Anatomy of a digital assistant

Stefan Wellsandt[1], Karl Hribernik[1], and Klaus-Dieter Thoben[1,2]

[1] BIBA - Bremer Institut für Produktion und Logistik GmbH at the University of Bremen, Hochschulring 20, 28359 Bremen, Germany
{wel, hri}@biba.uni-bremen.de
[2] University of Bremen, Faculty of Production Engineering, Badgasteiner Straße 1, 28359 Bremen, Germany
tho@biba.uni-bremen.de

Abstract. Why is it helpful to have a digital assistant? This question's answer is not simple nor easy to find because artificial intelligence (AI) assistants, such as Alexa, Bixbi, or Siri, are amorphous compound technology and multi-purpose tools. Most of an assistant's components provide unique benefits on their own. Mobility, voice interaction, the delegation of administrative tasks, and rapid data analysis are typical benefits, but they are not exclusive to digital assistants. Understanding an assistant's benefits helps assistant designers and decision-makers who need to assess whether an assistant is a suitable workforce enhancement tool. Academic literature often describes an assistant's benefits superficially. This article presents an overview of a preliminary catalog of these benefits in manufacturing. It covers central access, customization, delegation and guidance, eyes-free and hands-free interactions, mobile assistance, the support of multiple interface types, permanent accessibility, and speed. We conclude that the cataloged benefits need more evidence, preferably created during experiments in natural manufacturing environments, to explore and experience the factors that determine the use of a digital assistant. These factors include trust in AI systems, impacts on teams and individuals, training and education, and capabilities of open and closed technologies. Disadvantages, limitations, and risks concern reduced worker autonomy, constrained language understanding, increased dependency on software, and harmful exploitation.

Keywords: Virtual Assistant, Conversational AI, Voice Interface, Industry 4.0, Smart Manufacturing

1 Introduction

Why is it helpful to have a digital assistant? This question's answer is not simple nor easy to find because artificial intelligence (AI) assistants, such as Alexa, Cortana, Bixbi, or Siri, are *amorphous compound technology* and *multi-purpose tools*. Most of an assistant's components provide unique benefits on their own. Mobility, voice interaction, the delegation of tasks, and rapid data analysis are typical benefits, but they are not exclusive to digital assistants. Therefore, developing an assistant is challenging because customers can object that they do not benefit from all components immediately. For example, if an organization wants to delegate repetitive tasks to artificial intelligence, its solution may not need a voice interface or run on a mobile device. An assistant's compound nature may overfit in this case.

Answering the question above begins with an assistant's anatomy. Its purpose is to identify the assistant's components, functions, and benefits. Assistant designers and customers can jointly use this knowledge to assess if a digital assistant is a helpful tool. The starting point for this anatomy can be the *business domain*. A digital assistant for employees will likely integrate deeply into the company's information systems to be as helpful as possible. Therefore, such an assistant would be expensive to develop and deploy, and due to the novelty of its technology, it may fail to be helpful initially.

Besides, a failing assistant is often more costly for a business than a consumer. Disruptions of deeply integrated software can propagate and affect many people, machines, processes, and even collaborating organizations. Therefore, businesses are typically more skeptical about digital assistants.

Among the business domain, *manufacturing* is highly relevant for digital assistants. In 2017, the European Union's manufacturing sector had around 2 million enterprises, employed 28.5 million persons, and generated 1,820 billion Euros value-added [1]. The future workforce in the EU will be smaller due to demographic changes, diverse due to immigration, and threatened to be replaced by automation solutions [2]. Enhancing the remaining workforce's skills and capabilities through artificial intelligence offers a way out of this challenging situation. It allows companies to employ people they would otherwise not hire. At the same time, digital assistants provide several opportunities to save employees' time and, consequently, contribute to increasing work efficiency.

This paper aims to summarize the benefits of business-focused digital assistants in manufacturing. Its results may help assistant designers and decision-makers in companies assessing whether an assistant is a suitable tool at work. The remainder of this paper has four sections. Section 2 presents conversational and technology-based agents, digital assistant types, and other related work. Section 3 describes a use case that outlines challenges and problems in manufacturing environments. Section 4 presents the identified benefits that digital assistants could provide, while Section 5 concludes the results and suggests future work.

2 Related work

Academic literature about digital assistants is extensive and heterogeneous. The Scopus database finds more than 12,000 entries that contain digital assistant, virtual assistant, conversational agent, or software robot in the abstract, title, or keywords. These entries belong to computer science (6,500+), engineering (4,300+), medicine (2,400+), mathematics (1,500+), and social sciences (1300+). Scopus indicates 889 entries for 2020 and a peak of 1,332 entries for 2004 – which is the same year Scopus launched. The numbers above demonstrate that the scientific foundation for digital assistants is strong. Literature likely contains critical information to answer why these assistants are useful.

2.1 Conversational and technology-based agents

The main conceptual foundation for this article is the so-called **conversational agent** (CA). Literature mentions various terms describing software with similar functions. Some terms are synonymous with the CA, and others articulate a specific subtle distinction [3]. A CA is a dialog system embedded in personal technologies and devices [3]. It can support spoken and written natural language as input and output. The minimal architecture of a CA provides the following functions [4, 5]:

Speech recognition transcribes voice utterances – the result is a text. *Meaning extraction* uses the transcript to understand intents and entities, i.e., context. *Data queries* acquire additional information to fulfill the intent. The *Dialog manager* tracks the dialogue state and decides how to respond [6]. It uses the dialog state to interpret the final meaning of an utterance. *Response generation* selects or creates the specific response text. Finally, the *speech output* synthesizes voice from that text.

A CA supports task completion in real-time and develops knowledge about the user to act on their behalf. It does not focus on anthropomorphism nor representing a specific person and is, therefore, no human avatar [7].

The second key concept is the **technology-based agent**, a system that observes, interprets, decides, and learns to act upon its environment [8]. This agent interacts with humans and machines to achieve shared goals. Autonomy and capabilities determine how such agents support humans. Less autonomous agents can retrieve information for the user and automate tasks in decision-making. If a user delegates more responsibility to an agent, it can execute various tasks without further human involvement. An agent with limited capabilities uses static patterns to react to the inputs it receives from the environment. Sophisticated agents learn to operate in initially unknown environments and improve their capabilities.

2.2 Digital assistants

Digital assistants are socio-technical systems and an application class [9]. The former considers individual users, their goals, related tasks, and technology that processes data and allows human-computer interaction. The latter means the assistant is an orchestration of different components that provide specific functions.

Knote et al. [10] investigated 115 assistants and identified 31 design characteristics grouped into ten dimensions. Relevant characteristics include, for instance, communication mode, the direction of interaction, query input, response output, assistance domain, command complexity, adaptivity, and embodiment. Knote et al. [10] used them in a k-means clustering method to identify assistant categories. **Table 1** summarizes the results of their investigation.

Table 1. Assistant categories identified through empirical analysis [10]

Category	Features
Adaptive Voice Assistants	Speech, optical sensors, screen outputs, execute services upon request, general-purpose, adaptive, computer-generated human-like voice

Chatbot Assistants	Text, images, videos, screen interaction, task-oriented support, special purpose, present information to users, virtual characters
Embodied Virtual Assistants	Human-like, speech, screen outputs, virtual characters, special purpose, adjust to user autonomously, anthropomorphism
Passive Pervasive Assistants	Unobtrusive, collects data from sensors, initiates interaction with user, observes user's tasks and context, autonomous, special purpose
Natural Conversation Assistants	Speech, imitate human natural language interactions, execute services upon request, static behavior, understands compound commands

The characteristics and categories above provide a first direction for the benefits of digital assistants. *General-purpose* assistants support humans in tasks, such as information retrieval, calendar management, working with communication channels, and controlling smart devices. An assistant with *special-purpose* knowledge supports humans in fulfilling clearly defined tasks, such as filling an issue report. Some assistants are *adaptive* and learn from interactions. They can improve their language understanding and interpretation skills and adjust their behavior to improve user experience. Finally, *embodied* assistants use anthropomorphism to increase user acceptance – a typical feature is generating a human-like voice.

Maedche et al. [9] point out that machines, such as digital assistants, are ideal for repeatable and highly structured tasks. They are good at collecting, storing, and processing data and they make accurate predictions provided the environment is relatively stable. On the other hand, humans are better suited to solve abstract problems and manage fragmented information efficiently. Besides, they are more aware of context and can use intuition, empathy, and ethics in decision-making. Maedche et al. [9] also argue that the desired collaboration level between humans and digital assistants will occur within a continuum of autonomy. On one end, the human decides, and, on the other end, the assistant decides – the space in between covers all possible cases where AI assistants support humans and vice versa.

2.3 Operator 4.0 and software robots

Romero et al. [11] introduced the idea of the so-called Operator 4.0. This name summarizes an operator's new roles in Industry 4.0, Smart Manufacturing, and similar visions for future manufacturing. The *Smart Operator* role refers to operators that collaborate with digital assistants in the following areas: a) searching and retrieving from a digital library; b) scheduling and setting reminders for actions or events; c) store and visualize planning data to support humans in problem-solving; d) mobility and location assistance; e) interfacing with connected devices; f) detecting and diagnosing errors and problems; g) suggest troubleshooting tools and strategies; h) track operator and machine behavior to build predictive models; and i) notify about the need for proactive actions.

Rabelo, Romero and Zambiasi [12] demonstrated how operators could benefit from software robots, referred to as softbots. The authors recognized that softbots largely overlap with the software agent concept [13]. Abner et al. [14] argue that a softbot can respond to a user request directly, perform pre-defined scheduled tasks, and proactively communicate the status of fully delegated tasks. Their softbot provides descriptions, diagnostics, predictions, and prescriptions, which users access via sequential workflows or independently.

2.4 Studies on the use of voice assistants

Rzepka [15] investigated when and why users choose speech interaction over traditional user interfaces. She applied the Means-End Chain theory to understand individual decision-making processes. Her findings for a group of 31 users indicate that fundamental objectives are: faster task completion, easy access to the underlying system, joy of the interaction, minimize physical effort, and minimize deliberate thinking. The author points out that her study focused on the private use of digital assistants, and future work should address assistants' use in an organizational context.

3 Use case

We define a use case in this section to clarify challenges in manufacturing that digital assistants could address. The case outlines the work situation in production with a focus on information-intensive processes. Such processes may naturally benefit from digital assistants. The following description grounds on our experience from two assistant prototypes. Their focus is on predictive maintenance, augmented analytics, and cognitive assistance during on-the-job training in production. We added labels in parentheses to highlight challenges and connection points for digital assistance.

In production environments, workers operate, maintain, and repair machines to manufacture products. These persons work under *time pressure* to meet production performance goals (A). Performing these jobs requires *skills and competencies*. Workers receive vocational and on-the-job training to acquire these – this training is costly, takes weeks, months, or years, and limits the available workforce (B). Beyond their initial training, workers continuously learn and develop their skills and competencies (C).

Workers experience *fatigue* and require *recreation* – physically demanding work or highly repetitive tasks exhaust workers faster. Organizations can use a shift system to guarantee permanent operations. However, shift systems are costly because the organization must employ persons with similar skills, competencies, and knowledge (D).

Besides, reaching the performance goals in complex manufacturing processes typically requires workers to *access information* about machines, products, and processes (E). Some information is available after time-consuming data processing and analysis only (F). Information differs in complexity, may change quickly, and is accessible through software or printed media; it can be structured, semi-structured, or unstructured. Related software typically has different user interfaces and might be installed on desktop computers. The former requires workers to *learn* using these interfaces (G), and the latter requires them to *move* between the workplace and the computer (H). Both are time-consuming, and moving to access an interface *interrupts and delays tasks* (I).

We describe the following benefits from two perspectives: machine operators and technicians. *Machine operators* are responsible for one or more machines. They operate them, perform simple maintenance tasks, and collaborate with production line managers and technicians to solve problems. *Technicians* are responsible for maintaining and repairing machinery – they are highly mobile and may visit several machines per day. Besides, they analyze the hardware systematically to identify the root causes of problems and address them through repair, replacement, or other on-site measures. Such tasks may require both hands, and technicians must pay attention to avoid injuries or damage to the machine. Typically, they use software for complex analyses, and they collaborate with people from different professions to identify and discuss causes and solutions.

4 Results

The related work above contains various arguments for using a digital assistant in manufacturing. Unfortunately, authors often describe benefits superficially in a single sentence or list them without further justification and explanation. This lack of detail makes it difficult for assistant developers and decision-makers to articulate and identify benefits that solve or partially address specific problems. This section outlines an assistant's benefits and connects them to technology and practical manufacturing challenges. The collection is not comprehensive and represents a work-in-progress.

The order of the following benefits is alphabetical to create a neutral overview. Parentheses with a letter indicate how benefits connect to the challenges in the use case.

Central access (E, G, H, I). A digital assistant interfacing multiple information systems can become a central access point to these systems from the user's perspective. Users access the assistant via one or more personal devices. Central access minimizes learning different interfaces and moving between workplace and suitable computers.

Customization (A, C). Employees differ in skills, competencies, motivations, physical capabilities, and personality. Providing customized assistance can address individual needs and preferences. It can increase the acceptance of using a digital assistant. Higher acceptance could minimize opportunity costs that emerge when employees reject potential assistance. Furthermore, such costs could incur when provided assistance is inefficient due to an individual's characteristics.

Delegation (A, D, F). Tasks can be highly structured and repetitive. As a result, employees perceive these tasks as boring, and employers seek their delegation to a computer. The former affects an employee's satisfaction with their work situation, while the latter saves time. Companies can use this time-saving in two ways. First, to *increase employees' efficiency* because they can perform more of the remaining tasks in this time. Second, to *empower employees* by assigning them new tasks that focus on solving abstract problems and managing fragmented information efficiently.

Users can delegate a variety of tasks to a digital assistant. The assistant performs the delegated tasks with different degrees of autonomy and can perform synchronously with a user's task or asynchronously. *Synchronous* performance is beneficial during imminent work situations where users and assistants collaborate [16]. *Asynchronous* task performance can be necessary when tasks require a long execution time. An assistant can significantly outperform humans in performing specific tasks.

An example is calendar management, where the user can perform the task easily but may decide to delegate it to have time for more valuable activities. A second example is the delegation of root cause analysis. Skilled employees can analyze failures to identify root causes but decide to leave this task to a digital assistant. The assistant may be significantly more efficient than a human because it can process large amounts of data reliably and quickly.

Eyes-free (A, I). For many tasks, it can be beneficial if the employee's eyes focus on the objects of interest. Changing this focus may result in oversight, which can have no impact. Sometimes oversight has an impact, though, ranging from minor follow-up costs to severe injuries or death. Employees can use their voice to interact with a digital assistant without switching their eyes' focus. Its impact is difficult to pinpoint, but it includes less oversight due to the avoided shift of eye focus and avoiding the costs of inaccessible information because the eyes focus on the objects of interest.

Guidance (B). Assistants can guide employees through complex tasks, effectively reducing related skill requirements. This reduction would allow producers to hire less skilled people, reduce training costs, and increase the potential workers' supply. Typically, these employees receive lower salaries and, therefore, further cost savings.

Hands-free (A, I). Performing tasks can require that employees use both hands simultaneously. In these situations, graphical user interfaces are nearly impossible to use.[1] Instead, an employee can use voice to interact with a digital assistant. Its impact is difficult to pinpoint, but it includes aspects such as: saving the time spent on using the graphical interface, increased safety due to avoided work interruptions, and reducing costs of making information accessible while hands are busy. The latter includes, for instance, the time spent by co-workers that must deliver the needed information.

Mobile assistance (A). Employees may need to move during their work to access different locations. A digital assistant can support these employees either while they move or at the target location. Mobile assistance is beneficial when notifications reach a person quickly to minimize follow-up costs – i.e., actual costs and opportunity costs. It is also beneficial if a person must act quickly for the same reason. Acting includes, for instance, delegating a task. Besides, mobile assistance covers on-site support.

Multiple interface types (E). Employees may need specific forms of assistance for tasks and situations. Specific interface types, e.g., voice, text, haptic, or visual elements, can be adequate for some but not all tasks and situations. For instance, an information retrieval task can return a table with measurements. A voice interface conveys the table's contents much slower than an interface that uses visual elements to display the table. A digital assistant can have two or more interfaces to account for the variability among assistances, tasks, and situations.

Permanent accessibility (D). Human co-workers experience fatigue and require recreation time. Their services are, therefore, not accessible to others at all times. A digital assistant has no downtime provided the infrastructure has energy management[2]

[1] Nearly because graphical interfaces may use hands-free technologies, such as eye tracking.
[2] This includes, for instance, recharge strategies for mobile devices and permanent power supply.

and redundancy to compensate for maintenance, repair, overhaul, and breakdown. Users can – at all times – benefit from an assistant, either by requesting support or because the assistant can communicate with the user whenever necessary.

Augmented data analytics is one application for manufacturing that benefits from an assistant's permanent accessibility. For example, users can ask the assistant to perform root-cause analysis any time a factory system fails. Since the assistant has no working hours, it can immediately respond, perform the analysis, and report the result. Besides, an assistant that continuously monitors measurements can notify one or more users about unusual measurements.

Speed (A). Performing tasks faster mostly has the benefit that an employee can spend the saved time on other tasks. Besides, taking less time can minimize follow-up costs, as outlined above. A digital assistant can accelerate task performance in different ways. Speed advantages of functions that involve arithmetic and logical operations are often easy to argue because the time saving is so significant. The benefit of voice interactions is harder to quantify, but Ruan et al. [17] identified that using voice is almost three times faster than typing on a QWERTY keyboard. These results are only indicative because the authors performed their experiments in a controlled environment with little noise. Besides, the error rate in the final transcribed text was higher when using voice. Other areas that can create speed benefits are the flatter navigation structures of voice interfaces and the effects of learning efficient assistance.

5 Conclusion

The results above are our first attempt to answer why digital assistants are helpful in manufacturing. Indeed, they are preliminary, and the remaining vague expressions require evidence and discussion before assistant designers and decision-makers can use them effectively.

Future research should investigate benefits through experiments in *natural work environments* to explore and experience the various factors determining digital assistants' use. These factors are concepts, such as trust in AI systems, impacts on teams and individuals, training and education, and capabilities of open and closed technologies.

Besides, future work has to clarify the **disadvantages and limitations** of using assistants in manufacturing. There is an inherent risk that workers lose their *autonomy* when assistants influence or take over their tasks. Human-in-the-loop designs could ensure that workers always remain essential for the process and participate in decisions.

An assistant has technological and designed limits *understanding language*. For example, it may not understand the jargon in manufacturing and need human help resolving ambiguity. Worker training must create awareness for these constraints and teach how to talk effectively with a digital assistant. Developing and performing this training may be costly and time-consuming.

When digital assistants contribute substantial work in manufacturing, producers become *dependent*. The assistant must work reliably – sometimes even in extreme situations, such as a blackout or network breakdown. It should be replaceable by another digital assistant (e.g., to avoid vendor lock-in) or by a human to cover situations where the assistant is unavailable.

Finally, every additional information and communication technology in an organization increases the risk that third parties *exploit* it. Scenarios range from industrial espionage through eavesdropping to the corruption of an assistant to disrupt production or harm employees.

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. Eurostat (2020) Manufacturing statistics - NACE Rev. 2. https://ec.europa.eu/eurostat/statistics-explained/pdfscache/10086.pdf. Accessed 16 Mar 2021
2. Smit S, Tacke T, Lund S et al. (2020) The future of work in Europe: Automation, workforce transitions, and the shifting geography of employment
3. Luger E, Sellen A (2016) "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. In: Kaye J (ed) Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, pp 5286–5297
4. Wyard P, Simons A, Appleby S et al. (1996) Spoken language systems - Beyond prompt and response. BT Technology Journal 14
5. McTear M, Callejas Z, Griol D (eds) (2016) The Conversational Interface: Talking to Smart Devices, 1st ed. 2016. Springer International Publishing, Cham
6. Harms J-G, Kucherbaev P, Bozzon A et al. (2019) Approaches for Dialog Management in Conversational Agents. IEEE Internet Computing 23:13–22. 10.1109/MIC.2018.2881519
7. Pütten AM von der, Krämer NC, Gratch J et al. (2010) "It doesn't matter what you are!" Explaining social effects of agents and avatars. Computers in Human Behavior 26:1641–1650. 10.1016/j.chb.2010.06.012
8. Seeber I, Waizenegger L, Seidel S et al. (2020) Collaborating with technology-based autonomous agents. INTR 30:1–18. 10.1108/INTR-12-2019-0503
9. Maedche A, Legner C, Benlian A et al. (2019) AI-Based Digital Assistants. Business & Information Systems Engineering 61:535–544. 10.1007/s12599-019-00600-8
10. Knote R, Janson A, Söllner M et al. (2019) Classifying Smart Personal Assistants: An Empirical Cluster Analysis. In: Bui T (ed) Proceedings of the 52nd Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences
11. Romero D, Stahre J, Wuest T et al. (2016) Towards an operator 4.0 typology: a human-centric perspective on the fourth industrial revolution technologies. In: Dessouky M, Dessouky Y, Eldin HK (eds) Proceedings of the 46th International Conference on Computers and Industrial Engineering, pp 608–618
12. Rabelo RJ, Romero D, Zambiasi SP (2018) Softbots Supporting the Operator 4.0 at Smart Factory Environments. In: Moon I (ed) Advances in production management systems: Smart manufacturing for Industry 4.0 : proceedings. Part II, vol 536. Springer, [S. l.], pp 456–464
13. Rabelo RJ, Zambiasi SP, Romero D (2019) Collaborative Softbots: Enhancing Operational Excellence in Systems of Cyber-Physical Systems. In: Camarinha-Matos LM, Afsarmanesh

H, Antonelli D (eds) Collaborative Networks and Digital Transformation: 20th ifip wg 5.5, vol 568. SPRINGER NATURE, pp 55–68

14. 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
15. Rzepka C (2019) Examining the Use of Voice Assistants: A Value-Focused Thinking Approach. In: Santana M, Montealegre R, Rodriguez-Abitia G et al. (eds) Proceedings of the 25th Americas Conference on Information Systems (AMCIS 2019), pp 1–10
16. Norman D (2017) Design, Business Models, and Human-Technology Teamwork. Research-Technology Management 60:26–30. 10.1080/08956308.2017.1255051
17. Ruan S, Wobbrock JO, Liou K et al. (2018) Comparing Speech and Keyboard Text Entry for Short Messages in Two Languages on Touchscreen Phones. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1:1–23. 10.1145/3161187