

# What Do You Call Your Analytical Endeavours?

An Analysis of Term Usage in German Job Openings from 2017 to 2022

Finja Below

Department of Computer Science  
NORDAKADEMIE  
Elmshorn, Germany  
finja.below.i16a@nordakademie.org

Uwe Neuhaus

Department of Media Education  
Europa-Universität Flensburg  
Flensburg, Germany  
uwe.neuhaus@uni-flensburg.de

Michael Schulz

Department of Computer Science  
NORDAKADEMIE  
Elmshorn, Germany  
michael.schulz@nordakademie.de

**Abstract**—This paper examines which terms are used for analytical information systems in business practice and how their use has developed over time. For this purpose, a total of 3,000 German job openings from 2017 to 2022 are subjected to a frequency analysis. Eight relevant generic terms are identified and divided into first- and second-order generic terms based on their definition of relevance to practice. A detailed examination of the common occurrence of the generic terms and the change in their use over time shows that the traditional term Business Intelligence continues to dominate in practice and that the terms Artificial Intelligence and Data Science are becoming increasingly established in a more technical context. More specific terms, such as Advanced Analytics, Machine Learning and Data Mining, on the other hand, are becoming less important for describing analytical information systems.

**Keywords:** *Analytical Information Systems; Business Intelligence; Artificial Intelligence; job openings; frequency analysis.*

## I. INTRODUCTION

Procedures used for data analysis to support decision-making in a business context have existed for decades [4] [6]. Over time, different terms for such Analytical Information Systems (AIS) have emerged. These emphasize certain aspects, focus on a specific application area, or pick up on currently "fashionable" terms [11]. When new terms emerge, it usually takes a while for standardized definitions and a common image to take hold [8]. In addition, the use of terms often differs in the scientific literature and in business practice [11].

This article analyzes which terms are used for AIS in business practice. German job openings from 2017 to 2022 are used for the analysis. It is assumed that job openings are a suitable source of data, as (a) they contain current data and (b) companies need to formulate tasks and requirements for applicants realistically in order to find suitable employees [14].

The aim of this work is to identify the Generic Terms (GT) used currently in business practice. These are determined with the help of a frequency analysis. In addition, the relationship between the GT and the changes in term usage over time are examined. Not part of this article is the explanation of the identified practice definition and relevance shifts over time. These will be analyzed in later research.

This paper is divided into five sections. Following this introduction, Section 2 defines the terms relevant to this article. Next, Section 3 explains the data basis used for the study. Subsequently, Section 4 presents the results of the data analysis. The article concludes with an outlook on future research in Section 5.

## II. DEFINITIONS OF TERMS

The job openings studied are from the AIS sector. With such systems, analyses can be carried out to support management in decision-making. They are usually based on large volumes of data that have to be processed depending on the analysis objective [7]. The examination of job openings described in the following sections revealed that a variety of GT is used in business practice. In particular, the terms used emphasize the scope of the analyses, their degree of complexity, or their methodology. In the following, the most frequently occurring GT are presented.

The term Business Intelligence (BI) describes concepts and methods for supporting managers in decision-making with the help of IT systems [11]. The main objective of BI is to make data available and prepare it for analysis [9]. As a result, there is a close connection to the term Business Analytics (BA). It is often regarded as an extension of BI to include more advanced forms of analysis [9]. According to other sources, on the other hand, the term represents an addition to BI of these advanced analysis methods (cf. e.g. [1]).

To emphasize the increasing complexity of analytics, some authors use the term Advanced Analytics (AA) [3]. With its mathematically and algorithmically sophisticated methods, AA goes beyond the traditional data analysis of BI and BA [1]. The term Data Analytics (DA), on the other hand, emphasizes the importance of increasingly large, multifaceted, and frequently changing data sets for decision making. DA encompasses the entire process from data collection and analysis to the presentation of results [13].

While the terms BI and BA emphasize the economic context, the terms Data Science (DS), Artificial Intelligence (AI), Data Mining (DM) and Machine Learning (ML) focus more on the methods used. DS is an interdisciplinary field in which insights are extracted semiautomatically from sometimes complex data [12]. It thus has great intersections with AI. This term covers methods that enable a system to

interpret data and use it to learn patterns [10]. Subfields of AI are DM, which provides methods and algorithms to discover previously unknown relationships in large data sets [2], and ML, which allows computers to improve automatically based on experience [5].

### III. DATA BASIS

This study continues the work of [14] and adds a temporal component to it.

Job openings collected once a year from 2017 to 2022 on the largest job portals in Germany serve as the data basis. Since the goal of this study is to identify GT used in business practice, a search that was too narrow - e.g., for terms such as BI or AI - could have resulted in relevant information being overlooked. For this reason, "analy\*" was used as the search term, with the asterisk replaced by any string that combined to form a German- or English-language word. Only job openings in which the entered term was mentioned in the job title and which were written in German were considered.

Furthermore, job openings that clearly do not fall into the area of analysis with the goal of decision support were eliminated for the following investigations. This eliminated primarily job openings from the natural sciences and those that focused on classic software development tasks. This elimination step was carried out independently by all three authors in order to minimize the degree of subjectivity in the compilation of the data basis. Divergent decisions were discussed to arrive at a mutually accepted result. As a final data set, 500 job openings were used for each year under consideration.

Based on the resulting 3,000 job openings, a frequency analysis was performed. In this work, terms were chosen as the unit of analysis, formed either by single words or by n-grams.

Only terms that appear in more than 5% of all job openings were considered relevant GT. Candidates that are mentioned in 3-5% of the job openings and thus just miss the threshold are the terms Predictive Analytics, Digital Analytics, Web Analytics, Big Data Analytics, and Data Engineering.

### IV. DESCRIPTIVE DATA ANALYSIS

Figure 1 shows the development of the GT under investigation from 2017/2018 (starting point of the arrows) to 2021/2022 (end point of the arrows). For the analysis, the data of two consecutive years were combined. Data of the years 2019/2020 are not shown for reasons of clarity. The value on the x-axis indicates what proportion of sole use of GT is accounted for by the respective term. Sole use means that none of the other GT considered was mentioned in the job openings. In 2021/2022, sole mention of GT was identified in 298 job openings (2017/2018: 293, 2019/2020:353).

The y-axis shows the proportion of the respective terms in job openings in which more than one GT was mentioned.

The area below the diagonal (Area I) thus contains those GT whose share of sole use is greater than the share of their use in combination with other GT. They can be defined as GT of the first order. They are GT that are sufficiently comprehensive and precise to fully describe an analytics area.

The area above the diagonal (Area II), on the other hand, contains those GT whose share of common use with other GT is higher than the share of their stand-alone use. They are defined as second-order GT; depending on the context, they often cannot stand alone but require other GT for specification.

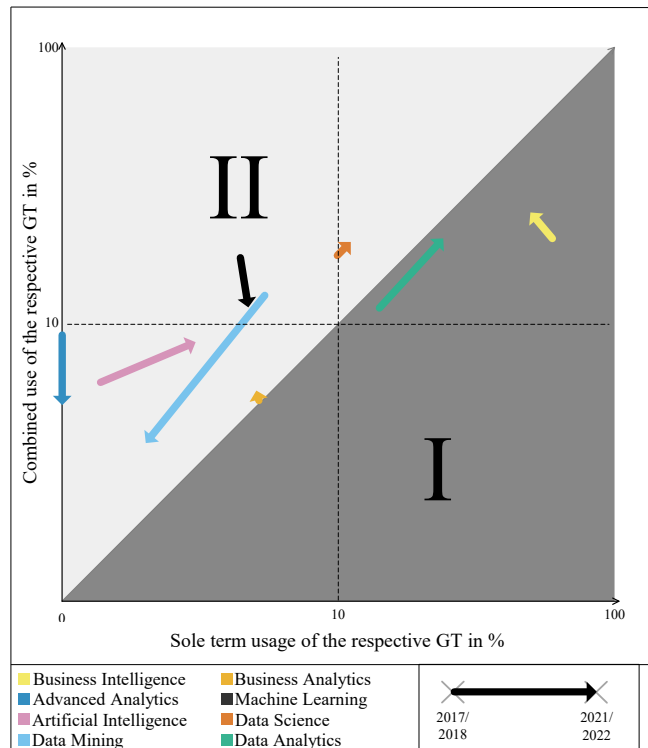


Figure 1. GT development from 2017/2018 to 2021/2022.

The diagram shows the practice definition and practical relevance shift of GT in the period under consideration:

- The *practice definition* of GT is said to be stable if the ratio between sole and joint use does not change, i.e., there is no shift over time or a shift parallel to the diagonal.
- An increase (decrease) of the *practical relevance* can be determined by the fact that the distance of an GT to the origin of the coordinate system increases (decreases) with respect to Figure 1.

Area I of Figure 1 contains the two terms that form GT of the first order: DA and BI. Even though BI alone has lost shares, it remains the most frequently used term overall. In 2021/2022, it was included in 32.5% of job openings (2017/2018: 32.8%, 2019/2020: 39.3%). DA, on the other hand, did not change in its practice definition during the period under consideration, but its relevance did: While the term was included in 12.6% of job openings in 2017/2018, it was already 21.4% in 2021/2022.

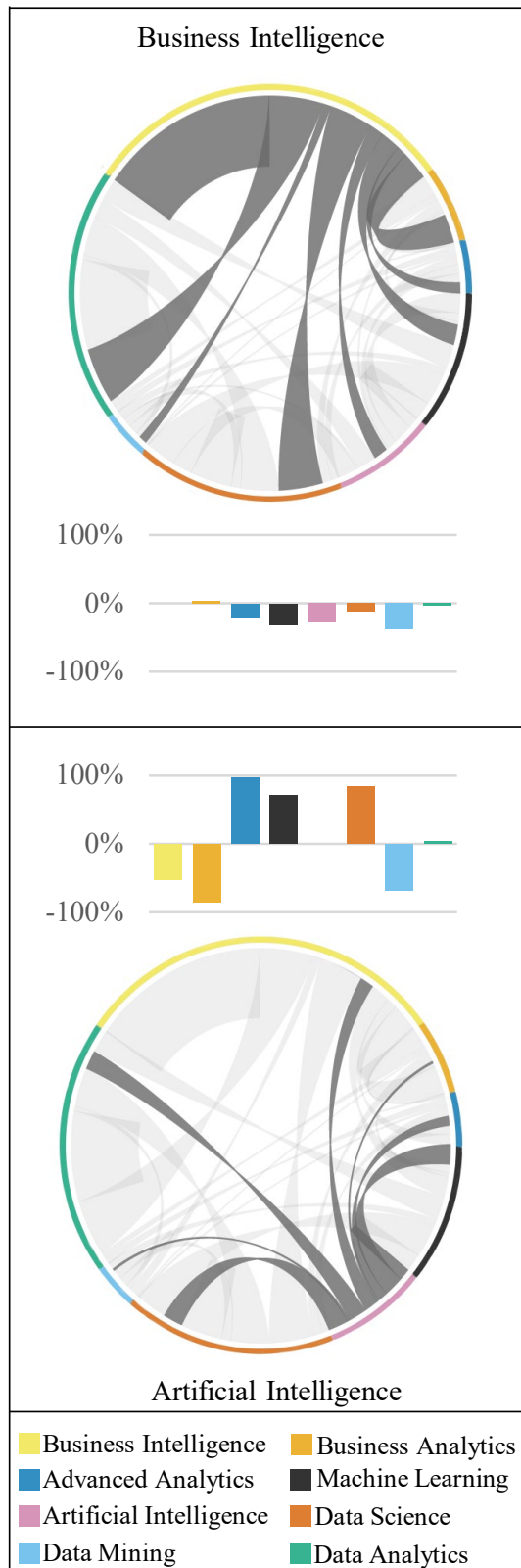


Figure 2. Interrelationships of the GT.

The greater number of GT examined are positioned in Area II of Figure 1 and are thus second-order GT. Only the practice definition and relevance of the term BA remained almost stable over the six years under consideration. The terms AA, which was insignificant as a stand-alone term during the entire period under consideration, and DM have lost relevance while the practice definition has remained unchanged.

The term ML has also lost importance overall, but has gained slightly as a stand-alone term; its practice definition has shifted accordingly. The terms AI and DS, on the other hand, gained overall relevance. The practice definition of DS has remained stable, while AI has become more established as a stand-alone term.

Figure 2 uses Chord diagrams for the years 2021/2022 to show which GT appear in the job openings in combination with which other GT. The bar charts show the deviation of the identified individual relationships from the relationships when all GT are considered (a priori probability). The first part of the figure shows the BI relationships, since this is the most frequently used first-order GT, and the second part the AI relationships, since the relevance of this second-order GT increased the most in the period under review.

The frequency with which GT is mentioned is made clear by the size of the respective occupied circle section in the Chord diagram. In the case of BI, it can be seen that the term is most frequently mentioned on its own (34%). If it occurs together with other terms, these are in particular DA (20%), DS (16%) and BA (10%), where the correlation is slightly weaker in each case than would have been expected when considering the a priori probability. Mention of BI together with more specialized and more technical generic terms such as ML (7%), AI (5%) and DM (3%), on the other hand, are significantly less frequent.

The generic term AI rarely stands alone (6%); it is very often accompanied by other GT that specify the focus of the job openings. The most frequent combinations are DS (22%), ML (22%) and DA (21%). The first two also occur significantly more frequently than might have been expected (DS +84%, ML +71%). At -52% significantly less frequently than would have been expected when considering the a priori probability, AI is mentioned in the job openings together with BI (15%). The term AA occurs 97% more frequently in combination with AI than would have been expected.

## V. CONCLUSION AND FUTURE RESEARCH

In this article, the terms most intensively used by companies at present for AIS were identified. For this purpose, 3,000 job openings serves as analytical data basis. Furthermore, it was examined to what extent the practical definition and practical relevance of GT changed from 2017 to 2022. Using the terms Business Intelligence and Artificial Intelligence as examples, it was shown which GT are frequently or rarely used in combination with other GT.

In the future, more in-depth text analytics will be applied based on the described data set from 2017 to 2022. The job openings contain further information that allows a better understanding of the focus and developments of the use of AIS in operational practice. This includes the qualifications required of applicants in the various phases of data analysis,

the analysis software skills needed, and the planned areas of application and types of tasks. Frequently, the ads also name the driving developments underlying the analytics desire (e.g., digitalization, Industry 4.0, or Internet of Things). It is planned to extract and cluster this information so that its relevance can be determined and significant changes over time can be shown.

This more detailed investigation will also make it possible to work out more concretely how the GT under consideration are used in operational practice. On this basis, an attempt can then be made to explain the shifts in practice definition and relevance presented in this article and to determine whether there has been a fundamental change in data analysis for decision support or whether it is simply a shift in the trend of terms used.

what you're doing? The term 'Analytics' in operational practice", *NORDBLICK*, no. 6, pp. 4-29, 2018.

#### REFERENCES

- [1] P. Chamoni, "Data Science - Entwicklungslinien und Trends" [in English: "Data Science - Lines of Development and Trends"] in D. Frick, A. Gadatsch, J. Kaufmann, B. Lankes, C. Quix, A. Schmidt and U. Schmitz, (Eds.), *Data Science: Konzepte, Erfahrungen, Fallstudien und Praxis* [in English: *Data Science: Concepts, Experiences, Case Studies and Practice*], Springer Fachmedien, pp. XII–XVI, 2021.
- [2] J. Cleve and U. Lämmel, *Data Mining* (3. ed.). De Gruyter, 2020.
- [3] C. Dittmar, "Die nächste Evolutionsstufe von AIS: Big Data" [in English: *The next evolutionary stage of AIS: Big Data*] in P. Gluchowski and P. Chamoni (Eds.), *Analytische Informationssysteme: Business Intelligence-Technologien und -Anwendungen* [in English: *Analytical Information Systems: Business Intelligence Technologies and Applications*], (5. ed., vol. 51). Springer Gabler, pp. 55-65, 2016.
- [4] G. A. Gorry and M. S. S. Morton, "A framework for management information systems", *Sloan Management Review*, vol. 13, no. 1, pp. 56-79, 1971.
- [5] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects", *Science*, Vol 349, Issue 6245, pp. 255-260, 2015.
- [6] P. G. Keen, "Decision support systems: a research perspective" in *Decision support systems: Issues and challenges: Proceedings of an international task force meeting*, pp. 23-44, 1980.
- [7] H. Kempter, *Betriebliche Informationssysteme: Datenmanagement und Datenanalyse* [in English: *Business Information Systems: Data Management and Data Analysis*]. Kohlhammer Verlag, 2017.
- [8] J. H. Kroeze, M. C. Mathee and T. J. D. Bothma, "Differentiating data-and text-mining terminology" in J. Eloff (Ed.), *ACM Conferences. Proceedings of the 2003 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology*, South African Institute for Computer Scientists and Information Technologists, pp. 93-101, 2003.
- [9] E.-P. Lim, H. Chen and G. Chen, "Business Intelligence and Analytics", *ACM Transactions on Management Information Systems*, vol. 3, no. 4, pp. 1–10, 2013.
- [10] P. Mikalef and M. Gupta, "Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance", *Information & Management*, vol. 58, no. 3: 103434, 2021.
- [11] S. Negash, "Business Intelligence", *Communications of the Association for Information Systems*, vol. 13, no. 15, pp. 177-195, 2004.
- [12] M. Schulz, et al., "Introducing DASC-PM: A Data Science Process Model", *ACIS 2020 Proceedings*, vol. 45, 2020.
- [13] C.-W. Tsai, C.-F. Lai, H.-C. Chao and A. V. Vasilakos, "Big data analytics: a survey", *Journal of Big Data*, vol. 2, no. 1, pp. 1-32, 2015.
- [14] U. Neuhaus and M. Schulz, "Wie nennt ihr, was ihr da tut? Der Begriff "Analytics" in der betrieblichen Praxis" [in English: "What do you call