Capturing Value from Data: Revenue Models for Data-Driven Services

Ronny Schüritz
Karlsruhe Institute of Technology
ronny.schueritz@kit.edu

Stefan Seebacher
Karlsruhe Institute of Technology
stefan.seebacher@kit.edu

Rebecca Dorner
Karlsruhe Institute of Technology
rebecca.dorner@student.kit.edu

Abstract

Undisputedly, the amount of data is growing exponentially and huge opportunities exist to exploit them. New service business models are being built around value propositions based on data and analytics. Suitable revenue models need to reap the benefits of these value propositions. However, the question of how to best turn a value proposition into revenue for data-driven services is not systematically addressed in literature.

We provide an overview of possible revenue models for data-driven services. Based on a sample of 100 start-ups, we apply qualitative analysis to identify different revenue models for newly established data-driven services such as subscription, gain sharing and multi-sided revenue models.

This paper will contribute to the fundamental understanding of how companies can capture value from data-driven services. It should give guidance on the design and selection of appropriate revenue models and, thus, inspire new forms of revenue generation from the use of data.

1 INTRODUCTION

The amount of data is more than doubling every two years and is predicted to reach 44 zetabytes (44 trillion gigabytes) in 2020 [1]. It is widely acknowledged that reaping the value of data and analytics will be a key source of competitive advantage in the future [2].

One of the main motivations of companies to invest in analytics projects is to develop new business models [3]–[5] and, thereby, bring to bear entirely new "data-based" or "data-driven" business models [6], [7]. These new data-driven business models create value for customers through the generation, aggregation and analysis of data [6]. In addition to value creation, capturing the value generated via an adequate revenue model is a crucial part of a successful business model [8]–[12]. The connection between the survival of a company and a viable revenue model is illustrated when looking at the burst of the dot.com bubble, where one reason for companies to go bankrupt was relying on defective revenue models [13].

In the following, we provide a qualitative analysis that explores possible revenue models for data-driven services. We have based our analysis on a sample of 100 start-up companies that exclusively provide data-driven services. Using a hybrid coding technique, we identified a selection of revenue models and characteristics that are repeatedly used by these companies. In this paper we will describe the identified revenue models in detail, illustrate each of them using a representative case and close with recommendations for potential linkages between revenue models and generic types of data-driven services. Thus, this paper aims to contribute a comprehensive set of revenue models for data-driven services and at the same time to support practitioners in the choice of a revenue model.

The paper is structured as follows: In section 2, we will provide theoretical foundations to the understanding of revenue models and data-driven services. Section 3 covers how we collected 100 use cases and how we approached the analysis from a methodological perspective. Section 4 describes the identified revenue models in detail and illustrates each revenue model in combination with a use case from the sample. Section 5 discusses the observations and provides guidance for companies that want to engage in data-driven services. Section 6 briefly summarizes the results, provides implications for managers, reveals limitations of the study and develops an agenda for further research.

2 THEORETICAL FOUNDATIONS

2.1 Analytics

The foundations for today's field of analytics were laid in the 1990s when both statistical methods and data mining techniques became increasingly popular for analyzing data. Starting with the analysis of mostly structured data in relational database systems, analytics has so far undergone a major development [14].

The advent of new sources of data, such as the social web or cyber-physical systems, as well as changes in the digital lifestyle of people have led to an ever-increasing amount of data that is proliferated worldwide. Although a large proportion of generated data is not suited to be used in analytical processing, usable data for analysis is to grow from 20% in 2013 to 35% in 2020 [1]. The immense growth and availability of produced data is closely connected to the term "Big Data", comprising the features "volume", "variety", "velocity", "veracity", "variability" and "value" [15]. Although the term is widespread, it is used somewhat objectively: Whether data is labeled "Big

URI: http://hdl.handle.net/10125/41810
ISBN: 978-0-9981331-0-2
CC-BY-NC-ND

Data" depends on both the relative point in time and the experience of a firm in coping with large amounts of data. What was considered a data challenge a couple of years ago may today be well solved through advances in ICT. Since the particular understanding of "Big Data" depends on the individual viewpoint, we will continue to simply use the term "data" instead of "Big Data" in the following [7].

Although analytics has gained attention within recent years, the field of analytics still lacks of a uniform definition [16]. In fact, while there is no agreed-upon definition of analytics, some may use "analytics" and "data mining" interchangeably [17] or use "analytics" as a synonym for "business intelligence" [18]. It is further not uncommon to distinguish "basic analytics" and "advanced analytics" [19], [16]. Following this view on analytics, we refer to it as processing of large amounts of data through the use of methods from statistics and operations research in order to derive descriptive, predictive, or prescriptive decision support [20].

2.2 Data-Driven Services

Although the use of data and analytics is viewed to offer new ways of growth and competitive advantage [18], there is only little progress in integrating analytics within service systems [16]. Referring to Davenport [21], the use of analytics has developed from a traditional - mostly internal - decision support perspective to a stage where analytics helps creating additional value for customers by enriching products and services through the exploitation of data.

The rising supply of data opens up opportunities for the creation of entirely new (data) services [4]. Chen et al. [14] conceptualize two general areas of application using data and analytics, namely "Data-as-a-Service" (DaaS) and "Analytics-as-a-Service" (AaaS). Whereas the former focuses on providing raw and aggregated content, the latter denotes services that employ a rich set of common analytics components and infrastructure, adjusted to fit industry- and company-specific requirements.

Some authors claim that since data-driven services position data and analytics at the center of their service offerings, novel "data-driven" or "data-based" business models arise. As the central feature of these business models is that data is considered to be their key resource [4], [6], the question comes to mind: Is there a specific or distinct point where data becomes the key resource?

There is a fluent transition from existing to "data-driven" business models. Incentivized by the unused potential which data and analytics have to offer, more and more companies will include data and analytics in their business models, gradually evolving to more "data-driven" business models that offer some kind of data-driven service or infusing their current service offering [22]. At the same time we can see entirely new business models come to life that use data as their key resource and utilize the possibilities of analytics [6], [14]. We call the offerings of these business models *data-driven services*.

2.3 Revenue Models

A large number of publications have shown the rising interest of researchers and practitioners for business models [23]. The business model is a "heuristic logic" [8], "an architecture" [24], "a representation" [11]," or "model" [12] that articulates what the business has to offer and how customers are benefitting from using it [13], how the business "creates value through the exploitation of business opportunities" [23] and how this value is then captured and turned into profit.

Consequently, while sometimes confused with the business model [25], the "revenue mechanism" [13], "revenue model" [26] or "profit formula" [27] is a crucial element of the business model in most representations [8], [10], [11], [26]. It describes how a business model generates revenue through the provision of its service. The notion to understand the mechanics of revenue models more deeply rose from the interest in the business model [13], [26]. This was also caused by the fact that e-business created new revenue models that did not exist before. The end of the dot.com era illustrated the importance of equipping viable business models with suitable revenue models: Many companies failed to develop revenue models to turn the value they delivered to their customers into adequate revenue streams [13].

Dependent on the level of detail with which the concept of a revenue model is described, there exist different approaches for its definition [28]. Amit and Zott [29] define a revenue model's purpose as to illustrate "the specific modes in which a business model enables revenue generation". In this sense, Osterwalder [26] further details the definition by specifying a revenue model as a "logic of what, when, why and how" delivered value is transformed into earnings. Furthermore, he states that a revenue model may consist of different revenue streams which in turn can rely on varying pricing mechanisms hence a distinction between a revenue and a pricing model can be made [30]. As a revenue model comprises the economic activities which are applied to turn a company's service into revenue [26], a pricing model aims at deriving the market price for a respective good or service, using pricing strategies and price finding mechanisms, while meeting certain boundary conditions (e.g. regarding a firm's strategy, customer segments) [31].

Wixom [32] states that data can be monetized by either solely selling the data, bartering it in exchange for a product or service or wrapping it around a product or service. However, so far, hardly any research has been conducted addressing the actual revenue models of data-driven services in particular. Nonetheless, more general approaches investigating revenue models are available. There are different criteria by which revenue models can be distinguished such as by the product or service which is being sold [33], by the role of the customer [34], by the way the value is determined [35] or whether the revenue is generated directly or indirectly [36]. Especially in non-academic literature one can find many more detailed distinctions and listings of revenue models [37]. However, many of the listed revenue models are lacking clear

differentiations and, therefore, cannot be considered disjoint. Some show a mixing of revenue and pricing models, or simply are missing empirical evidence. While there is limited research on our particular focus, revenue models for data-driven services, we may draw upon revenue model work that has been done in the area of e-commerce research. Based on the performed structured literature review based on vom Brocke et al. [38] the following categorization of revenue models can be synthesized.

The most common and known revenue model is *asset sale*, where the ownership right of a product is transferred in return for money [9]. This is the case for many transactions and it is the common revenue model which can be found in retail, where a consumer buys fast-moving consumer goods. *Usage fee* refers to the revenue model in which a service is provided to a customer who in return pays for it. The higher the customers' usage of the service, the higher the fee they need to pay. This kind of revenue model can typically be found in service settings like hotels [9]. Both of these revenue models were also often combined into a so-called *single transaction* by different authors [30], [33]. They define this as selling a product or a service [33] or solely referring to products, missing to acknowledge the intangibility of a service.

If providers choose a *subscription model*, they sell the continuous access to a service and the customer pays as long as the subscription is valid [9], [33], [34], [39]. This may also be called a membership fee depending on the context.

In a *lending*, *renting* or *leasing model* someone is granted permission to exclusively use an asset for a defined period of time like renting a house or leasing an automobile, whereas in *licensing*, another revenue model, the user is granted the permission to use protected intellectual property. This right can be sold to numerous customers simultaneously like the right to use a patented technology [9].

Advertising as a revenue model is based on fees for advertising products, services, or brands. It has been a popular revenue model on the internet and has been accounting for the major portion of income of most internet companies in the past [40].

A *brokerage fee, commission or transaction cut* is a fee paid to a third party that facilitates the match for a successful transaction between two parties. Matchmaking platforms or real estate agents often rely on that revenue model to generate income [9].

Some authors suggests [35], [41] in a specific context (e.g. IT Services, procurement, etc.) a *gain sharing model* in which the provider is paid based on the gains (e.g., savings) the customer can realize by using that service.

Free is sometimes mentioned as a revenue model especially on the internet, where services are provided to customers free of charge [34]. Nevertheless, as a company needs to create revenues, it needs to have other revenue streams from which it generates adequate earnings [42]. Therefore, free is not self-sufficient and only works in conjunction with another revenue model.

3 METHODOLOGY

To explore the extent to which revenue models of pure data services differ from general revenue models, we followed a rigorous process of qualitative analysis.

In the following section 3.1 we specify how we collected a random sample of 100 start-up companies that offer a pure form of data-driven services and their respective revenue models. This is followed by a description of the coding mechanism and process that has been applied to analyze the collected data set (section 3.2).

3.1 Source of data

Our collected dataset consists exclusively of pure data-driven services including their respective revenue models. As mentioned before, we argue that there is a fluent transition from existing to data-driven business models. By analyzing pure data-driven services, whose business models thoroughly depend on data as a key resource, we avoid any influence from factors of revenue models for product-service bundles.

We focus on start-up companies using pure data-driven services to reflect that these services are at an early stage of development and application [16] and that start-ups tend to be the first ones to leverage the opportunities of novel technologies [43].

For the purpose of accessing information on start-up companies, we leverage the database of the company AngelList (www.angellist.com). Among other things, AngelList provides an online platform, enabling both start-up companies to raise money and investors to invest into beneficial business concepts. Start-ups can advertise their projects using profiles on the platform, thereby releasing information about their company and proposal. Additionally, companies categorize themselves by indicating their thematic focuses using tags [44].

As a first step to identify companies that fit the definition of pure data-driven services, the entire set of tags is assessed. Carrying out group discussions with four researchers who share a sound understanding of the topic, the list of keywords is reduced to the search keys displayed in Table I. Since companies are not limited in the amount of tags they use to describe their proposals and therefore may use more than one of our chosen tags, our selection is expected to have a minimal overlap, while simultaneously covering a high proportion of proposals in the area of pure data-driven services.

As a second step, AngelList's database is searched for each of the identified keywords. During the search the platform's implemented search algorithm evaluates and sorts results according to relevance, displaying the 400 most important start-ups. Thus, each search results in a list of 400 duplicate-free entries. In order to ensure mutual exclusiveness across lists, duplicates are eliminated resulting in an overall sample of 1552 proposals (cf. Table I). Since duplicates are removed randomly, the number of entries per keyword varies in size.

In order to avoid any interference with data and the respective selection process, we apply random selection to determine our final sample. Although random sampling might

not reveal the entire information on a phenomenon, potentially leaving out rich or important data [45], it helps to overcome systemic bias, also known as the researcher bias [46].

Table I. Entries per keyword

Keyword	Number of Cases
Data Mining	386
Data Services	342
Big Data	275
Business Analytics	319
Data Analytics	230
Sum	**1552**

Each of the five lists is randomized separately before the start of the final selection. Starting from the top of each randomized list, a start-up is selected in case it meets our understanding of data-driven services, otherwise it is dismissed. This procedure is conducted until 20 relevant startups out of each list are selected, resulting in a final set of 100 companies.

For each of the 100 start-ups, AngelList lists a description provided by the start-up itself. In order to enrich our data basis and to ensure a sound foundation additional information on the companies, their business proposals as well as on their revenue models is obtained, using the homepage of the start-up itself and publicly available sources such as TechCrunch (http://techcrunch.com).

For 39 of the cases neither the provided description nor the homepage itself nor other websites provided sufficient information on the revenue model. In these cases, the research team has contacted the start-up via email to receive additional information on their revenue model. Six of the start-ups replied, of which four provided sufficient information, resulting in a useful data set of 65 start-ups.

3.2 Coding mechanism

In order to perform a textual analysis of our collected data, consisting of several sources such as short descriptions from AngelList's platform, start-up websites, and tech news websites, a hybrid coding approach is applied.

Since there is a rich variety of coding techniques in qualitative research, selecting the appropriate coding method depends on the pursued research goals. For example, in case there exists a conceptual framework as the underlying basis of a research inquiry, Saldaña [47] recommends the use of provisional lists of codes. If theory development is intended instead, coding mechanisms such as "Open Coding" should be used [48]. Furthermore, it is possible to combine different coding methods. With his "paradigm of choice", Patton [49] refers to a pragmatic, but reasonable choice of methods as well as their interplay.

On the one hand, as revenue models in general have already been assessed to some extent by fellow researchers, we may apply provisional coding, starting with a given set of codes. Following Saldaña [47], the provisional codes are identified through the investigation of related work. On the other hand, the role of revenue models in data-driven services in particular has not yet been subject to research. Since our research goal is to explore new aspects or even entirely new types of revenue models with regard to data-driven services, a more open and elementary method (Initial Coding) will be used complementarily. The combination of both Provisional and Initial Coding forms our hybrid coding approach in order to fulfil our research objectives.

Synthesizing our literature review on revenue models, the start list of codes on revenue models consists of the following codes: Asset sale, lending/ renting/ leasing, licensing, advertising, brokerage/ commission, usage fee, and subscription. After developing this first set of codes, two individual researchers (coders) start to analyze the collected data in order to identify relevant information on revenue models by either placing existing codes or developing new codes. A computer-based coding process is applied, using the coding software MAXQDA (http://www.maxqda.com/).

The act of coding and thereby the choice of codes is dependent on the individual researcher, since it is an analytical and interpretative process which is based on the perceptions of a single person [50]. In order to minimize any bias resulting from subjectivity while enhancing validity and rigor of this study, continuous discussion sessions between the two researchers involved in the coding process are arranged. Disputes are solved by mutual discussion or by involving a third researcher.

At the end of the coding process the results are again discussed and consolidated in order to identify critical elements of revenue models for data-driven services and to derive typical patterns of such revenue models.

4 Revenue Models in Data-Driven Services

Following the coding approach laid out in the previous sections, the analysis of start-ups that provided data-driven services leads to a set of four distinct revenue models, illustrated in figure 1.

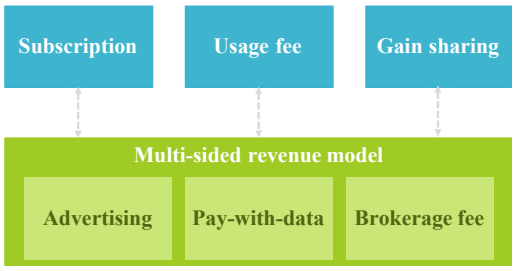

Figure 1 Observed revenue models in data-driven services

Every observed start-up applies at least one of the revenue models such as *subscription*, *usage fee*, or *gain sharing*. Some even have more than one revenue model in place as they offer different data-services to different target customers. Some

observed companies have a *multi-sided* revenue model: The start-up offers its service to at least two different target customers which, however, are interrelated to each other. One of the target groups is offered the service in return for bearing some advertisement or giving away its data while the other group is offered one of the basic revenue models.

The following sections will describe each revenue model in detail and elicit what kind of services typically draw upon a particular revenue model. An exemplary case will illustrate each revenue model.

4.1 Subscription Model

Subscription has been identified as the most popular revenue model within our sample. With 44 of 65 analyzed start-ups it is used as the main revenue stream for data-driven services. In a subscription model the customer pays a fee in a periodical way to have access to a service. In our dataset, we observed a tendency for short term contracts, which were mostly extended on a monthly basis. However, we have also observed annual models.

Most subscription models are also coined by some form of feature differentiation model. That means the customer can choose between different subscription models that vary in their respective features which are included within a service (e.g. "basic" vs. "professional"). We have witnessed differentiation either based on functionality or on volume. Regarding *functionality limitations*, many data-driven services limit the functionality in a basic tier by not offering support, cutting back on offered analytical methods, or do not offer APIs to other services. Concerning *volume-based limitations*, the companies in our sample establish some kind of tier model with regards to the request the user can make to the service, the amount of data which can be processed, the number of devices (e.g., phones, sensors, etc.) that can be added to the service, the number of users that can use the service or the number of customers that can be analyzed with the service. In some instances, the most basic tier of the subscription model is free of charge, offering the service with some of the above described limitations. This aims at attracting customers through word-of-mouth with some of them willing to upgrade to a higher tier. Such practices are usually referred to as a freemium approach [51].

Companies which are tying their offers to a subscription model can benefit from the continuous data collection as well as from using the gathered data to improve the provided service across customers. Companies that use subscription as their revenue model usually offer services in which the customers benefits from the continuous collection and analysis of data. This benefit is only created for the period that the service is used. In our sample, we have two distinctions in which subscription is chosen as the revenue model: (1) Collection and analysis of individual business data that is created through the service delivery of the business customer in order to track, monitor, and optimize performance (e.g., customer behavior data) and (2) collection of generally available data that is constantly being provided to the customer (i.e. trending news articles).

Both types of data outdate over time after their collection but they differ in the way the customer benefits from them. The more individual data is collected and analyzed, the more value is generated for the customer which leads to the offering of volume-restricted tiers. In case focus is set on the provision of general data, access to such a database may be useful for a variety of users, therefore a respective subscription model may be account-restricted. The offered subscription tier is thus depending on the kind of benefit the customer will receive.

One example for a data-driven service that uses subscription as their main revenue model is the company *AmigoCloud*. *AmigoCloud* is a mapping technology company that provides a mobile geographic information system solution that helps organizations and individuals to collaboratively collect, edit, visualize, publish and analyze geospatial data. *AmigoCloud* offers a subscription model which can be either purchased on a monthly or yearly basis offering a certain discount. The subscription model relies on four tiers: individual, small office, professional and enterprise. The tiers are distinguished in volume of users who collaboratively use the software, the number of private projects that can be worked on and the online storage that is provided. In addition, some features that mainly focus on the integration of other datasets are limited to professionals and enterprises.

4.2 Usage fee

With five instances, the usage fee represents another revenue model in our sample. In this revenue model, the customer pays for the use of the service depending on how much he uses it. The usage can be determined by different factors such as volume of data which is to be analyzed, the amount of end-customers or customer visits and the number of queries, channels or requests, as well as number of accesses granted. While not many instances could be identified using a pure usage fee model, almost all subscription models show aspects of this revenue model by limiting the volume or features across tiers.

Usage fee models were chosen for data-driven services that increase the benefit with the volume of data being analyzed or provided (i.e. every additionally tracked and targeted user is additional revenue potential) or for which the benefit lies in the access of the data as it outdates over time.

VoiceBase provides solutions that focus on speech analysis in customer call centers. Based on the inquiries customers state during their calls, *VoiceBase* offers both automatic transcription and a variety of analytical techniques such as semantic analysis, the identification of topics and non-compliances in interaction as well as sales leads. *VoiceBase*'s revenue model charges the user each minute of analyzed speech a specified rate.

4.3 Gain sharing

A gain sharing revenue model is identified in three different start-ups. This rather innovative revenue mechanism is based on

the idea that the provider is paid based on the success of the service they provide, usually picking up a certain percentage of the generated value at the customer end. We have seen this revenue model mostly in combination with a recommender system that engages end-customers to purchase.

A gain sharing model certainly does make sense when the result and created value of a data-driven service are measurable and quantifiable. The benefit is realized as long as the service is used and aims to enhance an existing process.

Predictry is a start-up in our sample that offers a recommender solution to its e-commerce customers in order to increase their sales and conversion rates. By using *Predictry*, the online shop will be extended by a recommender function that makes personalized recommendations to shop visitors. *Predictry* takes advantage of the gain-sharing revenue model thereby tying its revenue to achieved performance. In the course of a successful recommendation, *Predictry* receives 3-6% of sales revenue. There are no limits on number of recommendations or on functionality.

4.4 Multi-sided revenue model

In contrast to the above listed revenue models the combination of at least two mutually dependent revenue streams, which involve at least two different target customer groups, form what we call a *multi-sided revenue model*.

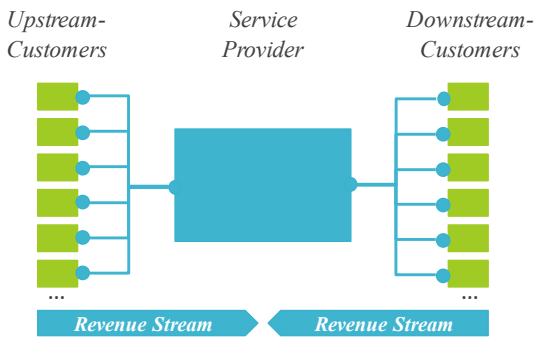

Figure 2 Generalized composition of a multi-sided revenue model

In this context, data and information of one group of customers may establish the foundation for creating another group's revenue stream to the related service provider. Within our sample we were able to identify 15 data-driven services that apply a multi-sided revenue model. In the course of multi-sided revenue models, we apply the terms upstream- and downstream-customers in order to distinguish between two different customer groups (cf. Figure 2). The upstream-customer gives monetary compensation to the provider in order to receive a service that is based on the personal or operational data of the downstream-customer.

A distinction between four different groups of multi-sided revenue models is drawn, all of which consist of an upstream-customer and downstream-customer. Multi-sided revenue models can utilize the described revenue models (subscription, usage fee and gain sharing) but also feature some exclusive models that can only be seen and used in multi-sided revenue models. Within our sample of multi-sided revenue models, downstream-customers can mostly use a service without paying a direct fee. Instead they have the choice to grant access to parts of their private data, which could be viewed as a form of indirect payment. This results in the following four groups of *multi-sided revenue models*:

1. To use the data-driven service, downstream-customers expose themselves to advertisement (*endure-ads*) which is personalized by their data.
2. To receive tailor-made offerings by upstream-customers of the service provider, downstream-customers can choose to give out parts of their private data (*data-tailored-offering*).
3. Private data of downstream-customers is used to attract upstream-customers of the service provider, hence creating a data-marketplace (*buy-and-sell-data*).
4. Downstream-customers agree to share their private data in exchange for access to the service provider's offering (*pay-with-data*).

Within our sample there has not been a single reported case of free-to-use-services, which was not part of one of the described multi-sided revenue models, supporting the arguments of Berman [52] who states that a free service has to be paid by someone somewhere.

4.4.1 Endure-ads

When looking at the endure-ads group, downstream-customers can take advantage of a data-driven service, for which no monetary payment is needed. Instead they expose themselves to advertisement. Therefore, one of the service provider's tasks is to act as an intermediary between its two customer groups. The service provider aims at attracting users (downstream-customer) for its data driven service, since a broad client base is needed to convince its business customers (upstream-customer) to issue advertisements. Revenue may be generated either through one-time advertising fees for each issued advertising or through the establishment of a subscription model (cf. 4.1). To achieve a high business customer satisfaction, resulting in higher earnings, the service provider utilizes data to enable targeted or personalized advertising. In this context, an additional source of income might be established through the integration of a brokerage fee. Since the service provider gets paid a specified amount of money each time a downstream-customer expresses interest in a product of an upstream-company by clicking a respective advertising banner, the matchmaking provider is interested in predicting and allocating the most promising advertisement to its respective downstream-customer.

An endure-ads model with focus on advertising may be chosen by companies whose data-driven service is suitable for generating customer profiles for reinforced advertising messaging.

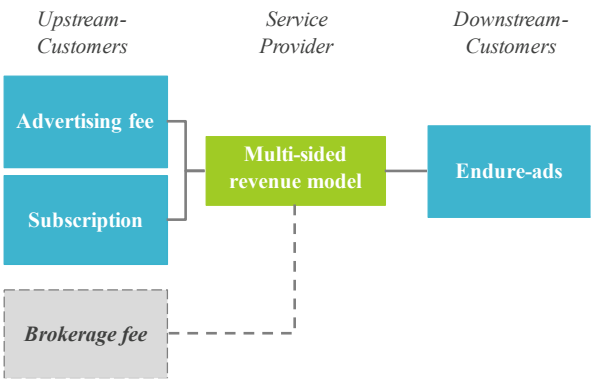

Figure 3 Composition of multi-sided revenue models: *Endure-ads*

Samba TV is exemplary for a start-up, whose data-driven service depends on a multi-sided revenue model focused on advertising. On the one hand, *Samba TV* offers television viewers a platform to engage with TV shows, for example by offering additional content on actors, and to receive recommended viewings free of charge. On the other hand, through the utilization of the platform, user profiles are generated, which in turn are used by *Samba TV* to attract advertisers.

4.4.2 Data-tailored-offering

Within the data-tailored-offering group, downstream-customers grant access to certain parts of their private data for the purpose of tailored offers by upstream-customers. Therefore, the service provider is running a platform at which interaction between its two customer groups is facilitated. Upstream-customers get access to data through a subscription model or a usage fee approach. Again, a brokerage fee may be included in order to generate additional revenue.

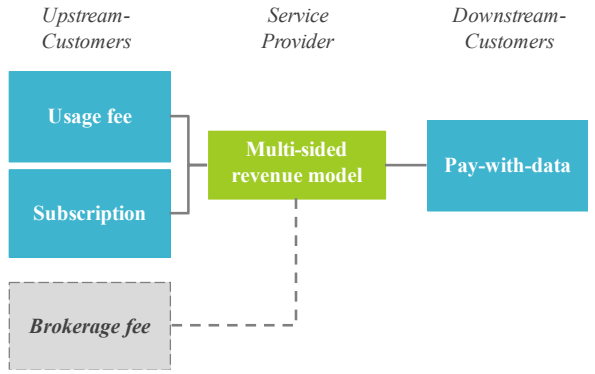

Figure 4 Composition of multi-sided revenue models: *Data-tailored-offering*

In general, adopting the data-tailored-offering approach may be beneficial for a company, whose data-driven service includes a platform or database and is based on highly scalable and repetitive transactions.

The start-up SchoolSparrow provides a platform for searching and comparing real estates in Chicago. Its data-driven service integrates information on the performance of public schools, school districts and available apartments and houses in Chicago. Thereby, its downstream-customers can choose their housings depending on available public schools in a district. Furthermore, *SchoolSparrow* receives a brokerage fee for each successful matching between renter and landlord.

4.4.3 Buy-and-sell-data

In contrast to the previous group of multi-sided revenue models, there is no interaction between an upstream- and a downstream-customer on a service provider's platform. In this case, the service provider acts as a data broker. The provider is engaged by its downstream-customers to find prospective buyers for their data. The service provider is financially connected to its downstream-customers on the basis of a profit-related payment, which might be initialized after each successful matching between its customer groups.

In general companies with a strong expertise in advertising, focusing on the creation of customer profiles, might choose the buy-and-sell-data approach. These companies are rather interested in ad placement than in the creation of advertising content.

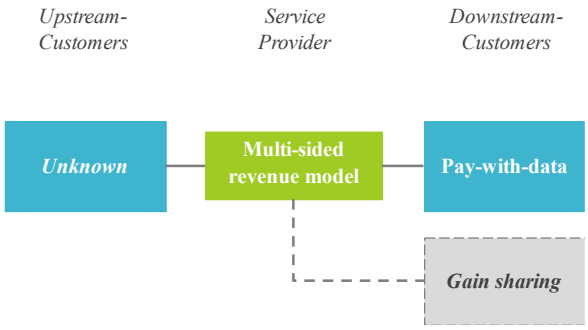

Figure 5 Composition of multi-sided revenue models: *Buy-and-sell-data*

One example of a company who uses such a revenue model is *Zeotap*. Telecom operators engage *Zeotap* with their customer data to find prospective advertisers, increasing their revenue from advertisements. In turn, advertisers get access to a comprehensive dataset, thereby being able to personalize their advertisements. When using *Zeotap's* service, telecom operators agree to take part at a revenue sharing model.

4.4.4 Pay-with-data

In the pay-with-data case, downstream-customers are interested in using the provided data-driven service, while granting access to parts of their personal data. The latter is collected, in some cases analyzed and eventually offered to upstream-customers, who in turn have to pay a subscription or usage fee. Looking at our pay-with-data cases, a direct interaction between upstream- and downstream-customer has not been observed.

In general, applying a pay-with-data approach might be useful for companies, who do not want to rely on integrating advertisements in their services or whose services generate downstream-customer data that is useful if collected from a

large amount of users. The created value is based on the reports that the company can offer to the market through value adding services.

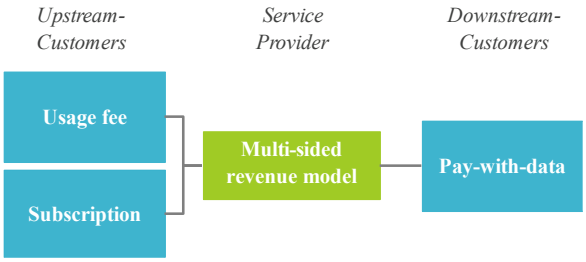

Figure 6 Composition of multi-sided revenue models: *Pay-with-data*

The start-up *GameAnalytics* is an example for a service provider where downstream-customers (game studios), use the "free" service where their provided data is collected and analyzed. Using the collected data, *GameAnalytics* offers aggregated industry reports and insights which can be purchased by upstream customers.

5 Discussion

Our analysis of start-ups that provide data-driven services has given us a deeper understanding on revenue mechanisms of these services. As we analyzed start-ups, some of the analyzed companies have not developed a revenue model yet. This is not surprising as some of these start-ups are in an early phase of developing their business model. Those which are more advanced in the process have developed a wide value proposition that includes data as well as analytics services. In order to capture this value, the regarded start-ups focus on a subset of known revenue models, such as subscription and usage fee, while leaving out other well-established approaches like asset sale, lending, renting or leasing. In our case studies, gain sharing has moved from a so far solely theoretical use to a practical application. While multi-sided revenue models could be observed before, in the context of data-driven services multi-sided revenue models have developed new elements, such as pay-with-data.

In e-commerce it was very common to offer services for free and generate revenue through adding advertisements on homepages. This was observed in only two instances of our sample. Data-driven services do not build upon advertisement as their main revenue stream. Instead companies utilize revenue models such as subscription - a model that is well known from the media industry. This makes sense when considering that the customer who uses the services has the option to use the service via an API which makes it difficult to expose the customer to advertisement.

We were able to identify multi-sided revenue models in which the companies take advantage of more than one target group. In these instances, businesses exploit the fact that data is easily transferable and is of higher value if consolidated or analyzed. In these models companies use the data generated with one target group to generate revenue with another target group. However, while companies can surely utilize more than one revenue stream and model, multi-sided revenue models cannot be seen as a pure complementary to other models, because in multi-sided revenue models the different revenue streams are mutually dependent on each other.

Taking what we have seen from these start-ups, we may give guidance towards the choice of the revenue model for companies that want to pursue data-driven services. Therefore we have included the following tables (cf. Table II, IV), in which we summarize the characteristics of the observed data-driven services and the corresponding revenue models.

Table II. Guide: Basic revenue models

Revenue model	Characteristics of data-driven service
Subscription	Continuous data collection and/ or analysis through data-driven service
	Customer perceives value during period of service usage
	Data needs to be kept up-to-date, since it outdates after a given period of time
	Continuous collection of data can be used to improve service across customers
	Types of data: ▪ Data on individual business performance → *apply volume-restriction* ▪ General information or data → *apply access-restriction*
Usage fee	Perceived value for customer increases with every additional service usage
	Data needs to be kept up-to-date, since it outdates after a given period of time
Gain sharing	Results of service are measurable and can be associated with service execution
	Application of the service to enhances an existing process of the customer

Table III. Guide: Multi-sided revenue models

Revenue model	Characteristics of data-driven service
Endure-ads	Service execution facilitates the creation of customer profiles
	Integration of ads is facilitated during service execution
Data-tailored-offering	Operation of a platform or database is included in the service
	Service is based upon scalable and/ or repetitive transactions
Buy-and-sell-data	Service execution facilitates the creation of customer profiles
	Establishment of new interaction between customers/ customer groups through service
Pay-with-data	Collection of a vast amount of data through service execution
	Service transforms collected data into value adding services

6 CONCLUSION

Our qualitative analysis of 100 start-ups that offer data-driven services has identified a set of revenue models that are commonly used by them: subscription, usage fee, gain sharing and four different kinds of multi-sided revenue models. Each of them is presented with a detailed description and a representative case from the sample.

A subscription model has been found to be the most frequently used revenue model in our sample, but also more innovative approaches such as gain sharing and multi-sided revenue models are gaining traction among start-ups with data-driven services.

6.1 Managerial Implications

Looking at how start-ups that offer data-driven services capture value with distinct revenue models may give rise to immediate implications for managers.

First, our qualitative analysis revealed a set of revenue models for data-driven services. This overview may contribute to a better grounded and systematic discussion of such revenue models. Second, every identified revenue model is described and analyzed with the purpose to expose common traits of the service that chose the respective revenue model. This adds a valuable orientation for enterprises when engaging in data-driven services (cf. Table II, III).

6.2 Limitations

The data source and the method of sampling has exposed the study to a set of limitations.

We rely on AngelList as our source of data for the initial selection process of start-ups. Therefore, our identified companies are restricted to the ones listed in the database. This becomes significant in three ways. First, if a start-up does not approach AngelList for funding it cannot be selected for our sample. Second, the geographical composition of its user base is also limited to the area in which AngelList is known. Third, since companies provide the tags, which we used to identify data-driven services, themselves, there is a chance that we missed start-ups that have not tagged themselves as such.

By choosing a random selection approach for extracting the regarded start-up companies, we minimized researcher bias (cf. 3.1) within the selection process. However, one can argue that by adapting a random selection approach one might not yield a rich and diverse sample.

Finally, all analyzed companies are start-ups. While it can be argued that start-ups may not have a profound business model yet and their revenue model still needs to provide proof of success, pure data-driven services that are not influenced by the decision made for other products or services of a company can particularly be found there. Furthermore, some of the analyzed start-ups have long surpassed their infancy, having built up teams of engineers and already generating revenue on their own. While they still should be treated with cautiousness, some of the start-ups have developed feasible business models proving their viability.

6.3 Future Research

Looking at revenue models of data-driven services opens up a wide range of questions that need to be addressed in further research.

First, in order to further explore revenue models and create a comprehensive list of them, a broader analysis needs to be carried out. In this piece of research, 100 cases were randomly chosen and analyzed, but no additional extreme cases were taken into account. A more diverse sampling method may reveal additional insights.

Second, some of the revenue models need a deeper analysis to understand their dynamics. Multi-sided revenue models seem very appealing as value of data may increase when combining different data sources. Further, it is not yet investigated how pricing mechanisms work for the different models.

Third, our research is focused on revenue models for pure data-driven services of start-ups. It would be interesting to see on the one hand how established organizations, that already offer data-driven service for a longer period of time, and on the other hand organizations that have a wider range of products and services generate revenue with data-driven services.

7 REFERENCES

[1] V. Turner, J. Gantz, D. Reinsel, and S. Minton, “The digital universe of opportunities: Rich data and the increasing value of the internet of things.” IDC, 2014.

[2] D. Opresnik and M. Taisch, “The value of Big Data in servitization,” *Int. J. Prod. Econ.*, vol. 165, pp. 174–184, 2015.

[3] S. Lavalle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, “Big Data, Analytics and the Path From Insights to Value,” *MIT Sloan Manag. Rev.*, vol. 52, no. 2, pp. 21–32, 2011.

[4] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, “Big data: The next frontier for innovation, competition, and productivity.” McKinsey Global Institute, 2011.

[5] L. Kart, N. Heudecker, and F. Buytendijk, “Survey Analysis : Big Data Adoption in 2013 Shows Substance Behind the Hype.” Gartner, 2013.

[6] P. M. Hartmann, M. Zaki, N. Feldmann, and P. A. Neely, “Big Data for Big Business? A Taxonomy of Data-Driven Business Models Used by Start-Up Firms,” 2014.

[7] H. Demirkan, “Innovations with Smart Service Systems : Analytics , Big Data , Cognitive Assistance , and the Internet of Everything,” *Communications of the Association for Information Systems*, vol. 37, 2015.

[8] H. Chesbrough and R. S. Rosenbloom, “The role of the business model in capturing value from innovation : evidence from Xerox Corporation ’ s technology spin-off companies,” *Ind. Corp. Chang.*, vol. 11, no. 3, pp. 529–555, 2002.

[9] A. Osterwalder, Y. Pigneur, A. Smith, and T. Movement, *Business Model Generation*. New Jersey: John Wiley & Sons, 2010.

[10] M. M. Al-Debei and D. Avison, “Developing a unified framework of the business model concept,” *Eur. J. Inf. Syst.*, vol. 19, no. 3, pp. 359–376, May 2010.

[11] S. M. Shafer, H. J. Smith, and J. C. Linder, “The power of

business models," *Bus. Horiz.*, vol. 48, no. 3, pp. 199–207, May 2005.

[12] M. Morris, M. Schindehutte, and J. Allen, "The entrepreneur's business model: toward a unified perspective," *J. Bus. Res.*, vol. 58, no. 6, pp. 726–735, Jun. 2005.

[13] D. J. Teece, "Business Models, Business Strategy and Innovation," *Long Range Plann.*, vol. 43, no. 2–3, pp. 172–194, Apr. 2010.

[14] H. Chen, R. H. L. Chiang, and V. C. Storey, "Business Intelligence and Analytics: From Big Data To Big Impact," *Mis Q.*, vol. 36, no. 4, pp. 1165–1188, 2012.

[15] C. L. Philip Chen and C. Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," *Inf. Sci. (Ny).*, vol. 275, pp. 314–347, 2014.

[16] H. Fromm, F. Habryn, and G. Satzger, "Service Analytics: Leveraging Data Across Enterprise Boundaries for Competitive Advantage," *Glob. Prof. Serv.*, pp. 139–149, 2012.

[17] R. Kohavi, N. J. Rothleder, and E. Simoudis, "Emerging Trends in Business Analytics," *Communcations ACM*, vol. 45, no. 8, pp. 45–48, 2002.

[18] T. H. Davenport and J. G. Harris, "Competing on Analytics: The new Science of Winning," *Harv. Bus. Rev.*, vol. 84, no. 1, pp. 98–107, 2007.

[19] J. Kobielus, "The Forrester Wave^{TM}: Predictive Analytics And Data Mining." Forrester, 2010.

[20] D. Delen and H. Demirkan, "Data, information and analytics as services," *Decis. Support Syst.*, vol. 55, no. 1, pp. 359–363, 2013.

[21] T. H. Davenport, "Analytics 3.0," *Harv. Bus. Rev.*, vol. 91, no. 12, p. 64, 2013.

[22] R. Schüritz and G. Satzger, "Patterns of Data-Infused Business Model Innovation," in *IEEE 18th Conference on Business Informatics (CBI)*, 2016.

[23] C. Zott, R. Amit, and L. Massa, "The Business Model: Recent Developments and Future Research," *J. Manage.*, vol. 37, no. 4, pp. 1019–1042, May 2011.

[24] B. Electronic, C. Conference, A. Osterwalder, and Y. Pigneur, "An e-business model ontology for modeling e-business," *15th Bled Electron. Commer. Conf. June 17-19*, p. 12, 2002.

[25] G. George and A. J. Bock, "The Business Model in Practice and its Implications for Entrepreneurship Research," *Entrep. Theory Pract.*, vol. 35, no. 1, pp. 83–111, 2011.

[26] A. Osterwalder, "The Business Model Ontology - A Proposition in a Design Science Approach," *Business*, vol. Doctor, pp. 1–169, 2004.

[27] M. W. Johnson and C. M. Christensen, *Reinventing Your Business Model*, no. December 2009. Harvard Business School Publishing Corporation, 2010.

[28] T. Bieger, D. zu Knyphausen-Aufseß, and C. Krys, *Innovative Geschäftsmodelle - Konzeptionelle Grundlagen, Gestaltungsfelder und unternehmerische Praxis*. Heidelberg: Springer, 2011.

[29] R. Amit and C. Zott, "Value creation in E-business," *Strateg. Manag. J.*, vol. 22, no. 6–7, pp. 493–520, Jun. 2001.

[30] B. Skiera and A. Lambrecht, "Erlösmodelle im Internet," *Handb. Produktmanagement*, 2007.

[31] S. Klein and C. Loebbecke, "The transformation of pricing models on the web: examples from the airline industry," *13th Int. Bled Electron. Commer. Conf.*, 2000.

[32] B. Wixom, "Cashing in on data," *Research Briefing of the Center for Information System Research*, vol. 14, no. 8. 2014.

[33] P. Timmers, "Business Models for Electronic Markets," *Electron. Mark.*, Jul. 1998.

[34] B. Mahadevan, "Business Models for Internet-Based E-Commerce: An Anatomy," *Calif. Manage. Rev.*, vol. 42, no. 4, pp. 55–69, Jul. 2000.

[35] W. Buchholz, "Netsourcing business models-Geschäftsmodelle für Einkaufsplattformen," in *Die Supply Chain im Zeitalter von E-Business und Global Sourcing*, Paderborn, 2001, pp. 37–52.

[36] B. Wirtz, *Electronic Business*, 2nd ed. Wiesbaden: Gabler, 2001.

[37] Board of Innovation, "27 Revenue Models," 2016. [Online]. Available: http://www.boardofinnovation.com/27-revenue-model-options-for-b2c-b2b-companies/. [Accessed: 09-Jun-2016].

[38] J. vom Brocke, A. Simons, B. Niehaves, K. Riemer, R. Plattfaut, and A. Cleven, "Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process," *Proc. 17th Eur. Conf. Inf. Syst. Verona*, pp. 2206–2217, 2009.

[39] T. Weber, "Software-Referenzarchitekturen für Geschäftsmodelle des e-Business unter besonderer Beachtung ihrer Erlösmodelle.," Universität Leipzig, 2006.

[40] D. S. Evans, "The Online Advertising Industry: Economics, Evolution, and Privacy," *J. Econ. Perspect.*, vol. 23, no. 3, pp. 37–60, 2009.

[41] G. Satzger and A. Kieninger, "Risk-reward sharing in IT service contracts - A service system view," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, 2011.

[42] E. K. Clemons, "Business Models for Monetizing Internet Applications and Web Sites: Experience, Theory, and Predictions," *J. Manag. Inf. Syst.*, vol. 26, no. 2, pp. 15–41, Sep. 2009.

[43] P. Criscuolo, N. Nicolaou, and A. Salter, "The elixir (or burden) of youth? Exploring differences in innovation between start-ups and established firms," *Res. Policy*, vol. 41, no. 2, pp. 319–333, 2012.

[44] AngelList LLC, "AngelList," 2016. [Online]. Available: www.angellist.com. [Accessed: 02-Jun-2016].

[45] B. Flyvbjerg, "Five Misunderstandings About Case-Study Research," *Qual. Inq.*, vol. 12, no. 2, pp. 219–245, 2006.

[46] R. B. Johnson, "Examining the validity structure of qualitative research," *Education*, vol. 118, no. 2, p. 282, 1997.

[47] J. Saldaña, *The coding manual for qualitative researchers*. London: SAGE, 2009.

[48] J. M. Corbin and A. Strauss, "Grounded theory research: Procedures, canons, and evaluative criteria," *Qual. Sociol.*, vol. 13, no. 1, pp. 3–21, 1990.

[49] M. Patton, *Qualitative Research and Evaluation Methods*, 3rd ed. Thousand Oaks, CA: SAGE, 2002.

[50] K. Charmaz, *Constructing grounded theory: a practical guide through qualitative analysis*, vol. 10. London: SAGE, 2006.

[51] B. F. Schenck, "Freemium: Is the Price Right for Your Company?," *Entrepreneur*, 2011. [Online]. Available: http://www.entrepreneur.com/article/218107. [Accessed: 07-Jun-2016].

[52] S. Berman, *Not for Free: Revenue Strategies for a New World*. Boston, MA: Harvard Business Review Press, 2011.