UNDERSTANDING THE ANATOMY OF ANALYTICS-BASED SERVICES – A TAXONOMY TO CONCEPTUALIZE THE USE OF DATA AND ANALYTICS IN SERVICES

Fabian Hunke  
_Karlsruhe Institute of Technology_, fabian.hunke@kit.edu

Christian Engel  
_University of St. Gallen_, christian.engel@unisg.ch

Ronny Schüritz  
_Karlsruhe Institute of Technology_, ronny.schueritz@kit.edu

Philipp Ebel  
_University of St. Gallen_, philipp.ebel@unisg.ch

Follow this and additional works at: https://aisel.aisnet.org/ecis2019_rp

Recommended Citation

https://aisel.aisnet.org/ecis2019_rp/25

This material is brought to you by the ECIS 2019 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in Research Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
UNDE RSTANDING THE ANATOMY OF ANALYTICS-BASED SERVICES – A TAXONOMY TO CONCEPTUALIZE THE USE OF DATA AND ANALYTICS IN SERVICES

Complete Research

Hunke, Fabian, Karlsruhe Institute of Technology, Karlsruhe, Germany, fabian.hunke@kit.edu
Engel, Christian, University of St. Gallen, St. Gallen, Switzerland, christian.engel@unisg.ch
Schüritz, Ronny, Karlsruhe Institute of Technology, Karlsruhe, Germany, ronny.schüritz@kit.edu
Ebel, Philipp, University of St. Gallen, St. Gallen, Switzerland, philipp.ebel@unisg.ch

Abstract

The abundance of data accompanied by advances in analytics technologies increasingly drive companies to introduce analytics-based services, i.e. customer-facing services in which data and analytics help customers make decisions. Despite its growing application in practice, theoretical and conceptual work on analytics-based services is still scarce. In this paper, we develop a taxonomy of analytics-based services unveiling their conceptually grounded and empirically validated characteristics. Applying an established taxonomy building method, we draw upon an analysis of 85 use cases of analytics-based services. The results of an expert evaluation indicate both the usefulness and robustness of our taxonomy. The developed taxonomy of analytics-based services contributes in two ways: First, we add to the descriptive knowledge on this new service type, establish a common language among researchers and equip them with the means to analyze analytics-based services in a structured manner – thus laying the foundation for a deeper theorizing process in the future. Second, we provide a concrete conceptualization of analytics-based services for practitioners for initial guidance in new service development.

Keywords: Analytics-based Service, Taxonomy, Service Advancement.

1 Introduction

For decades, companies have gathered, stored, accessed, and analyzed data to support the internal decision making processes (McAfee and Brynjolfsson 2012; Watson 2009). Yet, nowadays companies increasingly strive to innovate their customer-facing service offerings through the application of data and analytics (Demirkan et al. 2015). Such analytics-based services are a special form of digital services that build upon data and analytics to extend existing service portfolios. They provide customers with information-rich data, deliver insights, or even act and make decisions for the customer, thereby creating new customer value. In contrast to digital services, they do not only provide a service in a digital manner, but provide the customer with information that helps them make better decisions (Schüritz et al. 2017; Wixom and Ross 2017; Wixom and Schüritz 2017). The majority of companies invest in this opportunity to enrich their offerings with analytics-based services (Gottlieb and Rifai 2017). For example, agricultural biotechnology corporations use databases containing map-data about their customer’s fields, weather simulations, and soil observations to decide which seed is expected to grow best on the customer’s individual field (Marshall
et al. 2015). Heavy equipment manufacturers predict the failure of machine parts and proactively maintain equipment in order to reduce operational downtime for their customers (Porter and Heppelmann 2014). As a consequence of these new technological opportunities, there is an increasing interest in the field. However, there is still limited research on the use of data and analytics for the advancement of service and scholars call for more empirical studies in that field (Lim et al. 2018b; Ostrom et al. 2015). The existing research is not yet exhaustive as we do not know the variety of possibilities in which data and analytics can be used to advance services. Practitioners alike would benefit from a clear understanding of how analytics-based services use data and analytics to create new customer value and how they differ in order to foster their systematic development.

In this paper we lay the foundation to understand this novel service type by developing a taxonomy that helps to conceptualize analytics-based services. Taxonomies have proven to enable researchers and practitioners to understand, analyze, and structure the knowledge within emerging research fields (Nickerson et al. 2013). By identifying common characteristics that can be used to distinguish analytics-based services from another and consolidating these characteristics into one unified taxonomy, we intend to answer the following research question: What are the conceptually grounded and empirically validated characteristics that describe analytics-based services?

For the development of the taxonomy we use a structured and well-used method by Nickerson et al. (2013). We conduct a structured literature review to identify the relevant concepts and use 85 cases of analytics-based services to empirically revise our taxonomy. We identify six key dimensions that help to distinguish and explain analytics-based services. Each dimension consists of a distinct set of characteristics. We evaluate the taxonomy’s general applicability and usefulness by letting independent experts applying it to a set of cases.

The developed taxonomy contributes to the existing body of knowledge in the field by establishing a common understanding of analytics-based services. Such a shared language contributes to structure the field of research helping researchers to position their work within this field. In addition to that, the shared understanding resulting from our taxonomy allows the materialization of ideas and considerations leading towards the development of a design theory on analytics-based services. For practitioners, the taxonomy may provide a useful tool and initial guidance for assessing the strengths and weaknesses of a proposed service and for benchmarking new analytics-based services against the competitor’s services.

The paper is structured as follows. First, we provide an overview of the extant literature on using data and analytics to advance service and taxonomies as a means for designing new services in section 2. Section 3 describes our general approach to develop the taxonomy. In Section 4 we present the development stages of the taxonomy and provide a detailed account for the data sources and methods for each step. We close the paper by discussing the implications of our research, reflecting its limitations and describing possible next steps.

2 Related Work

2.1 The Use of Data and Analytics in Services

Conceptualizing the use of data and analytics in services as a means to understand how to create value for the benefit of the customer is perceived a highly relevant topic in service research (Ostrom et al. 2015). Driven by advances in information and communication technologies leading to an ever-increasing amount of available data, several researchers have called for in-depth studies on the impact on services (Maglio and Lim 2016; Ostrom et al. 2015).

Chen et al. (2011) discuss two basic service applications with data and analytics, "Data-as-a-Service" and "Analytics-as-a-Service". The first focuses on providing raw and aggregated content, which is provided in new data-driven services (Hartmann et al. 2016). The latter refers to services that provide customers with a broad set of customizable analytics that enable them to analyze large amounts of data and to create their own analytics-based offerings (Fromm et al. 2012). Organizations may also use data and analytics
to amplify its existing value proposition offered to customers by enriching its services with data and analytics. The terms used for such services vary in the literature (e.g. "data-enriched" (Davenport 2013) or "data-infused" (Schüritz and Satzger 2016)), but generally refer to an increased integration of data and analytics into an organization’s core business. In this regard, data and analytics is often understood as a key to understanding and attracting customers. Huang and Rust 2013 (p.255) point out that analytics applied on customer data collected from IT-related services make it possible to “figure out [...] why customers make decisions they make and why they behave in a certain way”. Data and analytics are acknowledged to facilitate a much deeper access to the customer by transforming customer-related data into information that directly supports the service offering (Saarijärvi et al. 2014; Wixom and Schüritz 2017). An energy supplier, for instance, could amplify its underlying service with accurate, real-time energy consumption data presented to the customer, thereby serving the customer’s sustainability concerns, additionally (Saarijärvi et al. 2014).

Driven by trends like the internet of things (Atzori et al. 2010), researchers also investigate how to enrich existing product offerings. Tangible products are increasingly equipped with sensor technology which enables them to sense their own condition and their external environments and thus allow for real-time data collection (Allmendinger and Lombreglia 2005; Wünderlich et al. 2015). Based on these increasingly ‘smart’ objects, organizations can use information from the collected data to offer contextual and preemptive services, predominantly referred to as smart services (e.g. Beverungen et al. 2017), which gives them the opportunity to strategically differentiate themselves in the market (Porter and Heppelmann 2014). Similar, researchers point out the opportunity to harvest advances in information technology by enriching products with information-intensive service offerings (Lim et al. 2018b). These services allow for a value-add based on data and analytics by ‘wrapping’ contextual information around the core product (Woerner and Wixom 2015). Such smart, information-heavy services generally require an intelligent object (Allmendinger and Lombreglia 2005) and the resulting solutions often simultaneously embody the characteristics of products and services within smart product-service systems (Mont 2002; Valencia et al. 2015).

These studies demonstrate that the proliferation of data and analytics provide organizations with numerous opportunities to create customer value and many studies have discussed distinct benefits from a phenomenological view point. However, considering largely untapped data sources which hold promise for new business opportunities (Balducci and Marinova 2018), research still seems far from being exhaustive in this field. Further, little is known how different service concepts relate to or even build on each other, suggesting that, conceptually, using data and analytics in service is still underexplored. Yet, a conceptual understanding is essential, in particular when researchers and practitioners investigate the systematic design of new services in this field (Edvardsson and Olsson 1996; Goldstein et al. 2002). We argue that research on using data and analytics in services needs to take a broader view in order to establish a common understanding. Thus, as a comprehensive term for the use of data and analysis in services, we refer to analytics-based services and define it as a new service type in which the application of analytical methods (‘analytics’) on data significantly contributes to new customer value as it provides customers with situational insights, or even solves problems autonomously by making decisions.

2.2 Taxonomies as a Means for Designing New Services

Service design is seen as a design-led approach to new service offerings that integrates contributions from the fields of service marketing, service operations, and information technology (Blomkvist et al. 2010; Meroni and Sangiorgi 2011). It represents a multidisciplinary research field that brings to life new ideas by drawing upon human-centered, creative, and iterative methods to create new service offerings (Sangiorgi et al. 2017). Several service design methods exist in literature, such as multilevel service design (Patrício et al. 2011) or casebook design (Kim et al. 2012), and thus service design could serve as a basis to stimulate analytics-based service innovation. However, existing methods and tools neither focus on the context in which data and analytics are key to value creation nor consider newly emerging data sources,
e.g. sensor data. Thus, actionable insights to systematically design new analytics-based services are still lacking (Schüritz et al. 2017).

When it comes to the design of analytics-based services, literature lacks a common understanding of the role of data and analytics to enable the innovation of analytics-based services. The design of new services requires some sort of knowledge on the ‘anatomy’ of services and their key factors that need to be considered to achieve the desired differentiation of a new service from existing ones in the design process. In this sense, the ability to classify services among common characteristics is a prerequisite to the design of new services (McKelvey 1982). Scholars have introduced several taxonomies, i.e. unified classification schema (Nickerson et al. 2013), as a means to describing and understanding services (e.g. Lovelock 1983), gain managerial insights (e.g. Lim et al. 2018b), or identify archetypes of services (e.g. Glushko 2010). In the context of technological advancements, Williams et al. (2008) proposed a taxonomy of digital services unveiling the four key design dimensions and the respective service provider’s objectives. Hartmann et al. (2016) introduced a taxonomy of data-driven business models which illustrates different ways data and analytics afford value creation in data-driven services. Bridging the research on digital and data-driven services, Rizk et al. (2018) developed a taxonomy to deeper understand the key aspects in utilizing data and analytics in digital services.

These studies demonstrate that further knowledge is required to conceptualize the use of data and analytics in services and that taxonomies are a valid means to fulfill this purpose. However, the application of extant studies is still limited to the context of analytics-based services. In fact, we could not identify a framework during our literature review that takes into account e.g. multiple data sources, different types of analytics, or different value creation opportunities to provide affordances for analytics-based services. Thus, we argue that a comprehensive taxonomy concerning analytics-based services is required which illustrates the key factors that need to be considered from a service design perspective.

3 Methodological Approach to Taxonomy Development

We aimed to unveil the conceptually grounded and empirically validated characteristics that allow to describe analytics-based services. For that purpose, we decided to develop a taxonomy and provide a first evaluation of it regarding its applicability. In terms of taxonomy development, we followed Nickerson et al. (2013) and adopted their approach to our study context. This method seemed appropriate for our purpose, since several IS studies have successfully applied this method in different study contexts (e.g. Gimpel et al. (2018); Williams et al. (2008)), suggesting its robustness in developing taxonomies. For the evaluation, we designed an own approach as the suggested method did not provide this.

3.1 Preliminary Instructions

The taxonomy development method suggested by Nickerson et al. (2013) constitutes an iterative approach which allows researchers to build taxonomies both conceptually, based on literature, and empirically, based on real-world objects.

It starts with two preliminary instructions before the actual building process starts. First, the researcher defines a meta-characteristic which all dimensions and characteristics following in the development process will be a logical consequence of. Afterwards, the researcher defines ending conditions that need to be fulfilled entirely for the development process to terminate. The following objective ending conditions have been adapted from Nickerson et al. (2013): (1) All objects are analyzed. (2)-(4) Every dimension, every characteristic within its dimension, and every combination of characteristics is unique. (5) Each characteristic is associated with at least one object from our dataset. (6)-(8) There has been no variation (merge, split, new addition) of dimensions, characteristics, or cases in the last iteration of the development process. In addition to that, we also derived the following subjective ending conditions for the taxonomy: (9) It is concise having a limited number of dimensions and characteristics. (10) It is robust containing enough dimensions and characteristics to allow for a clear distinction between analytics-based services.
(11) It is comprehensive allowing to classify all cases in our dataset. (12) It is extendable allowing to add new dimension or characteristics in case novel types of analytics-based services appear in the future. (13) It is explanatory providing useful explanations to understand the nature of analytics-based services.

### 3.2 Taxonomy Building Process

The actual taxonomy building process consists of two independent modules that are applied consecutively in several iterations. First, the conceptual module allows the researcher to conceptualize the dimensions of the taxonomy without examining actual objects. Little guidance is provided to the researcher in terms of how to proceed in this step other than to use his respective knowledge on the field. Second, the empirical module allows the researcher to derive dimensions and characteristics based on a dataset of real-world objects. For that purpose, the module instructs to identify a subset of the objects available which are then analyzed with regard to common characteristics. Next, the identified characteristics are grouped into dimensions to create a (revised) taxonomy. Regardless of the module chosen, the researcher needs to assess if the ending conditions are fulfilled completely, afterwards. In case the process does not terminate, the researcher chooses between the two modules again to conduct a subsequent iteration of the taxonomy building process.

### 3.3 Evaluation Approach

Following the taxonomy development process, we evaluated the taxonomy’s general reliability to classify analytics-based services. Thereby, we intended to lend credibility on our research approach leading to a robust taxonomy in the sense that it ensures consistent results in characterizing analytics-based services according to its dimensions when applied by different users.

For that purpose, four experts, recruited in their capacity as researchers from our research institution, independently classified a subset of use cases from our dataset using the developed taxonomy. The experts where familiar with the research field and the concept of analytics-based services in general; yet, they had not taken part in the taxonomy development process itself. Due to the complex and time-consuming coding process (i.e. each case needs to be validated with regard to 20 characteristics), we randomly selected a new subset of ten cases that had to be classified.

### 4 Results

This section summarizes the results from our study. First, we report on our taxonomy development process by describing each conducted iteration, highlighting important insights we derived during the process, and illustrating the interim status of the taxonomy ($T$) and its dimensions ($D$). Second, we introduce our final taxonomy of analytics-based services in detail. Third, we confirm both the usefulness and robustness of our taxonomy and illustrate its application.

#### 4.1 Taxonomy Development

**Meta-characteristic:** The main purpose of our taxonomy is to determine the nature of analytics-based services, in order to support researchers and practitioners in the design of new analytics-based services. Therefore, we defined “relevant for the description of an analytics-based service” as our meta-characteristic.

**1st iteration:** For our first iteration, we decided to apply the conceptual module of the method. It allowed us to build upon the already existing work on analytics-based services in the literature. Since no further guidance was given regarding how to conceptualize the initial taxonomy dimensions, we decided to conduct a systematic literature review (Webster and Watson 2002). Striving to characterize analytics-based services, we used precise, quite narrow search strings to identify papers dealing services that build
upon data and analytics to create customer value as their central unit of analysis – only in this case a paper was perceived as relevant for our context. Thus, we included different wording alternatives for analytics-based service as shown in Table 1. This helped us to exclude literature that rather deals with e.g. digital transformation processes within organizations resulting from big data or advances in analytics technologies instead of analytics-based services as a novel service-type. To cover a representative set of relevant research, the search scope was set on the full text of the articles. We included five scientific databases (AISeL, Emerald, ScienceDirect, EBSCOhost, SCOPUS) in the search process to achieve a representative coverage of contributions in leading IS journals as well as related disciplines such as computer science and manufacturing. Additionally, it allowed us to include scientific papers from leading IS conferences. Thereby, we were able to consider more recent studies that had not reached a journal outlet, yet. Overall, we identified 82 relevant papers that were published between 2006 and 2018.

<table>
<thead>
<tr>
<th>Search string</th>
<th>AISeL</th>
<th>Emerald</th>
<th>Science</th>
<th>Ebsco</th>
<th>Scopus</th>
<th>Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;data-driven service&quot;</td>
<td>23</td>
<td>11</td>
<td>27</td>
<td>44</td>
<td>50</td>
<td>33</td>
</tr>
<tr>
<td>&quot;data-enabled service&quot;</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>&quot;data-enriched service&quot;</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>&quot;data-enriched product&quot;</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>&quot;data-based service&quot;</td>
<td>9</td>
<td>3</td>
<td>174</td>
<td>12</td>
<td>35</td>
<td>5</td>
</tr>
<tr>
<td>&quot;service analytics&quot;</td>
<td>34</td>
<td>6</td>
<td>337</td>
<td>54</td>
<td>38</td>
<td>13</td>
</tr>
<tr>
<td>&quot;analytics-based AND service&quot;</td>
<td>58</td>
<td>21</td>
<td>33</td>
<td>88</td>
<td>39</td>
<td>6</td>
</tr>
<tr>
<td>&quot;analytics-as-a-service&quot;</td>
<td>12</td>
<td>0</td>
<td>64</td>
<td>81</td>
<td>74</td>
<td>9</td>
</tr>
<tr>
<td>Sum (w/o duplicates)</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Added through forward/backward search</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total sum</td>
<td>82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Summary of articles reviewed during 1st iteration.

Using a concept-matrix (Webster and Watson 2002) to analyze the identified literature, we identified four central topics stressed in the context of using data and analytics to advance service. Adopting these topics, we derived four initial taxonomy dimensions; namely data as a source for new value creation, the analytics type used to analyze the data, the role of service to the existing portfolio of the company, and the role of the service user in service related interaction. Due to the purely conceptual nature of this first iteration, the taxonomy did not fulfill several ending conditions.

\[
T = \{ \begin{align*}
   D_1 & \text{ Data} & \mid D_1 = \{ \text{empty} \} \\
   D_2 & \text{ Analytics type} & \mid D_2 = \{ \text{empty} \} \\
   D_3 & \text{ Role of service to existing portfolio} & \mid D_3 = \{ \text{empty} \} \\
   D_4 & \text{ Service user role} & \mid D_4 = \{ \text{empty} \} 
\end{align*}
\]

2nd iteration: For our second iteration, we chose the empirical module of the method. For that purpose, we collected a set of real-word use cases of analytics-based services. In order for an organization to engage in the design and development of analytics-based services, it requires a set of different technologies to collect, store, process, analyze or visualize data (Fayyad et al. 1996). To efficiently build a large dataset for our context, we assumed that companies would develop these services with the support of software vendors who naturally possess this kind of knowledge. We therefore opted for software vendors as a means to systematically identify analysis-based services from different companies. More precisely, we used the publicly available customer success stories of software vendors. In these short reports, software vendors present successfully completed projects including information on the company partner, the project task and the solution developed and brought to market. To make the data collection process as transparent as possible, we selected the customer references of the three globally leading software vendors according to their annual revenue, Microsoft, Oracle and IBM (Chitkara and McCaffrey 2016), as a sample frame.
We collected all customer references on their individual websites that were tagged with “analytics”, “machine learning”, “big data”, or “artificial intelligence”. By gathering all cases in a first step, we expected to eliminate selection biases by the researcher. In addition to that, we increased the sample’s diversity consisting of cases from different regions, industries, and company sizes. Subsequently, we screened and purposefully sampled the collected cases along the following criteria: First, as a central aspect of analytics-based services (e.g. Davenport 2013, also cf. section 2.1), each case was required to describe a service in which data and analytics were used for the benefit of an (external) customer. Second, cases needed to provide information about the type of data used, the analytics technology applied as well as the solution outcome. This filtering resulted in a final set of 85 cases which were perceived as sufficient to serve as a basis for the further development of our taxonomy building process. Table 2 provides an overview of the sample structure by industry field, region, and company size. With regard to determining the company size, we followed the recommended definition provided by the European Union (2003).

<table>
<thead>
<tr>
<th>Industry field [%]</th>
<th>Region [%]</th>
<th>Company size [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>16.0 Retail</td>
<td>7.4 North America</td>
</tr>
<tr>
<td>Healthcare</td>
<td>14.8 Energy and utilities</td>
<td>7.4 Europe</td>
</tr>
<tr>
<td>Education</td>
<td>12.3 Public</td>
<td>4.9 Asia-Pacific</td>
</tr>
<tr>
<td>Information and communication</td>
<td>11.1 Financial</td>
<td>3.7 Africa</td>
</tr>
<tr>
<td>Entertainment</td>
<td>11.1 Financial</td>
<td>2.5</td>
</tr>
<tr>
<td>Travel and Transportation</td>
<td>8.6</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Sample structure of the dataset used for the taxonomy development.

In line with the method’s empirical module, we started by selecting a first subset of 20 cases which we randomly selected from our dataset. We then analyzed the cases identifying common characteristics among them, which we grouped into dimensions afterwards. Several ending conditions were still not met by our taxonomy (e.g. characteristics were split), indicating that it had still not met a sufficient level of maturity. Thus, a third iteration was needed.

\[
T = \{ \\
D_1 \text{ Data source} \\
D_2 \text{ Analytics type} \\
D_3 \text{ Role of service to existing portfolio} \\
D_4 \text{ Service user role} \\
\}
\]

\[
D_1 = \{ \text{Customers; Non-customers; Processes; Objects; Processes; Physical surroundings; Natural surroundings} \}
\]

\[
D_2 = \{ \text{Descriptive; Predictive} \}
\]

\[
D_3 = \{ \text{New; Extension} \}
\]

\[
D_4 = \{ \text{Mere recipient; Collaborator} \}
\]

**3rd iteration:** For our third iteration, we conducted the empirical module again analogous to the previous iteration analyzing 27 new cases from our dataset. We were still able to identify new characteristics which confirmed our decision to split our initial dataset into subsets. For instance, we noticed that it did not seem sufficient to address data-related characteristics within a single dimension describing the data source. Instead, we decided to split this dimension into a dimension characterizing the generation of data, and a dimension characterizing the target the data is expected to contain information about. This allowed us to distinguish cases on a much more granular level in that perspective. Still, our taxonomy required an additional iteration to eventually fulfill all ending conditions. However, we noticed considerable improvement with regard to the number of ending conditions met up to this point.

\[
T = \{ \\
D_1 \text{ Data generator} \\
D_2 \text{ Data target} \\
D_3 \text{ Data origin} \\
D_4 \text{ Analytics type} \\
D_5 \text{ Role of service to existing portfolio} \\
D_6 \text{ Service user role} \\
\}
\]

\[
D_1 = \{ \text{Customer; Non-customer; Object; Process} \}
\]

\[
D_2 = \{ \text{Customer; Non-customer; Object; Process; Environment} \}
\]

\[
D_3 = \{ \text{Internal; External} \}
\]

\[
D_4 = \{ \text{Descriptive; Diagnostic; Predictive; Prescriptive} \}
\]

\[
D_5 = \{ \text{Stand-alone; Wrapped around product; Wrapped around service} \}
\]

\[
D_6 = \{ \text{Data receiver; Data provider; Integrator} \}
\]
**4th iteration:** For our fourth iteration, we decided to apply the empirical module again. Since we noticed that the majority of ending conditions were fulfilled already, we expected little changes to be made with regard to the characteristics and dimensions already included in the taxonomies. Thus, we decided to include the remaining 38 cases in this iteration. In addition to the instructions of the empirical module, we furthermore focused on clear definitions and formulations of the characteristics and dimensions already identified. For that purpose, we discussed extreme examples, and ambiguous examples of analytics-based services described in the subset and revised the wording of dimensions and characteristics wherever necessary.

$$T = \{D_1 \text{ Data generator} \mid D_1 = \{\text{Customer; Non-customer; Process; Object}\}$$

$$D_2 \text{ Data target} \mid D_2 = \{\text{Customer; Non-customer; Process; Object; Environment}\}$$

$$D_3 \text{ Data origin} \mid D_3 = \{\text{Internal; External}\}$$

$$D_4 \text{ Analytics type} \mid D_4 = \{\text{Descriptive; Diagnostic; Predictive; Prescriptive}\}$$

$$D_5 \text{ Portfolio integration} \mid D_5 = \{\text{Stand-alone; Wrapped around product; Wrapped around service}\}$$

$$D_6 \text{ Service user role} \mid D_6 = \{\text{Recipient; Provider of data; Interactor}\}$$

All ending conditions were satisfied after this iteration which prompted the taxonomy building process to end. The ending conditions and their fulfillment after each iteration are listed in Table 3.

<table>
<thead>
<tr>
<th>Ending Condition</th>
<th>1st (concept.)</th>
<th>2nd (empir.)</th>
<th>3rd (empir.)</th>
<th>4th (empir.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) All objects within the sample were analyzed</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(2) Every dimension is unique</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(3) Every characteristic within the dimension is unique</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(4) Every combination of characteristics is unique</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(5) At least one object assigned to each characteristic</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(6) No new dimension added</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(7) No dimension or characteristic was merged or split</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(8) No objects were merged or split</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(9) Concise</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(10) Robust</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(11) Comprehensive</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(12) Extendable</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(13) Explanatory</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 3. Summary of the iterations and ending conditions.

### 4.2 Analytics-based Service Taxonomy

In this section, we present the final taxonomy that resulted from our research process. We, thereby, provide a first answer to our research question regarding the characteristics that describe analytics-based services. The developed taxonomy consists of six key dimensions, each entailing a certain set of characteristics. The six key dimensions as well as their definitions and the literature contributing to their development are consolidated in Table 4. Even though Nickerson et al. (2013) recommend characteristics to be mutually exclusive in taxonomies, we decided to allow characteristics to be non-exclusive within data-related dimensions. Literature argues that, in service contexts, data increases in value when coupled with other data (Porter and Heppelmann 2015). In line with this, we learned from the use cases we studied that analytics-based service create more complex insights for their customers when drawing on multiple data sources. We visualize our final taxonomy in Fig. 1. In the right column we describe, whether a dimension is either exclusive (E), i.e. one characteristic observable at a time, or non-exclusive (N), i.e. potentially multiple characteristics observable at a time.
The **Data Generation** dimension in our taxonomy addresses the entities, which generate the data for the analytics-based service. From the empirical iterations of our development process we learned that companies draw upon four distinct types of data generators – Customers, non-customers, processes, and objects. Nowadays, customers have turned into “walking data generators” (Chen et al. 2012, p. 5), thereby becoming an increasingly important source of information for analytics-based services (Loebbecke and Picot 2015). Considering the service’s perspective, we defined customers as direct clients using an analytics-based service including business-to-business-to-customer configurations. Contrarily, we define non-customers being all humans, who generate data for an analytics-based service, but do not consume the service themselves directly. With digitalization actively transforming industries, physical objects are increasingly equipped with sensors and actuators (Atzori et al. 2010). Thus, physical objects turn into data generators becoming a valuable source of data for analytics-based services (Wunderlich et al. 2015). Accordingly, this accounts for business processes – from production processes within manufacturing to consumption processes of customers – indicating key process indicators, for instance.

**Data Origin** describes where the generated data comes from. We found that analytics-based services truly differ from each other whether the data is characterized as internal or external. An internal origin represents data that comes from within the company offering the service. This may range from machine data to event-based data. In contrast, external data refers to data that comes from outside the company. Our analysis of the use cases showed that this particularly referred to publicly available data or purchasable data (e.g. weather data).

The **Data Target** dimension specifies about whom or what the generated data contains information. Sticking to the statement above – customers have turned into walking data generators – customers might not only generate data regarding themselves, but also about other targets as well. Thus, we adopted the same characteristics from the data generator dimension as we noticed that these characteristics hold true for data targets as well. Additionally, we added environment as a possible data target, which refers to the surrounding of respective data generators, e.g. weather data such as temperature, or wind velocity.

**Analytics Type** refers to the type of analytics applied in an analytics-based service. Scholars distinguish between four characteristics, namely descriptive, diagnostic, predictive, and prescriptive (Han and Kamber 2006; Porter and Heppelmann 2015). Descriptive analytics focus on reporting and visualizing data which often results in aggregated reports or accumulated visualizations. Diagnostic analytics follow a deductive approach aiming to derive answers to why something happened. In turn, predictive analytics follow an inductive approach. This characteristic comprises of advanced analytics focusing on predicting what will happen according to existing data. Prescriptive analytics take it a step further and investigate what should be done based on existing data and its resulting predictions; thus, prescriptive analytics heavily build on simulation - and optimization techniques.

**Portfolio Integration** describes the integration of the analytics-based service into the existing portfolio of the provider and clarifies how the analytics-based service is related to existing products or services.

---

**Figure 1. Taxonomy of analytics-based services.**

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Characteristics</th>
<th>E/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Generator</td>
<td>Customer</td>
<td>Non-customer</td>
</tr>
<tr>
<td>Data Target</td>
<td>Internal</td>
<td>External</td>
</tr>
<tr>
<td>Data Origin</td>
<td>Customer</td>
<td>Non-customer</td>
</tr>
<tr>
<td>Analytics Type</td>
<td>Descriptive</td>
<td>Diagnostic</td>
</tr>
<tr>
<td>Portfolio Integration</td>
<td>Stand-alone solution</td>
<td>Wrapped around product</td>
</tr>
<tr>
<td>Service User Role</td>
<td>Recipient</td>
<td>Provider</td>
</tr>
<tr>
<td>Dimension</td>
<td>Definition</td>
<td>Contributing literature</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Data Generator</td>
<td><strong>Who or what generates the service-relevant data?</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Customer: Direct client of the service (including customers from B2B2C).</td>
<td>Loebbecke and Picot</td>
</tr>
<tr>
<td></td>
<td>Non-customer: Humans not considered customers as defined above.</td>
<td>(2015); McAfee and</td>
</tr>
<tr>
<td></td>
<td>Process: Structured activities or tasks performed by people or devices in</td>
<td>Brynjolfsson (2012);</td>
</tr>
<tr>
<td></td>
<td>Object: Physical objects equipped with sensors.</td>
<td>Porter and Heppelmann</td>
</tr>
<tr>
<td></td>
<td>sequences which are digitally logged.</td>
<td>(2014); Troilo et al.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2017)</td>
</tr>
<tr>
<td>Data Origin</td>
<td><strong>Where is the data infusion channel located?</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Internal: Data comes from within the service providing company and no</td>
<td>Hartmann et al. (2016);</td>
</tr>
<tr>
<td></td>
<td>third party providers are involved.</td>
<td>Lim et al. (2018a);</td>
</tr>
<tr>
<td></td>
<td>External: The service-relevant data comes from an involved third party</td>
<td>Woerner and Wixom (2015)</td>
</tr>
<tr>
<td></td>
<td>outside of the service providing company.</td>
<td></td>
</tr>
<tr>
<td>Data Target</td>
<td><strong>About whom or what does the generated data contain information?</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Customer: Direct client of the service (including customers from B2B2C).</td>
<td>Huang and Rust (2013);</td>
</tr>
<tr>
<td></td>
<td>Non-customer: Humans not considered customers as defined above.</td>
<td>Kumar et al. (2013);</td>
</tr>
<tr>
<td></td>
<td>Object: Physical objects equipped with sensors.</td>
<td>Loebbecke and Picot</td>
</tr>
<tr>
<td></td>
<td>Process: Structured activities or tasks performed by people or devices in</td>
<td>(2015); Porter and Heppelmann (2014)</td>
</tr>
<tr>
<td></td>
<td>Environment: Surrounding of the respective data generator.</td>
<td></td>
</tr>
<tr>
<td>Analytics Type</td>
<td><strong>Which kind of analytics is applied in the service?</strong></td>
<td>Chen et al. (2012);</td>
</tr>
<tr>
<td></td>
<td>Descriptive: “what happened?”</td>
<td>Davenport (2013);</td>
</tr>
<tr>
<td></td>
<td>Diagnostic: “why did it happen?”</td>
<td>Delen and Demirkan (2013)</td>
</tr>
<tr>
<td></td>
<td>Predictive: “what will happen?”</td>
<td>Han and Kamber (2006);</td>
</tr>
<tr>
<td></td>
<td>Prescriptive: “what should be done?”</td>
<td>Porter and Heppelmann</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2015)</td>
</tr>
<tr>
<td>Portfolio</td>
<td><strong>How is the service related to existing products or services?</strong></td>
<td></td>
</tr>
<tr>
<td>Integration</td>
<td>Stand-alone solution: Service is offered separately</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wrapped around product: Service enriches an already existing product.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wrapped around service: Service enriches an already existing service.</td>
<td></td>
</tr>
<tr>
<td>Service User Role</td>
<td><strong>What role does the service user play in company/customer interaction?</strong></td>
<td></td>
</tr>
<tr>
<td>User Role</td>
<td>Recipient: Service user merely consumes the service; no further</td>
<td>Chen et al. (2011);</td>
</tr>
<tr>
<td></td>
<td>involvement. Provider of data: Service user provides the data the service</td>
<td>Grönroos and Voima</td>
</tr>
<tr>
<td></td>
<td>requires. Interactor: Service user actively interacts with provider</td>
<td>(2013); Lim et al. (2018a);</td>
</tr>
<tr>
<td></td>
<td>by integrating the service into his own processes.</td>
<td>Saarijärvi et al. (2014)</td>
</tr>
</tbody>
</table>

Table 4. Taxonomy definitions and contributing literature.

It can be split into two primary types: Stand-alone and integrated. A stand-alone solution refers to analytics-based services that are offered separately and do not require existing products or services to operate. Accounting for analytics-based services that are integrated into an existing product or service, we aimed for a more detailed distinction. In their research, Wixom and Ross (2017) describe how companies use data and analytics to enrich products or services, wrapping analytics-based services around them. Following these findings, we split integration into wrapped around product and wrapped around service; referring to services that enrich existing product or service offerings.

**Service User Role** describes the role that the service user plays in provider-customer interactions related to an analytics-based service. Our research unveiled three different roles the service user assumes in this context. First, we identified the recipient role, i.e. service users that merely receive and consume analytics-based services. Yet, building on Grönroos and Voima (2013) on value creation and co-creation, we also notice analytics-based services in which both service user and provider contribute to the creation of value. In a second possible role, the service user actively contributes as he takes over the role of a provider of data. He shares necessary data that is essential for the provision of the service itself. Third, taking an interactor role, the service user actively interacts with the service provider by integrating the
analytics-based service into his own processes. In this case, the analytics-based service creates value by e.g. enabling automated tasks, manual interventions or self-service options.

### 4.3 Evaluation and Application

We measured the inter-coder agreement within the expert group using hit ratios (Nahm et al. 2002). We counted agreement among the four experts for each characteristic within the dimensions. We then summarized the single hit ratios to receive the dimension-specific hit ratio. The experts achieved dimension-specific hit ratios of at least 80%. Moreover, three dimensions reached or exceeded hit ratios of 93%. Table 3 provides an overview of the dimension-specific hit ratios. On average, our taxonomy achieved an inter-coder agreement of 89% across the six dimensions.

This result lend credibility on the taxonomy’s ability to classify analytics-based service consistently among researchers. It furthermore indicates that analytics-based services can be described concisely using our taxonomy. Nevertheless, the evaluation also unveiled some of the taxonomy’s weak spots that need further attention in future research. In particular, a distinct classification with regard to the service user role seemed to cause difficulties among the experts. This dimension revealed the lowest hit ratio (80%).

To illustrate the usefulness of our taxonomy in describing analytics-based services and to make the classification of concrete use cases more transparent, we provide two examples from our dataset. ThyssenKrupp, a manufacturing company introduced an analytics-based service that is wrapped around its elevators (Fig. 2; Thyssenkrupp Elevator AG (2017)). Leveraging data that is generated by integrated sensors and telematics, the product continuously produces data about itself – in particular about its operating performance and components’ condition – giving service technicians an up-to-date overview of the elevator’s condition at all times. The data is directly made available to Thyssenkrupp’s databases. Using predictive analytics, Thyssenkrupp forecasts system failures and thus creates an efficient, predictive maintenance service extending its product portfolio. Although the elevators are connected to Thyssenkrupp’s databases, the service user, e.g. facility operators, is required to actively provide the data to Thyssenkrupp in order to enable the service. Therefore, the service user acts as a data provider for the service.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Hit Ratio [%]</th>
<th>Dimension</th>
<th>Hit Ratio [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Generator</td>
<td>88</td>
<td>Analytics Type</td>
<td>94</td>
</tr>
<tr>
<td>Data Target</td>
<td>93</td>
<td>Portfolio Integration</td>
<td>93</td>
</tr>
<tr>
<td>Data Origin</td>
<td>88</td>
<td>Service User Role</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 5. Evaluation results.

To illustrate the usefulness of our taxonomy in describing analytics-based services and to make the classification of concrete use cases more transparent, we provide two examples from our dataset. ThyssenKrupp, a manufacturing company introduced an analytics-based service that is wrapped around its elevators (Fig. 2; Thyssenkrupp Elevator AG (2017)). Leveraging data that is generated by integrated sensors and telematics, the product continuously produces data about itself – in particular about its operating performance and components’ condition – giving service technicians an up-to-date overview of the elevator’s condition at all times. The data is directly made available to Thyssenkrupp’s databases. Using predictive analytics, Thyssenkrupp forecasts system failures and thus creates an efficient, predictive maintenance service extending its product portfolio. Although the elevators are connected to Thyssenkrupp’s databases, the service user, e.g. facility operators, is required to actively provide the data to Thyssenkrupp in order to enable the service. Therefore, the service user acts as a data provider for the service.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Characteristics</th>
<th>E/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Generator</td>
<td>Customer, Non-customer, Processes, Objects</td>
<td>N</td>
</tr>
<tr>
<td>Data Target</td>
<td>Internal, External</td>
<td>N</td>
</tr>
<tr>
<td>Data Origin</td>
<td>Customer, Non-customer</td>
<td>N</td>
</tr>
<tr>
<td>Analytics Type</td>
<td>Descriptive, Diagnostic, Predictive, Prescriptive</td>
<td>E</td>
</tr>
<tr>
<td>Portfolio Integration</td>
<td>Stand-alone solution, Wrapped around product, Wrapped around service</td>
<td>E</td>
</tr>
<tr>
<td>Service User Role</td>
<td>Recipient, Provider, Interactor</td>
<td>E</td>
</tr>
</tbody>
</table>

Figure 2. Classification ‘Thyssenkrupp’ case.
SAP helps the German Soccer Association to increase its performance during soccer matches (DFB 2017). Using sensors placed in the player’s socks, data is continuously generated that reveals insights about the process of the match on an individual player’s level (e.g. passes, player movements, distance covered). The collected data is stored on a platform where SAP can directly access it. Aggregated descriptive information about the matches (e.g. each player’s metrics) are then provided to the coach-team allowing them to analyze the match to immediately improve tactics. Thus, the German Soccer Association takes an interactor role as it deeply integrates the service into its own processes.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Characteristics</th>
<th>E/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Generator</td>
<td>Customer</td>
<td>Non-customer</td>
</tr>
<tr>
<td>Data Target</td>
<td>Internal</td>
<td>External</td>
</tr>
<tr>
<td>Data Origin</td>
<td>Customer</td>
<td>Non-customer</td>
</tr>
<tr>
<td>Analytics Type</td>
<td>Descriptive</td>
<td>Diagnostic</td>
</tr>
<tr>
<td>Portfolio Integration</td>
<td>Stand-alone solution</td>
<td>Wrapped around product</td>
</tr>
<tr>
<td>Service User Role</td>
<td>Recipient</td>
<td>Provider</td>
</tr>
</tbody>
</table>

Figure 3. Classification ‘German Soccer Association’ case.

5 Discussion

5.1 Theoretical and Managerial Implications

Introducing a thoroughly developed, reliable taxonomy of analytics-based services entails immediate implications for research and practice. Our taxonomy illustrates in a simple yet powerful way the conceptually grounded and empirically validated characteristics of analytics-based services and groups them into six dimensions that play a key role in describing these services. Simple, because a set of six key dimensions is sufficient to individually describe each analytics-based service; powerful, because our taxonomy allows researchers to describe analytics-based services in a consistent manner and to distinguish them from each other. Thus, our taxonomy contributes to the existing body of knowledge on using data and analytics to advance service by establishing a common understanding of analytics-based services that has been missing so far. Adding descriptive knowledge on analytics-based services results in a shared language that contributes to structuring the field of research, thereby helping researchers to position their work within this field. In particular, our results help to build a better understanding of the six key dimensions and the characteristics they entail. Most importantly, a common understanding based on our taxonomy allows the materialization of ideas and considerations among scholars leading towards the development of a deeper theorizing process on analytics-based services. The taxonomy is theoretically sound, empirically verified, and offers a comprehensive perspective on the phenomenon of analytics-based services. In sum, our results provide deeper insights into the structure of analytics-based services and will help to systematize and synthesize research in this field. We therefore encourage researchers to build future work on the taxonomy presented in this study.

Second, a common understanding of analytics-based services also gives rise to implications for managers. Despite the taxonomy’s simplicity, it may prove to be highly effective in identifying the strengths and weaknesses of the analytics-based service they are responsible for. By recognizing which characteristics their own service shares with other services, even in seemingly independent industries, managers can use the taxonomy to look beyond their immediate competitors for new ideas. Thereby, the application of the taxonomy allows for the strategic differentiation of services in the market. For this purpose, the proposed
key dimensions may be used as lenses for the evaluation of analytics-based services. Practitioners may benefit from our taxonomy which provides them with initial guidance in the materialization of ideas and considerations in new service development. For this purpose, our taxonomy could serve as a basis for applying a morphological analysis for the systematic generation of ideas (Geum et al. 2016). New service concepts could then be easily defined using the predefined characteristics within each dimension.

5.2 Limitations

The research process poses a number of limitations which leaves, in turn, potential for future research. The adapted taxonomy development method follows a design science philosophy building an effective solution to an identified problem (Hevner et al. 2004). It does not necessarily provide an optimal solution and independent studies may generate different results. Also, we are aware, that the results from our evaluation are limited in their generalizability. Applying the taxonomy to a set of cases that were used for building the taxonomy only indicates its usefulness to classify analytics-based services consistently and future research should re-evaluate and adjust our taxonomy. Regarding our collected dataset of use cases, we relied on customer references that were publicly available on websites of software vendors that helped individual companies to build analytics-based services. We believe that we were able to capture a sufficient set of diverse use cases from different industries and regions. Yet, there are still many vendors that are not considered in this study and it may be beneficial to take into account customer references from smaller, more specialized software vendors to identify use cases of analytics-based services. This may lead to an increased diversity in the dataset, for instance by increasing the number of small companies offering analytics-based services in the dataset.

6 Conclusion and Outlook

Building on a well-established method in information systems literature introduced by Nickerson et al. (2013), we developed a taxonomy of analytics-based services. Overall, we conducted four iterations, one being conceptually based on a structured literature review and three iterations being empirically grounded on a heterogeneous set of 85 analytics-based service use cases. The sample consists of companies from different regions and of different sizes, covering eleven different industries. Our taxonomy of analytics-based services consists of six dimensions – data generator, data origin, data target, analytics type, portfolio integration, and service user role – each represented through a distinct set of characteristics, providing a means to conceptualize analytics-based services as a phenomenon. The evaluation of the taxonomy indicates its reliability to classify and distinguish cases of analytics-based services across researchers. Thus, following Nickerson et al. (2013), the taxonomy proved to be useful and it fulfills its purpose by providing conceptually grounded and empirically validated characteristics of analytics-based services. The taxonomy contributes to existing literature by extending the descriptive body of knowledge on analytics-based services. It establishes a common understanding among researchers, and thus enables them to analyze analytics-based services in a structured manner. Additionally, we provide practitioners with a means to conceptualize analytics-based services and with initial guidance in new service development. We also see potential for our taxonomy to serve future research, as it lays the foundation for a deeper theorizing process on the nature of analytics-based services. Having a reliable taxonomy at hand, future research may focus on collecting more cases of analytics-based services. Building on a larger sample would allow a more detailed analysis of analytics-based services and their characteristics. For example, using the taxonomy to classify cases along the six key dimension presented in this study, cluster analysis could identify configurations of characteristics that are common among analytics-based services leading to higher-order constructs such as analytics-based service archetypes (Greenwood 1993). Such archetypes could help to further theorize how data and analytics is used in customer-facing services to help customers
make better decisions.

References


Hunke et al. / Taxonomy of Analytics-based Services


