Valuating and Pricing of Vehicle Generated Data as a Marketable Product in the Automotive Industry

Frank Bodendorf, Tobias Meissner and Jörg Franke
Friedrich-Alexander-University of Erlangen-Nuremberg (FAU)
Institute for Factory Automation and Production Systems
Erlangen, Germany
e-mail: frank.bodendorf@faps.fau.de
e-mail: tobias.d.meissner@fau.de
e-mail: joerg.franke@faps.fau.de

Abstract— **Data-driven business models play a significant role in the digital transformation of traditional value-added industries. More and more existing and potential partners of automobile manufacturers show interest in the vehicle data generated by their products. However, there is still no monetary value assessment to support decisions regarding the release of data. Traditional pricing approaches for material goods are based on cost, margin, and volume. However, these bottom-up calculation approaches are not applicable to digital goods. The background to this is, among other things, the uncertainty about the potential sales volume, the difficulty of cost splitting, and the high cost degression per unit of digital goods. This paper provides decision support for selling data to third parties as an intangible product. It introduces a concept that allows to value data generated by a motor vehicle in order to determine potential prospects and prices for sale. The evaluation model developed supports the car manufacturer's negotiating position towards potential data buyers.**

Keywords - Automotive Industry; Car Data; Business Model; Value Estimation

I. INTRODUCTION

The use of valuable data will fundamentally change competition in the future [21]. “The expected growth of the value pool from car data and shared mobility could add up to more than USD 1.5 trillion by 2030“ [1]. Volume and quality of this “data treasure“ will create strategic as well as operational competitive advantages [13].

Today, data is generated in large quantities by the vehicle, recording values of thousands of attributes. On the one hand, the vehicle user (driver) has the opportunity to enter data in on-board systems and “exchange” them for services. He/she receives individually adapted functions, such as voice control, comfort settings when entering the car, navigational instructions in real time or other services [3][4]. On the other hand, a variety of sensors and computers in the vehicle, unnoticed by the driver, generates a steady stream of data, which among others serves for control purposes [8]. Examples are the anti-lock braking system or the automatic windscreen wiper and light regulation. According to an internal study the data usage can be divided into nine purpose oriented categories:

- Facilitating vehicle use
- Meeting regulatory and legal requirements
- Supporting marketing and advertisement
- Assessing IT security
- Improving technical processes
- Fulfilling terms of contract
- Innovating and developing products
- Ensuring road safety
- Offering services to third parties

All of these categories have in common that value is created through the use of collected vehicle data. On the one hand, this value is reflected in technical or qualitative improvements as well as in cost reductions of the company's internal processes. On the other hand, the use of vehicle data can also lead to an economic improvement of the business results and in particular to an increase in turnover [22]. This may be a result of higher sales figures of products, i.e., manufactured vehicles, which are more attractive through data-based functions (“data infused products”). In addition, it is possible to offer certain data for sale as an end product itself [12].

This paper focuses on selling data to third parties. The demand for vehicle generated data depends on the benefit seen or expected by the buyer. From the perspective of the data provider, it is important to determine the value of the data in order to estimate the demand potential on an external market and to create appropriate pricing models.

The paper consists of three main parts. First, it gives an insight into the state of the art of data evaluation and points out weaknesses. Second, a new methodology is introduced that combines business model analysis and value estimation approaches. Third, a use case is presented which illustrates the model application and leads through the main process steps of value estimation.

II. EXISTING VALUATION APPROACHES

The generic value of a data product in sales situations cannot be determined by a benefit that has already been

 ISBN: 978-1-61208-720-7

realized, since the data is not yet being used by the buyer at the time of the transfer. Therefore, an evaluation must be based on probable and potential benefits [12]. This value can be estimated by using qualitative and quantitative methodological approaches. For a corresponding systematic value determination of vehicle data, a number of existing evaluation methods is outlined. This is initially done by a tabular overview in Section A, followed by a more detailed description in Section B and C and by a discussion of limitations and transferability in Section III.A. Literature often speaks of data and information without exactly differentiate between these terms. Some authors see in information "refined data", e. g., by placing it in a context of meaning. In this paper, both terms are used synonymously.

A. *Overview of potential methods for data evaluation*

The identification and selection of potential valuation procedures is done through a combination of literature review and in-depth interviews. First, 20 sources of literature are used to collect a comprehensive set of possible valuation approaches. Subsequently, valuation approaches are selected and specified with the help of 50 in-depth interviews with experts from the divisions or departments in the areas of cost engineering, data strategy, data analysis, and purchasing. The consolidated results are shown in Table I [18].

TABLE I. QUALITATIVE AND QUANTITATIVE VALUATION APPROACHES FOR DATA

Method	**Characteristics**			
	Type	*Input*	*Operator*	*Output*
Data Product Scorecard	Qualitative	Data attributes	Scoring-method	Willingness to pay
Data Value Design Canvas	Qualitative	Data use case	Expert workshop, Canvas nine factors	Interactions / connections
Value determi-nation per user	Quantitative	Acquisition cost Number of users	Discount calculation	Price per user-dataset
Value improve-ment by data services	Quantitative	Data material	Statistical analysis, e. g, hypothesis testing	Increase in value or quality through the use of data
Value determi-nation by Laney	Qualitative and Quantitative	Data material	Gartner Valuation Model	Qualitiative and financial value
Value determi-nation by partners	Quantitative	Theoretical value, maturity, expiration of information	Intangible Assets Evaluation	Monetary information value
Pricing based on customer value	Quantitative	Different data bundels	Versioning, price differen-tiation, surcharge calculation	Price for data bundels

B. *Qualitative evaluation*

The appraisal of methods has been carried out by literature review and interviews as mentioned in section II.A. The following most popular methods for evaluating vehicle data are identified:

- Data Product Scorecard
- Data Value Design Canvas

The *Data Product Scorecard* is a method of pricing on data marketplaces. For this purpose, the customer's willingness to pay depending on various data properties must first be estimated. This qualitative evaluation of the data properties is made by the Data Product Scorecard from a simulated perspective of end users or potential buyers of the data [20]. As part of an evaluation workshop within the company, the role of the user is taken and each data characteristic given in the scorecard is rated with 0, 5 or 10 points.

The *Data Value Design Canvas* approach looks at the data value chain. The approach is based on the theory of Service Dominant Logic and the "Jobs-To-Be-Done" theory [2][14]. According to [16], the data value chain begins with the generation of data and extends up to the provision of information to the (paying) customer.

C. *Quantitative evaluation*

Many companies have problems finding the real economic value of their data [15]. For a rethink in the development of new business models [7] and the optimization of internal processes, the determination of this value, especially for the automotive industry, is of particular importance.

Value determination per user

When acquiring companies with data-driven business models who have not yet monetized their database but still offer data-based applications to the end user, the data value is often determined by the value of the application per user. The price of acquisition is divided by the total number of end users of the application. From this calculated price per user, the average user acquisition costs are subtracted [10][19].

Value determination by Laney

According to Laney the data is evaluated through quality-based and quantitative financial analysis [12]. In the quality-oriented evaluation, the output is a scoring value between zero and one, in the financial evaluation an absolute monetary value. The two-part consideration focuses on methods for improving the "Information Management Discipline" and deals with "Foundational Measures" as

- How correct, complete and exclusive is the data? (Intrinsic Value),
- How good and relevant is the data for specific purposes? (Business Value),
- How does this data affect key business drivers? (Performance Value).

On the other hand, the "Information Economic Benefit" of "Financial Measures" is examined:

- What would it cost us if we lose this data? (Cost Value),

- What could we get from selling or trading this data? (Market Value),
- How does this data contribute to our bottom line? (Economic Value).

Both considerations provide a quantitatively measurable contribution to the value of data and will be explained in more detail below. Based on a collection and analysis of existing valuation approaches according to Laney, the combination of a plausibility check and a requirements analysis leads to a new valuation perspective introduced into a new model presented in chapter III.

III. DEVELOPED METHODOLOGY

A. *Motivation of a new evaluation approach*

Value Determination by Laney comprises pure comparative methods. These do not calculate monetary values, yet they show some interesting perspectives. The Data Value Design Canvas lists nine factors which affect the value of data. The effects of information or data are considered generally. For example, information/data protects against unwanted events or promotes wanted events. Unwanted events always result in costs. Thus, the avoidance of unwanted events corresponds to a cost reduction. The realization of desired events effects an increase in sales as the most relevant example. The Data Product Scorecard assesses the willingness to pay of a customer. If the information considered is “perfect”, the customer is willing to pay the full price for this information.

The method according to Laney, the Data Value Design Canvas, and the Data Product Scorecard focus on the quality of the information or data. Both methods have in common that always certain use cases of data usage are considered.

By comparing the different approaches listed in Section II, some requirements can be derived for a new concept integrating different aspects:

- Quality factors must be taken into account.
- The willingness to pay is relevant for data sales. This depends on various factors, including the purchase motive, the perceived benefit, the reputation of the seller, and the individual purchase situation.
- Data has the potential to increase sales at the customer or to reduce costs for internal customer processes.
- Competition should be considered as an important factor.

B. *Evaluation model*

The developed “integrated methodology” for an innovative evaluation model meets these requirements from the fields of quality assessment [17], price differentiation [24], cost management, and competitive analysis in a combined way [22]. So, for a specific use case, it is possible to estimate a monetary value of data by integrating different aspects (see section III.A). Basically, Laney's approaches are not limited to any specific field of application [23].

The plausibility check of selling prices is achieved by combining qualitative tools based on methods such as Business Model Canvas or Data Canvas with a practical evaluation through quality workshops as well as quantitative calculations. These include, among other things, the valuation by Laney, a bottom-up cost calculation as well as profit split approaches.

The process model outlined in Figure 1 shows the process steps of the model for a use-case-specific value determination. The non-rivalry property of data enables multiple sales of similar data bundles or even the same dataset. The total value of the data bundle can be determined as the sum of the values across all (potential) use cases:

$$V_G = \sum_{i=1}^{n} V_i \qquad \text{(a)}$$

V_G	Total value of data bundle G
V_i	Individual value of data for the use case i
i	Use case index
n	Number of use cases

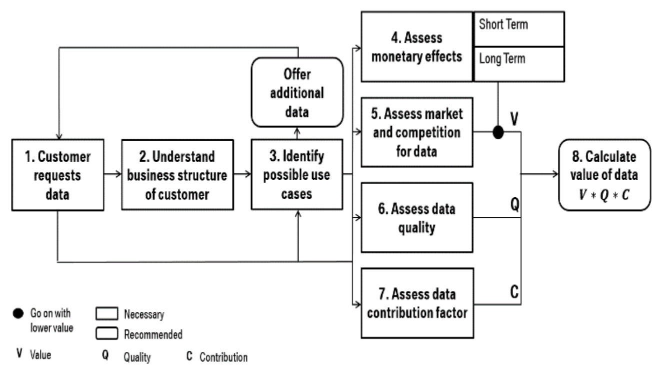

Figure 1. Process for the monetary valuation of data

As a first step in the process, data requests from potential business customers are collected or data is proactively offered to the customer. In order to be able to identify potential data needs, the customer's business model must first be understood. (1 + 2)

The way the customer translates the data into benefits can be identified through a systematic analysis of possible use cases. For this a combination of the Business Model Canvas and Data Value Design Canvas is suitable. With progressive understanding of the application, it is theoretically possible to offer different or additional versions of data bundles to the customer. (2 + 3)

In order to determine the customer's willingness to pay, the value of the use case must be understood in detail [9]. For this, a possible cost reduction or increase in sales by the data is to be determined.

For each use case, there is both a short-term and a long-term monetary benefit, which in individual cases can also be zero. The model of Figure 1 shows a parallel approach to Laney's business value calculation of information, which determines the data relevance for specific processes. (4)

In cases where there is competition on "data marketplaces" or the self-collection of data is significantly more favorable than granting the monetary benefit to a third party, these influences must be measured for further calculation. (5)

At the same time, the data is qualitatively evaluated based on selected criteria (see section II.B). The model considers the monetary value of quality criteria, following Laney's valuation ideas in the form of a quality factor Q. (6)

In addition, the data contribution factor C takes into account that other vehicle generated data or additional information may be necessary in addition to the offered vehicle generated data. (7)

After having carried out steps 1 to 7, finally the use-case-specific combined value of the data is calculated (8). The determined preliminary data value V is multiplied by the quality factor Q and the data contribution factor C. Both of them are discount factors. The calculations of Q, C, and V are illustrated in section IV. Hence, V is reduced according to low quality or insufficient amount of data. Interdependencies between the factors are possible but not taken into account. E. g., low data quality can lead to a higher demand for data from other sources. A lowered data quality factor can therefore further reduce the data contribution factor.

IV. USE CASE

The applicability of the developed methodology is experimentally tested for real sales situations. One of them is the offering of Road Segment Data (RSD) to a navigation maps provider. This data is needed to provide a high definition roadmap for autonomous driving.

A. Business structure analysis

The results of the structural analysis of the business case are outlined below. On this basis, further analysis leads to the data needs and the data bundles to be offered [11][20].

- *Value proposition:* build tomorrow's road network
- *Customers:* companies
- *Segments:* automotive industries
- *Channels:* direct contact
- *Revenue stream:* Selling refined information (HD map) for autonomous driving and location based services
- *Cost structure:* personnel, data transfer
- *Key resources:* street/road data, navigational data, real time traffic information data, road segment data (RSD)
- ***Data offering**:* road segment data captured by car cameras
  (edge markings, center markings, strip width, crash barriers, guide posts, signs, wild animal warning reflectors and barriers)

B. Monetary Effect

An internal study of the car the manufacturer says that the willingness to pay for the enriched HD map in the self-driving vehicle industry is 60 € per year for a highly autonomous or fully autonomous vehicle. For the year 2020, a global volume of 3.3 million high or fully autonomous vehicles is expected on the market. In 2035 this number is predicted to be 28 million vehicles. The market for navigation maps in vehicles is divided into market shares of 15% to 25% [6].

These assumptions lead in the worst case scenario (assuming a market share of 15%) to a potential turnover of at least 29.7 million € in 2020 (3.3 million * 60 € * 0.15) and 252 million € in 2035 (28 million * 60 € * 0.15).

C. Market competition

The market position is qualitatively described by the criteria "is it valuable", "is it rare", "is it hard to imitate", and "is the firm organized for success" according to [15]. The answers to the questions in a conducted survey are based on a competitive analysis of 26 navigation maps suppliers. The qualitative assessment shows that there is a "short term competitive advantage".

D. Data quality and contribution factors

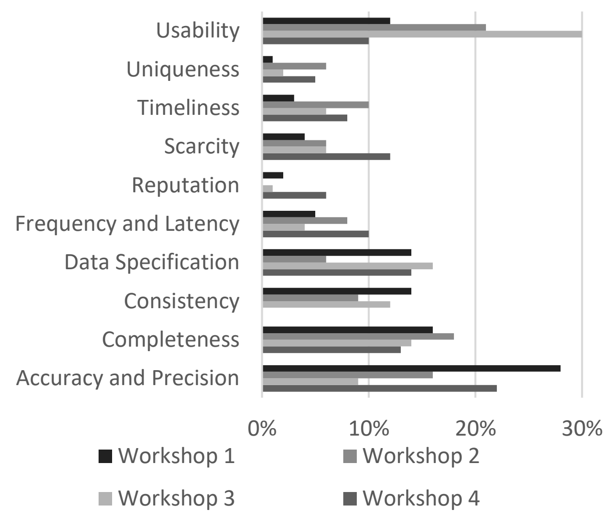

Figure 2 Evaluation of the quality criteria of data

In order to evaluate the data quality, 14 data scientists are interviewed in four expert workshops on given quality criteria. The results of the workshops are summarized in Figure 2 and transformed into a quality factor Q.

$$Q = \sum_{i=1}^{n} c_i * w_i \qquad \text{(b)}$$

Q	Quality factor
c_i	Evaluation factor of criterion i
w_i	Weight of criterion i
i	Index of criterion
n	Number of criteria

The evaluation factor values c_i of the criteria result from a pairwise comparison of all criteria in a preference matrix. The data contribution factor *C* expresses if all required data (contribution factor = 1), almost all data (contribution factor = 0.75), about half of the data (contribution factor = 0.5), few data (contribution factor = 0.25) or no data (contribution factor = 0) can be provided.

$$C = CM * w_{cm} + CK * w_{ck} \tag{c}$$

C	Data contributing factor
CM	Contribution factor of metadata
CK	Contribution factor of key data
w_{cm}	Weight of metadata
w_{ck}	Weight of key data

The use-case-specific factors CM und CK differentiate between key data und metadata (additional data). The expert based weights w_{cm} and w_{ck} rate the relative data contribution of each data type in the use case. Location based services are one example. They are based on customer preferences (metadata) on the one hand and GPS data (key data) generated by the vehicle on the other hand. Experts give weights of 0.2 for GPS data and 0.8 for preference data. It is assumed that the vehicle can deliver 90 percent of the GPS data needed, however none of the preference data. So, the contribution factor of the metadata is zero. The calculation of C results in the value of 0.18:

$$0.2 * 0.9 + 0.8 * 0 = 0.18$$

The corresponding calculation for the navigation maps provider gives the following result:

Quality factor Q = 0.8

Contribution factor C = 0.55

E. Selling prices

Through the product of monetary effect (*V*), quality (*Q*) and contribution factor (*C*) (see Figure 1) the value for the offered data for the use case "navigation map provider" is in the worst case scenario 13.07 million € (29.7 million € * 0.8 * 0.55).

F. Lessons Learned

On the one hand the presented evaluation model has been applied to several fictitious use cases with real information coming from companies interested in buying data but without any sales decision. On the other hand the evaluation model was tested on several specific sales situations and the outcome of the model, i. e. the post calculated data value, was compared to the real sales price. The monetary data values determined using the evaluation model show an average deviation of 8% from the sales prices negotiated in practice.

Thus, the created model seems to deliver a good approximation to price imaginations which are based on gut instincts so far. The model is capable to support price negotiations by providing a systematic methodological basis and a transparent multidimensional valuation framework.

V. CONCLUSIONS

The methods of data evaluation identified in the literature are individually not suitable for practical value determination of data and their pricing in sales situations. This paper presents a methodology that focuses on the selling of data as intangible products to external business partners.

The methodology can also be transferred to use cases within the company. In addition to determining the value of the data, decisions regarding the pricing model must be made.

However, for long-term strategies it is unclear to what extent currently recorded data is valuable in the future. Data that is still useless, because currently there are no use cases, can be highly relevant for future use cases. Because of the existing knowledge gap and missing empirical values, it is impossible to determine a value of data over the entire lifecycle, above all because of very uncertain future potentials.

This article exemplifies a possible evaluation and monetization of a small fraction of the total data available in the automotive industry.

Against the background of the huge amounts of data available there, it quickly becomes clear that due to technical limitations probably never all potential use cases can be implemented. There are various transmission options for vehicle generated data. The built-in memory can be read in authorized garages, updates can be transmitted at weekly or daily intervals, or data can be transferred in real time. There is always a technical limitation due to the restricted transfer rates or transfer options. Not all conceivable applications can be realized at the same time.

It is also an open question whether it makes sense to regard the vehicle as an open platform. In this case, an automobile manufacturer or even the automotive industry as a platform provider could probably sell the platform as a service (PaaS) to service providers who will pay for specific data accessed via the platform. As an analogy, platforms of Apple and Android can be considered. Third parties develop services to be offered on these platforms. The developed services (e. g., apps) increase the attractiveness of the platform. Depending on the design, there are direct and indirect network effects. With regard to autonomous driving, this approach may potentially increase the attractiveness of vehicles and vehicle fleets acting as such platforms. For example, in addition to many existing connected drive services, applications of third-party providers can be activated, which leads to an immense increase of the value of a ride and the driving experience for the customer. Here, completely new service ecosystems spanning and connecting different industrial sectors are appearing. To name only one step toward the future, the intelligent personal assistant from BOSCH enables the networking of car services and e-home services [5].

 ISBN: 978-1-61208-720-7

REFERENCES

[1] J. Balasubramanian, S. Beiker, S. Chauhan, T. Colombo, F. Hansson, S. Inampudi, R. Jaarsma and M. Kässer, “Car data: paving the way to value creating mobility : Perspectives on a new automotive business model”, Advanced Industries McKinsey, 2016

[2] L. A. Bettencourt and A. W. Ulwick, “The customer-centered innovation map", Harvard Business Review, vol. 86 (5), pp. 109-114, 2008

[3] BMW: *BMW ConnectedDrive Kundenportal – digitale Vernetzung zu Ihrem BMW*. [Online]. Available from: https://www.bmw-connecteddrive.de/app/index.html#/portal/store/Base_TolCarOffer [retrieved: 11, 2018]

[4] BMW: *NOW Mobilitätsdienstleistungen*. [Online]. Available from: https://www.bmwgroup.com/de/marken/now-mobilitaetsdienstleistungen.html[retrieved: 01, 2019]

[5] BOSCH: *Nur Fahren war gestern – der persönliche Assistent ist morgen*. [Online]. Available from: https://www.bosch-presse.de/pressportal/de/de/nur-fahren-war-gestern-%E2%80%93-der-persoenliche-assistent-ist-morgen-101568.html. [retrieved: 01, 2019]

[6] A. Brugger: *Global production forecast for semi- and fully automated vehicles*. [Online]. Available from: https://www.oliverwyman.com/con-tent/dam/oliver-wyman/global/en/files/who-we-are/press-releases/Oli-verWyman_Graphics_Value%20Pools%20Autonomous%20 Driving_EN_16072015_final.pdf [retrieved: 11, 2018]

[7] J. L. L. Francisco and J. Esteves, Value in a Digital World: How to assess business models and measure value in a digital world, Berlin: Springer, 2017. IS BN 978-3-319-51750-6

[8] W. Gründinger: *Data-Driven Business Models in Connected Cars, Mobility Services & Beyond.* [Online]. Available from: https://www.bvdw.org/fileadmin/bvdw/upload/publikationen/ connected_mobility/20180418_data_driven_business_models _Seiberth_Gruendinger.pdf [retrieved: 01, 2019]

[9] R. Harmon, H. Demirkan, B. Hefley , and N. Auseklis, “Pricing Strategies for Information Technology Services: A Value-Based Approach”, 42nd annual Hawai'i International Conference on System Sciences (HICSS 2009), Piscataway, NJ: IEEE Press, Jan. 2009, pp. 1–10, ISBN: 978-1-424-44197-6

[10] N. Henke, et al., The Age Of Analytics: Competing In A Data-Driven World, McKinsey Global Institute, 2016

[11] HERE: *Location-Based Services for the Autonomous Future.* [Online]. Available from: https://www.here.com/en. [retrieved: 01, 2019]

[12] D. B. Laney, Infonomics, How to monetize, manage, and measure information as an asset for competitive advantage, New York, NY: bibliomotion inc, 2018, ISBN: 978-1138-09038-5

[13] E. Leonidas, DataStreamX: *A Practical Guide to Pricing Data Products*. [Online]. Available from: http://cdn2.hubspot.net/hubfs/573334/Downloadable_Con-tent_(WP_or_Guides)/DataStreamX_Data_Product_Pricing_ Whitepaper.pdf [retrieved: 11, 2018]

[14] C. Lim, et al., “From data to value: A nine-factor framework for data-based value creation in infor-mation-intensive services”, International Journal of Information Management, vol. 39, pp. 121–135, 2018, ISSN: 0268-4012

[15] S.M. Liozu, W. Ulaga, Monetizing Data, A Practical Roadmap for Framing, Pricing & Selling Your B2B Digital Offers, Guetersloh, Germany: VIA Publishing, 2018, ISBN: 978-1-945815-04-1

[16] K. Mathis and F. Köhler, “Data-Need Fit - Towards data-driven business model innovation”, The Fifth Service Design and Innovation conference (ServDes.), pp. 458–467, 2016, ISBN: 978-91-7685-738-0

[17] D. Mcglivary, “Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information: Definitions of Data Categories”, Chief Data Officer & Information Quality Symposium, 2009

[18] T. Meißner, Wert und Verwertung von Daten der Hardware- und Softwarenutzung im Fahrzeug, University of Erlangen-Nuremberg, Master Thesis, Erlangen, 2018

[19] K. O'Neal, Quantifying the Value of Data: The First Step in Data, unpublished presentation

[20] A. Osterwalder and Y. Pigneur: Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers, Hoboken, New Jersey: John Wiley & Sons, 2010, ISBN: 978-0470-87641-1

[21] B. Schmarzo and M. Sidaoui, Applying Economic Concepts To Big Data To Determine The Financial Value Of The Organization’s Data And Analytics, And Understanding The Ramifications On The Organizations’ Financial State-ments And IT Operations And Business Strategies. [Online]. Available from: https://in-focus.dellemc.com/wp-content/uploads/2017/04/USF_The_Economics_of_Data_and _Analytics-Final3.pdf [retrieved: 01, 2019]

[22] C. Shapiro and H. R. Varian, Versioning, “The Smart Way to Sell Information”, Harvard Business Review, vol. 6, pp. 107 f., 1998

[23] F. Stahl and G. Vossen, “High quality information provisioning and data pricing”, 29th International Conference on Data Engineering Workshops (ICDEW 2013), 2013, pp. 290-293

[24] S. Viswanathan and G. Anandalingam, “Pricing strategies for information goods”, Sadhana Indian Academic Sciences Journal, vol. 30, 2005, pp. 257–274, ISSN: 0256-2499

Wirtschaftswoche: *GPS-Satellitennavigation: Wie Navigationsgeräte-Firmen um einen Milliarden-Markt kämpfen.* [Online]. Available from: https://www.wiwo.de/unternehmen/gps-satellitennavigation-wie-navigationsgeraete-firmen-um-einen-milliarden-markt-ka-empfen-seite-3/5426542-3.html [retrieved: 11, 2018]