# On the Value of Data: Multi-Objective Maximization of Value Creation in Data-Driven Industrial Services

Jürg Meierhofer
*School of Engineering*
*Zurich University of Applied Sciences*
CH 8401 Winterthur, Switzerland
juerg.meierhofer@zhaw.ch

Rodolfo Benedech
*School of Engineering*
*Zurich University of Applied Sciences*
CH 8401 Winterthur, Switzerland
rodolfoandres.benedech@zhaw.ch

Christoph Heitz
*School of Engineering*
*Zurich University of Applied Sciences*
CH 8401 Winterthur, Switzerland
christoph.heitz@zhaw.ch

***Abstract*—Data-driven value creation is a key topic in industrial services. However, designing such services in an optimal way represents a multidimensional and complex task. In this paper, we present a design methodology based on a simultaneous maximization of value creation for both the provider and the customer, allowing the identification of optimal service configurations. We apply this methodology to a use case of a manufacturer delivering services for its machines in the context of a pay-per-use business model.**

**The approach is based on modeling the value creation separately for both provider and customer, as a function of data-driven services which may be offered in different phases of the lifecycle. The model allows finding Pareto-optimal service configurations which provide value creation optimized simultaneously for both the provider and the customer. These optimal configurations are not easy to find with simpler methods because of non-linear effects in value creation along the lifecycle.**

***Index Terms*—smart services, data-intensity, customer lifecycle, mutual value creation**

## I. Introduction

In this paper, we develop a quantitative model for value creation by utilizing data along the different phases of the customer lifecycle. The mathematical model quantitatively describes how the data-driven services create value for the customer and the provider by specific service processes. For the combined optimization of the value for the provider and the customer, the model draws on the approach for multi-objective optimization that was proposed in [1] for service customization in the customer journey. The model is extended and applied to optimizing service value creation with smart, connected products along the customer lifecycle in the industrial manufacturing and elaborates on detailed service processes in order to create this value.

Applying the utility-oriented approach for data valuation in service ecosystems discussed in [2], we focus on the dyadic relation between a service provider and a beneficiary, as shown in Fig. 2. The provider utilizes data about technical (e.g., machines, equipment) and human (e.g, operator, technician) actors in order to create value additionally to the value generated by the equipment and basic after-sales services. This value may be provided to the customer in the form of advanced services [3], [4] that enable getting an output from the equipment that is better targeted at the customers jobs, pains, and gains [5]. Additionally, the provider can utilize the data for improving its own operations, in particular its product innovation and marketing process. Both these actions increase the effectivity of the provider and avoid waste.

In [1], the value creation for the customer and the provider are jointly considered by applying multi-objective optimization. The optimum service configurations can be identified via the Pareto front of a representation of both the value for the customer and for the provider for the different configurations (example in Fig. 4). Taking this concept further, this paper elaborates how specifically the utilization of data for value creation can be operationalized in the different phases of the customer lifecycle in order to optimize value creation. What does it mean for a provider-customer constellation to have different positions on the Pareto front and below? How does the design of the data-driven services impact this position? For example, which data-driven after-sales maintenance services should a provider invest in? How to trade-off the investment in after-sales services against additional pre-sales consulting services? Thus, the research question of this paper is as follows: *How to design data-driven services in the different phases of the customer lifecycle for optimizing value creation for provider and customer on an operational level?*

## II. Research Methodology

The quantitative model elaborated in this paper is based on a previous study [1], which conceptualizes value creation over time by modeling value creation for provider and customer as two separate dimensions. In this paper, we model the value creation as a function of time-varying state functions which, in turn, are influenced by service processes. The service value is modeled based on the principles of industrial service value creation. The detailed value creation is mathematically modeled taking into account the value drivers and the value detractors. This mathematical model for the quantification of the value creation is numerically evaluated for a specific application case. The application case is hypothetical, but derived in its essence from the empirical study described in [2]. It consists of a service constellation in which a manufacturer provides equipment to a customer with a pay-per-use business model.

## III. Model for Value Creation by Data-driven Services

### A. Conceptual Model for Data-Driven Value Creation along the Lifecycle

We consider a manufacturer which provides equipment (e.g., an industrial machine for material processing) to a business customer who uses it in its operations, e.g., for producing end customer products. We assume a use-oriented service business model according to [6], [7] in which the customer gets the right to use the equipment for a fee per time period. Additionally, the customer may order additional services from the manufacturer for improving the equipment performance for additional fees. This service model reduces the CapEx (capital expenditure) requirements for the customer because it can base its operations only on operational fees (OpEx, operational expenditure). This makes this model especially attractive for small and medium sized enterprises (SMEs) [7].

In order to model the economic impact, we consider the provider-customer relationship along the full customer lifecycle. Based on [8] and extending [9], we model the lifecycle by the four phases shown in Fig. 1.

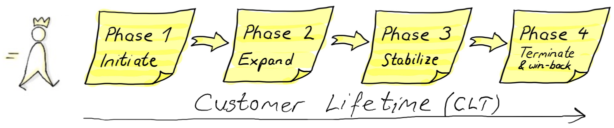


Fig. 1. Phases of the customer lifecycle (based on [8])

In phase 1 (initiate), the customer relationship is initiated by marketing the equipment to the customer and, in case of a successful conversion, getting the use-oriented contract signed and the equipment installed. In phase 2 (expand), the customer gets more and more acquainted with the equipment and learns how to use it optimally in its given context (learning how to create value-in-context [10]). During this phase, the equipment performance is increasing over time due to learning effects. Phase 3 (stabilize) represents the usage phase which typically lasts several years. In phase 4 (terminate and win-back), the customer decides to quit the use-oriented contract, while the provider tries to win it back. Successful win-back creates additional customer lifetime $CLT$, with corresponding profit.

### B. Enabling Use-oriented Service Models With Smart, Connected Products

According to [7], use-oriented service models profit from accurate and timely information about the equipment and its usage. For generating this information, the equipment is enhanced by sensors and connected to the provider by the internet-of-things (IoT). Thus, providers and customers create data from operating the equipment and the processes. The customer can share data about the condition of the equipment and the processes with the service provider (Fig. 2). This leads to the concept of smart, connected products [11]. Also business processes create data, typically in the form of time stamps from work flow tools, that can be shared with the service provider and combined with the data from the equipment, for enabling value-creating interventions.

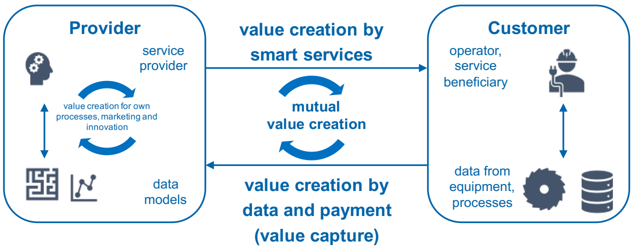


Fig. 2. Conceptual model for valuation of data. Source of figure: [2]

The provider can then create digital models representing the customer's products, processes, and operations (e.g., digital twins [12]). These models describe past, current or predicted future states of of the equipment and processes and thus may provide decision advisory for management actions. The impact of such data-based services manifests itself in improved availability, efficiency, or quality, thus creating functional or utility-oriented value [13], [14], which can easily be translated into financial value by applying cost factors, e.g., hourly rates of staff or cost of waiting time.

### C. Quantifying Value Generation on the Level of the Lifecycle

In this subsection, we model the value generation on the level of the lifecycle phases (see Fig. 1), i.e., without going into the details within the four phases. The detailed value creation within the four phases will be discussed in subsection III-F.

The value of a service firm has shown to be dominated by the value of its customer base, as revenues are generated from customer relations [15]. This definition of firm value represents a long-term strategic view on value creation. For quantifying the value of the customer base, we use the concept of customer lifetime value $CLV$ [15], defined as the total net profit (revenues minus costs) generated from a customer, over its lifetime. Thus, the value for the provider, $V_P$, becomes

$$V_P = \sum_{i=1}^{4} V_{P,i} \tag{1}$$

where each $V_{P,i}$ represents the value accrued in phase $i$ of the lifecycle. Therefore, modeling value creation for the provider as a function of data usage requires to derive a relation between specific ways of using data and their effect on $CLV$. Note that the extension of the customer lifetime is one way of increasing $CLV$ (see subsection III-F).

For quantifying the value from the perspective of the customer, and against the background of the use-oriented service model discussed in III-A, another logic seems more appropriate for the customer value $V_C$. In the studied setting of a pay-per-use utilization of equipment, the additional customer value from sharing data and benefiting from data-based services is evaluated by its effect on the OpEx, i.e., by considering yearly costs and benefits. Therefore, we define the value for the customer by the average yearly value created by the services - not including the basic fee for using the equipment - over the lifetime $CLT$.

$$V_C = \frac{1}{CLT} \sum_{i=1}^{4} V_{C,i} \tag{2}$$

Since the focus of this paper is put on value creation by additional, data-driven services, we do not take into account the value created by the base product and the standard service when calculating $V_P$ and $V_C$. The only point where the base annual fee paid by the customer will be catered for is in phase 4 (terminate), where an extension of the customer lifetime ($CLT$) will incur more base fee for the provider value $V_P$. This effect will be calculated in subsection III-F4.

### D. Different Services with Different Data-Intensities

The creation of value in each phase of the lifecycle for the provider and the customer ($V_{P,i}$ and $V_{C,i}$) depends on the design of the services in these phases. These services will be described in TABLE I in the description of the application case (section IV). In each phase $i$ ($i = 1...4$), there are different services with different characteristics. For each of these services, we define a normalized measure for the data-intensity per phase of the lifecycle, $D_i$, which represents an independent variable that impacts $V_{P,i}$ and $V_{C,i}$, i.e.,

$$V_{P,i} = f_p(D_i) \quad \text{and} \quad V_{C,i} = f_c(D_i) \tag{3}$$

The functions $f_p(\cdot)$ and $f_c(\cdot)$ will be elaborated in subsection III-F.

Without limiting the generality, we can consecutively number the services with their data-intensity $D_i$ increasing in a given phase. This is inline with the increasing level of services offerings from basic to intermediate to advanced services described by [4]. For example, in phase 3 (stabilize) of the application case (TABLE I), this means: level 1 is just standard base or intermediate service with data-intensity $D_3 = 1$. Also a standard service like, e.g., yearly maintenance, requires a minimum data-intensity $D_3$ like, e.g., a service logbook. Level 2 enables a monitoring service requiring some sensors on the equipment transmitting their data, which we denote as data-intensity $D_3 = 2$. Level 3 adds remote service, e.g., by parameter configuration over the IoT, which requires a higher data intensity $D_3 = 3$. Level 4 adds condition based maintenance on top, which requires analytical models which process the sensed data to new data (data-intensity $D_3 = 4$). Level 5, finally, adds performance optimization, which evaluates optimum parameter settings based on the data for achieving certain goals, e.g., minimizing wear and tear on components for a given output (data-intensity $D_3 = 5$). Levels 4 and 5 are enabler for advanced services according to [4].

Section III-E will discuss how $V_{P,i}$ and $V_{C,i}$ of Eq. 3 are modeled and computed for the lifecycle phases $i = 1...4$. In order to simplify the notation of the formulas, we will omit the costs and fees incurred for the services. However, in a real application case, each service will require the customer to pay a fee - thus reducing $V_C$ and increasing $V_P$, and it will also incur costs for the provider to deliver the service (e.g., to operate the resources for remote service) - thus reducing $V_P$. In the application case (section IV), these fees and costs will be taken into account.

### E. Optimization of Value for Provider and Customer

The research questions with the sub-problem statement how to optimize value creation for the provider and the customer leads to the question how to jointly optimize $V_P$ and $V_C$. The research question can be re-framed to: *"Which combination of services along the lifecycle with their required data-intensities $D_i$ ($i = 0...4$) for service value creation optimize the values $V_P$ and $V_C$ in a combined way?"*.

A straightforward approach would be to take $V_P + V_C$ and optimize the data-intensity per phase of the lifecycle to maximize this sum. However, the customer and the provider have totally different value contexts, which makes a direct summation $V_P + V_C$ not appropriate. This is also supported by the way $V_P$ and $V_C$ are calculated (Eqs. 1 and 2): the two value schemes have different dimensions - a cumulative value over the lifetime for $V_P$ and a per annum value for $V_C$, and also very different scales. As already applied in [1], an approach is given by multi-objective optimization (e.g., [16]). Applying this, for all possible combinations of the independent variables (i.e., the varying data-intensity in the phases of the lifecycle), their impact on the two target variables $V_P$ and $V_C$ is considered. The two target variables will be represented in a scatter plot as shown in the example of Fig. 4. The optimum can then be found on the Pareto front, as will be discussed in section IV-C.

### F. Value Creation by Data-Intensity along the Lifecycle

Utilizing operational data with different intensities $D_i$ creates different benefits and costs in each phase of the lifecycle (Eq. 3). Costs arise for the provider to deliver the services and for the customers to pay for them via a fee. These costs reduce the value on both sides. In the sequel, we discuss how data creates the value at the different intensities.

*1) Phase Initiate:* In the initiation phase, customers get information and offers of new products. Utilizing data from the customer in this phase typically means knowing the customer wishes and needs (in terms of jobs, pains, and gains [5]) more precisely, thus being enabled for more targeted marketing and offerings. We assume that the customer need is described by $M$ different dimensions and is represented by the vector $\vec{C}$. For example, these dimensions may describe the customer need for capacity, size, prize or energy consumption. The provider has a product portfolio consisting of $N$ different products each having properties in the same dimensions as the customer needs. The properties of a specific product $k$ of this portfolio are denoted by a vector $\vec{P_k}$. As for the customer need, the dimensions describe the product properties (e.g., the capacity). The customer expectation gap [17] for the initial offering of product $k$ can then formulated as the distance $d_{0,k}$ ($0 \leq d_{0,k} \leq 1$) between $\vec{C_k}$ and $\vec{P_k}$. For the initiate phase, we model the probability for the customer accepting an offered product $k$ with $p_{take,k} = 1 - d_{0,k}$. If the provider has data $D_1$ about the customer's needs, it offers the product $k_{max}(D_1)$ of the portfolio which has the shortest distance $d_{0,k}$ with conversion rate $p_{take,k_{max}}$ instead of an random product with $p_{take,k_{rand}}$. With $C_c$ being the marketing costs per customer contact, the value of the data for the provider becomes

$$V_{P,1} = f_p(D_1) = C_c \cdot \left( \frac{1}{p_{take,k_{rand}}} - \frac{1}{p_{take,k_{max}(D_1)}} \right) \quad (4)$$

For the sake of simplicity, we model the customer value of targeted marketing to be negligible.

*2) Phase Expand:* In the expand phase, the customer gets acquainted with the equipment and learns how to handle it and how to create more value with the product. We describe this phase by "new customer phase" with higher support and training needs by the customer and assume it may last typically a couple of weeks to months. We define $f_{learn}(\cdot)$ as the learning curve, which monotonically grows over time and cannot exceed the level 1 over time. If during this phase, the provider collects data about the customer's challenges with the product and offers targeted training, the learning curve grows faster. Therefore, $f_{learn}(t, D_2)$ is dependent on time and the data-intensity.

The learning curve has direct impact on the performance of the equipment by $P_E(t) = P_{Emax} \cdot f_{learn}(t, D_2)$ with $P_{Emax}$ being the maximum possible performance after learning. The equipment performance is a compound variable dependent on the availability and the failure rate [18], which we do not further elaborate on here for the sake of simplicity. The equipment performance $P_E(t)$ directly impacts the customer value by its cumulated output over time, $\int P_E(t)dt$. Thus, the value of applying targeted training based on data with intensity $D_2$ becomes

$$V_{C,2} = P_{Emax} \cdot \int_{T_{xp}} \left[ f_{learn}(t, D_2) - f_{learn}(t, D_2 = 1) \right] dt \quad (5)$$

with $T_{xp}$ the duration of the phase expand. $f_{learn}(t, D_2 = 1)$ denotes the learning curve which the customer achieves without targeted training, e.g., by using standard manuals. For the sake of simplicity, we model the provider value of targeted training to be negligible.

*3) Phase Stabilize:* The phase stabilize usually represents the longest phase and may last several years for capital equipment (e.g., industrial machines). Additional value for the customer $V_{C,3}$ is created by applying data-driven services such as, e.g., condition monitoring, machine health prediction, performance optimization, or remote maintenance. The literature [19] describes how data-driven services can create value for the provider and the customer by different co-creation patterns. For the customer, this helps keep up or increase the equipment performance $P_E(t)$. We model the different levels of intensity of data-driven services by the different levers of value creation introduced in [11]: 1) monitoring 2) control 3) optimization, and 4) autonomy. An example for 1) is the condition monitoring of machines. The service provider can remotely observe the condition of the machine running at the customer's site (e.g., by "remote monitoring"). On level 2), feedback is established to control the machine based on the results of the monitoring. This can lead, for example, to operating parameters being adjusted to improve the condition or efficiency of the machine. The optimization used at level 3) pursues a minimization or maximization goal such as, e.g., number of units produced per time. Autonomous systems at level 4) can, e.g., be completely self-organized processes. For quantifying this effect, we model the equipment performance during the phase stabilize as

$$P_E(t) = \alpha_{service}(D_3) \cdot P_{Emax}$$

with $\alpha_{service}(D_3)$ denoting the performance level achievable by the service that can be achieved with data-intensity $D_3$ additionally. Analogously to the expand phase (Eq. 5), the value for the customer now becomes

$$V_{C,3} = P_{Emax} \cdot \int_{T_{stab}} \left[ \alpha_{service,l}(D_3) - \alpha_{service}(D_3 = 1) \right] dt \quad (6)$$

with $T_{stab}$ the duration of the phase stabilize. $\alpha_{service}(D_3 = 1)$ denotes the performance level which is achieved by standard services without any additional level of data-intensity.

The value for the provider in this phase is given by

$$V_{P,3} = CLT \cdot (C_{ServiceFee} - C_{ServiceCosts}) \quad (7)$$

with $C_{ServiceFee}$ the annual fee collected from the customer for the additional, data-driven services and $C_{ServiceCosts}$ the provider's annual cost to deliver these services. Since both $C_{ServiceFee}$ and $C_{ServiceCosts}$ depend on the data-intensity, $V_{P,3}$ is a function of $D_3$.

*4) Phase Terminate:* We assume that, in the phase stabilize, the customer yearly evaluates whether the current provider is still adequate and leaves the contract with a probability $p_{churn}$ and moves to another provider. In particular, the hurdle to do so is relatively low given the chosen use-oriented service model. Thus, the customer lifetime ($CLT$) is geometrically distributed with parameter $p_{churn}$ and an expected value $E[CLT] = 1/p_{churn}$ years.

If the event occurs that the customer really decides to leave the contract, the provider can try to win back the customer by a customized offering. For being able to do so, the provider needs sufficiently detailed, data-based customer insights. We model this effect by the win-back probability

$$p_{winback} = f_{winback}(D_4)$$

with $D_4$ denoting the data-intensity in phase 4. If the win-back measure is not successful (probability $1 - p_{winback}$), the customer still churns and $CLT$ remains unchanged. However, if it is successful (probability $p_{winback}$), the customer returns to phase 3 (stabilize). For the sake of simplicity, we assume that it keeps an unchanged churn probability $p_{churn}$ after this. Given this, the effective churn probability becomes $p_{churn} \cdot (1-p_{winback})$ and the effective customer lifetime $CLT$ taking into account the win-back (wb) measures becomes

$$E[CLT|wb] = \frac{1}{p_{churn} \cdot (1 - p_{winback})}$$

and the increase of the expected $CLV$ with win-back measures (wb) as opposed to without (nowb) becomes:

$$\Delta CLT = \frac{p_{winback}}{p_{churn} \cdot (1 - p_{winback})}$$

which is a function of the data-intensity $D_4$. Given the relation between $CLT$ and provider value (Eq. 7), increasing $CLT$ yields additional value for the provider $V_{P,4}$. While Eq. 7 incorporates only the service fees, we take also the yearly fee for the base equipment into account here.

$$V_{P,4} = \Delta CLT \cdot (C_{BaseFee} + C_{ServiceFee} - C_{ServiceCosts}) \tag{8}$$

with $C_{BaseFee}$ the yearly base fee for the equipment and the standard service paid by the customer. With $p_{winback}$ and thus also with $\Delta CLT$ depending on the data-intensity, it is clear that also $V_{P,4}$ is a function of $D_4$. In our simplified assumption, we model no additional value for the customer for this, i.e., $V_{C,4} = 0$, since a customer would also encounter comparable yearly costs after churning to another provider.

## IV. Application of the Model for Value Creation

### A. Description of Application Case

A hypothetical case study was developed based on in-depth interviews with eight manufacturing firms, in which typical patterns for data driven services were identified [2].

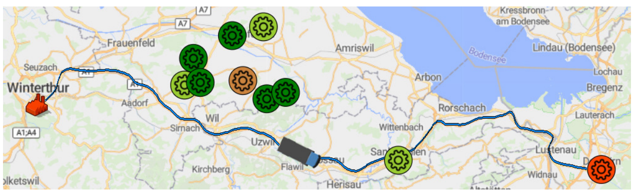


Fig. 3. Hypothetical example for constellation of providers (red factory) and customers (cogwheels in different colours indicating their condition). Service technician vehicle on the way to maintenance work (based on [2]).

We consider an SME which produces machines and offers data-driven (smart) industrial services over the lifecycle. The provider has an installed base of machines, which are operated at the premises of the customer (see the example in Fig. 3). The model incorporates utilizing data in different intensity levels for service value creation along the different phases of the lifecycle as described in subsection III-A. The overall value creation per customer is then computed according to Eqs. 1 and 2. The services implemented in the model are summarized in TABLE I.

TABLE I
DATA-DRIVEN SERVICES WITH DIFFERENT DATA-INTENSITY $D_i$ ALONG THE LIFECYCLE

| **Initiate** $D_1$ | **Expand** $D_2$ | **Stabilize*** $D_3$ | **Terminate** $D_4$ |
|---|---|---|---|
| 2. targeted offerings<br>1. no targeted offerings | 2. targeted training<br>1. no targeted training | 5. performance optimization<br>4. condition based maint.<br>3. remote service<br>2. monitoring<br>1. standard<br>*) 1. to 5. build cumulatively on each other | 2. targeted retention<br>1. no targeted retention |

### B. Numerical Model for the Application Case

The application model described in section IV-A is implemented in a numerical model using the MATLAB computing environment [20]. The value creation per lifecycle phase is implemented by applying the schemes described in subsection III-F. We assume that the customer gets the equipment without additional, data-driven services (just the base services) for an annual fee of $C_{BaseFee} = 7000$. In the customer's operations, the equipment has a maximum capacity $P_{Emax}$ to produce 1000 units of finished products per year and earns a margin of 100 per unit sold. The services listed in TABLE I have the following parameters:

- Phase 1, initiate: $C_c = 4000$, $p_{take,k_{rand}} = 0.05$, $p_{take,k_{max}}(D_1) = 0.2$, provider costs for applying targeted offerings, per customer: 500.
- Phase 2, expand: $T_{xp} = 0.5$ [years], capacity increase by targeted training: $\int_{T_{xp}} [f_{learn}(t, D_2) - f_{learn}(t, D_2 = 1)]\, dt = 0.3$ (Eq. 5), customer fee to get the training: 1000, provider cost to offer the training: 2900 (i.e., the provider subsidizes the training).
- Phase 3, stabilize: in TABLE II, capacity increase refers to $\int_{T_{stab}} [\alpha_{service}(D_3) - \alpha_{service}(D_3 = 1)]\, dt$ (Eq. 6). We assume for this hypothetical example that competition enforces to provide monitoring far below the cost, while remote service with its tangible properties can be charged substantially above costs, and, additionally, lowers the provider's process costs substantially. Condition based maintenance and performance optimization are considered as new development fields by the provider, which cannot yet be monetized at this early stage, i.e., fees are (far) below the costs.
- Phase 4, terminate: $p_{churn} = 0.1$ (hence: $CLT = 10$ [years]), $p_{winback} = 0.2$ (hence: $\Delta CLT = 2$ [years]), provider costs for applying win-back measures: 1000.

### C. Value Optimization for the Application Case

The numerical evaluation of $V_P$ and $V_C$ is done for the 40 different combinatorial selections of the services shown in TABLE I, resulting in 40 pairs of $(V_P, V_C)$, represented by

TABLE II
PARAMETERS NUMERICAL MODEL PHASE 3, STABILIZE

| **Service** | **Capacity increase** | **additional fee customer $C_{ServiceFee}$** | **additional cost provider $C_{ServiceCosts}$** |
|---|---|---|---|
| $D_3 = 5$ | 0.1 | 5000 | 10000 |
| $D_3 = 4$ | 0.05 | 500 | 8000 |
| $D_3 = 3$ | 0.05 | 1000 | 200 |
| $D_3 = 2$ | 0.01 | 100 | 1000 |
| $D_3 = 1$ | 0 | 0 | 0 |

the 40 points in the scatter plot shown in Fig. 4. Extending the discussion of [1], we can now specifically indicate for each point of the scatter plot by which service constellation of TABLE I these values are created.

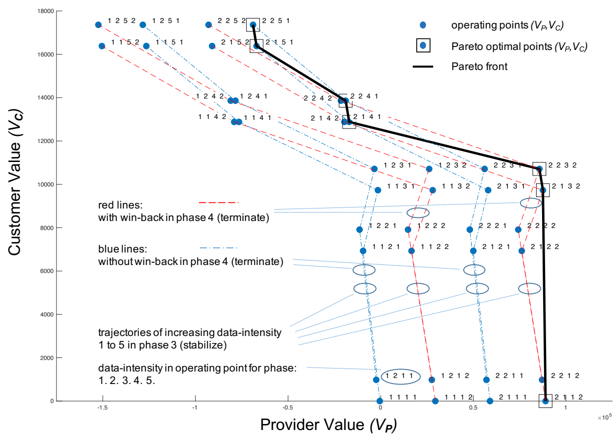


Fig. 4. Selected operating points and Pareto front for $V_P$ and $V_C$.

The numerical lables of the points indicate the data-intensities $[D_1, D_2, D_3, D_4]$. For the sake of visualization, the points in the scatter plot are connected along increasing data intensity $D_3$ in phase 3 (stabilize) by (dotted) dash lines. For each of these trajectories, the data-intensities in the other phases remain unchanged. Apparently, the points with $D_3 = 5$ and $D_3 = 4$ are at the top left end of these lines, i.e., creating high value for the customer $V_C$, but considerably less for the provider $V_P$. This can be interpreted given the numerical values described in subsection IV-B, because condition based maintenance and performance optimization are offered below their costs and provide a high value to the customer.

As described in subsection III-E, the optimum values can be found on the Pareto front of this scatter plot. The combinations of $(V_P, V_C)$ that are below the Pareto front can be improved both for the provider and the customer, hence are sub-optimal. Economically optimum operating points are therefore all on the Pareto front. As already discussed in [1], the provider still needs to decide where on this front it wants to operate, which is basically given by its strategy. Walking along the Pareto front from the bottom right in the figure to the top left, the first three optima up to data-intensity $D_3 = 3$ have the win-back measures in phase 4 enabled ($D_4 = 2$). In Fig. 4, the lables $[D_1, D_2, D_3, D_4]$ for these optima are $[2, 1, \mathit{1}, \mathit{2}]$, $[2, 1, \mathit{3}, \mathit{2}]$, and $[2, 2, \mathit{3}, \mathit{2}]$. Hence, up to data-intensity $D_3$ in phase 3, win-back in phase 4 clearly creates more value for the entire ecosystem. However, the next four optimal points have the win-back measures in phase 4 disabled. These are the points with data-intensity $D_3 = 4$ or 5 for phase 3 and $D_4 = 1$ for phase 4. Their lables $[D_1, D_2, D_3, D_4]$ are $[2, 1, \mathit{4}, \mathit{1}]$, $[2, 2, \mathit{4}, \mathit{1}]$, $[2, 1, \mathit{5}, \mathit{1}]$ and $[2, 2, \mathit{5}, \mathit{1}]$.

The fact that win-back measures would reduce value in these upper four Pareto optima can be explained by Eq. 8 and TABLE II: the high negative contribution margin - given by the high service costs $C_{ServiceCosts}$ as opposed to the low service fees $C_{ServiceFee}$ - penalise the entire ecosystem and diminish the value $V_{P,4}$ substantially.

Whether a provider customer constellation rather selects a Pareto optimal point towards the bottom right side or the top left, depends on the competitive situation of the market. In markets with high bargaining power of the customer, providers will rather select points with higher $V_C$ and lower $V_P$ and vice versa.

## V. DISCUSSION AND OUTLOOK

In this paper, the methodology introduced in [1] for optimally designing data-driven services of different levels of data-intensities was further developed to incorporate specific service designs.

By assigning specific services with their associated data intensities $[D_1, D_2, D_3, D_4]$, the model allows to determine exactly which service design will lead to which value creation for the provider and the customer. The scatter plot of these values $(V_P, V_C)$ (Fig. 4) helps to evaluate how mutual value creation changes when these data-intensities are varied. In particular, it becomes visible how varying the data-intensity of one phase (e.g., $D_3$ for phase 3) - while keeping the others fixed - impacts the value creation.

A hypothetical use case derived from an empirical study was developed incorporating specific services and their levels of data-intensity for the phases of the lifecycle. For this example and the chosen parameters, the model shows that applying win-back measures in phase 4 (terminate) creates high value for the provider by extending the customer lifetime, but is Pareto-optimal only for low data-intensities in phase 3 ($D_3 = 1, ..., 3$). Since higher data-intensities ($D_3 = 4$ or 5) have a negative impact on the contribution margin for the provider, it becomes favourable to apply no win-back measure in these cases. This is due to the assumption made for this example that the advanced services "condition based maintenance" and "performance optimization" are new development fields and therefore create high costs for the provider while the customers' trust in the services and willingness to pay are still low. Noteworthy, the model reveals that there are indeed Pareto optimal points with these advanced services enabled, but only if no win-back measures are applied.

The model obviously shows that due to non-linearities it is not easy to predict which service configuration is Pareto-optimal. For this reason, this integral approach is essential, since it takes into account all phases of the lifecycle, all service intensities as well as the impact on value and costs for providers and customers in a combined way.

The new model has several practical implications. It shows that heuristically chosen investment in leveraging data for service value creation does often not lead to conscious and optimized value creation. Companies need to be aware of the costs and benefits of using data in the different phases of the lifecycle, thus being enabled to design their services in an informed way. Additionally, companies become conscious that the optimum consists of a set of many possible solutions along the Pareto front, and that the choice of a specific solution is based on strategic criteria depending on the competitive situation. Moreover, companies can decide on development roadmaps for their data-driven services when moving along the trajectories shown in Fig. 4 and thus, e.g., consciously move towards providing more customer or provider value when departing from an initial solution (i.e., a specific point on the trajectory).

Further research should address open issues by:

- The development of methods for determining the parameters and functions of the model based on available company data.
- The development of methods for quantitatively assessing the effect of data-based services on operational behavior like, e.g., the capacity increases discussed in TABLE I. This may also incorporate the concepts available in the predictive maintenance literature and potentially will require Monte Carlo simulation approaches.
- Extending to an ecosystem perspective. The model proposed in this paper has its focus on the dyadic provider-customer relation. Given this, the model should be extended to cover mutual value creation among multiple actors in an ecosystem.

## Acknowledgment

The authors would like to thank ZHAW digital for the support of this work. This study was also partly supported by the grant ”ABH097” within the framework of the Interreg VI-programme ”Alpenrhein-Bodensee-Hochrhein” (DE/AT/CH/LI) whose funds are provided by the European Regional Development Fund (ERDF) and the Swiss Confederation. The funders had no role in study design, data collection and analysis, decisions to publish, or preparation of the manuscript.

## References

[1] J. Meierhofer and C. Heitz, “Service Customization: Optimizing Value Creation and Capture by Designing the Customer Journey,” in *2021 8th Swiss Conference on Data Science (SDS)*, Jun. 2021, pp. 32–38.

[2] J. Meierhofer, R. Benedech, L. Schweiger, C. Barbieri, and M. Rapaccini, “Quantitative Modelling of the Value of Data for Manufacturing SMEs in Smart Service Provision,” *ITM Web of Conferences*, vol. 41, p. 04001, 2022, publisher: EDP Sciences. [Online]. Available: https://www.itm-conferences.org/articles/itmconf/abs/2022/01/itmconf_iess2022_04001/itmconf_iess2022_04001.html

[3] C. Kowalkowski and W. Ulaga, *Service strategy in action: A practical guide for growing your B2B service and solution business*. Service Strategy Press, 2017.

[4] T. Baines and H. Lightfoot, “Made to Serve: How Manufacturers can Compete Through Servitization and Product Service Systems | Wiley,” 2013.

[5] A. Osterwalder, Y. Pigneur, G. Bernarda, and A. Smith, *Value Proposition Design: How to Create Products and Services Customers Want*. John Wiley & Sons, Oct. 2014.

[6] A. Tukker, “Eight types of product–service system: eight ways to sustainability? Experiences from SusProNet,” *Business Strategy and the Environment*, vol. 13, no. 4, pp. 246–260, 2004, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bse.414. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/bse.414

[7] F. Adrodegari, N. Saccani, and M. Rapaccini, “PSS Business Models: A Structured Typology,” in *The Palgrave Handbook of Servitization*, M. Kohtamäki, T. Baines, R. Rabetino, A. Z. Bigdeli, C. Kowalkowski, R. Oliva, and V. Parida, Eds. Cham: Springer International Publishing, 2021, pp. 57–71. [Online]. Available: https://doi.org/10.1007/978-3-030-75771-7_4

[8] J. Meierhofer, C. Heitz, and F. Hannich, “Optimizing Service Value Creation with Smart, Connected Products,” in *Proceedings of the 2021 Naples Forum on Service*, Capri, 2021, p. 13. [Online]. Available: https://naplesforumonservice.com/resources/#

[9] W. Reinartz, M. Krafft, and W. D. Hoyer, “The Customer Relationship Management Process: Its Measurement and Impact on Performance,” *Journal of Marketing Research*, vol. 41, no. 3, pp. 293–305, Aug. 2004, publisher: SAGE Publications Inc. [Online]. Available: https://doi.org/10.1509/jmkr.41.3.293.35991

[10] J. D. Chandler and S. L. Vargo, “Contextualization and value-in-context: How context frames exchange,” *Marketing Theory*, vol. 11, no. 1, pp. 35–49, Mar. 2011, publisher: SAGE Publications. [Online]. Available: https://doi.org/10.1177/1470593110393713

[11] M. E. Porter and J. E. Heppelmann, “How Smart, Connected Products Are Transforming Competition,” *Harvard Business Review*, Nov. 2014, section: Competition. [Online]. Available: https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition

[12] J. Meierhofer, L. Schweiger, J. Lu, S. Züst, S. West, O. Stoll, and D. Kiritsis, “Digital Twin-Enabled Decision Support Services in Industrial Ecosystems,” *Applied Sciences*, vol. 11, no. 23, p. 11418, Jan. 2021, number: 23 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2076-3417/11/23/11418

[13] J. C. Sweeney and G. N. Soutar, “Consumer perceived value: The development of a multiple item scale,” *Journal of Retailing*, vol. 77, no. 2, pp. 203–220, Jun. 2001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022435901000410

[14] D. L. Moody and P. Walsh, “Measuring the Value Of Information-An Asset Valuation Approach.” in *ECIS*, 1999, pp. 496–512.

[15] V. Kumar and W. Reinartz, *Customer Relationship Management*, ser. Springer Texts in Business and Economics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2018. [Online]. Available: http://link.springer.com/10.1007/978-3-662-55381-7

[16] K. Miettinen, “Introduction to Multiobjective Optimization: Noninteractive Approaches,” in *Multiobjective Optimization: Interactive and Evolutionary Approaches*, ser. Lecture Notes in Computer Science, J. Branke, K. Deb, K. Miettinen, and R. Słowiński, Eds. Berlin, Heidelberg: Springer, 2008, pp. 1–26. [Online]. Available: https://doi.org/10.1007/978-3-540-88908-3_1

[17] U. S. Karmarkar and G. Roels, “An Analytical Framework for Value Co-Production in Services,” *Service Science*, vol. 7, no. 3, pp. 163–180, Aug. 2015, publisher: INFORMS. [Online]. Available: https://pubsonline.informs.org/doi/abs/10.1287/serv.2015.0103

[18] S. S. Rao, *Reliability Engineering*. Pearson, Jan. 2016.

[19] M. Rapaccini and F. Adrodegari, “Conceptualizing customer value in data-driven services and smart PSS,” *Computers in Industry*, vol. 137, p. 103607, May 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361522000021

[20] “MATLAB Login | MATLAB & Simulink.” [Online]. Available: https://matlab.mathworks.com/