Majstorovic V, Marjanovic U et al. (eds) Advances in Production Management Systems. Towards Smart and Digital Manufacturing, vol 592. Springer International Publishing, Cham, pp 19–30
12. Listl FG, Fischer J, Weyrich M (2021) Towards a Simulation-based Conversational Assistant for the Operation and Engineering of Production Plants. In: 2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). IEEE, pp 1–4
13. Oracle (2022) Oracle Digital Assistant for ERP and SCM. https://www.oracle.com/chatbots/digital-assistant-for-erp-scm/. Accessed 24 Feb 2022
14. SAP (2022) SAP Conversational AI. https://www.sap.com/germany/products/conversational-ai.html
15. Spix Industry (2022) Spix: Industry-Ready Chatbot Technology. https://www.spix-industry.com/en/products-chatbot-for-industrial-software-development-spix/. Accessed 24 Feb 2022
16. SAP (2022) Augmented Analytics (Smart Features). https://www.sap.com/products/cloud-analytics/features.html. Accessed 14 Apr 2022
17. SAP Analytics (2020) Augmented Analytics in SAP Analytics Cloud
18. Nalbach O, Linn C, Derouet M et al. (2018) Predictive Quality: Towards a New Understanding of Quality Assurance Using Machine Learning Tools. In: Abramowicz W, Paschke A (eds) Business Information Systems, vol 320. Springer International Publishing, Cham, pp 30–42
19. Berger D, Zaiß M, Lanza G et al. (2018) Predictive quality control of hybrid metal-CFRP components using information fusion. Prod Eng Res Devel 12:161–172. https://doi.org/10.1007/s11740-018-0816-1
20. Bai Y, Sun Z, Deng J et al. (2018) Manufacturing Quality Prediction Using Intelligent Learning Approaches: A Comparative Study. Sustainability 10:85. https://doi.org/10.3390/su10010085
21. Gittler T, Relea E, Corti D et al. (2019) Towards predictive quality management in assembly systems with low quality low quantity data – a methodological approach. Procedia CIRP 79:125–130. https://doi.org/10.1016/j.procir.2019.02.026
22. Wellsandt S, Klein K, Hribernik K et al. (2022) Hybrid-augmented intelligence in predictive maintenance with digital intelligent assistants. Annual Reviews in Control. https://doi.org/10.1016/j.arcontrol.2022.04.001