Trustful Data Sharing in the Forest-based Sector - Opportunities and Challenges for a Data Trustee

Lennart Schinke[1,*], Martin Hoppen[1], Alexander Atanasyan[1], Xuebilian Gong[1], Frank Heinze[2], Kathrin Stollenwerk[3] and Jürgen Roßmann[1]

[1]Institute for Man-Machine Interaction, RWTH Aachen University, Ahornstraße 55, 52074 Aachen, Germany
[2]RIF Institute for Research and Transfer e.V., Joseph-von-Fraunhofer-Straße 20, 44227 Dortmund, Germany
[3]ComConsult GmbH, Pascalstraße 27, 52076 Aachen, Germany

Abstract

In addition to their economic value, multifunctional forests fulfill both ecological and social tasks. Thus, sustainable forestry impacts different areas. Although basic digitalization already exists in the forest-based sector and is being developed further, the integration of heterogeneous systems and the sharing of data between the stakeholders continues to be a challenge that requires mutual trust. To meet this demanding requirement, data trustees represent a possible approach. Thereby, the complexity of trustful data sharing is delegated to an intermediary that provides an easy entrance for data providers and consumers as well as software developers. The paper at hand presents the preliminary findings of an ongoing study on trustful data sharing in the forest-based sector in Germany. Based on a general outline of the concept of data trustees, this contribution examines various opportunities for the forest-based sector, identifies the associated challenges from a technical, user, and legal perspective, presents the use case of trustful sharing of harvester production data, and proposes an architecture to encounter the challenges while unlocking the opportunities.

Keywords

Data Trustee, Forestry, Trustful Data Sharing

1. Introduction

Sustainable forestry is a cornerstone of the Green Economy [1] and the European Green Deal [2]. It allows to meet the demand for wood for construction or as a basis for bio-based products. Besides economic aspects, multifunctional forests fulfill ecological and social tasks ("three pillars of sustainability"). While, fundamentally, digitalization already exists in the forest-based sector, one challenge is the integration of diverse systems and the sharing of data between stakeholders [3]. Due to the strong heterogeneity and different market interests of the involved actors (forest owners, service providers, buyers, environmentalists ...), mutual trust is a key requirement and challenge [4]. Data trustees represent an approach to delegate the complexity of trusted data sharing to an intermediary, providing an easy entrance for data providers and consumers, as well as software developers.

The paper at hand presents the preliminary findings of an ongoing study on trustful data sharing in the forest-based sector in Germany, focusing on the opportunities and challenges it presents and showing first results on how to deal with the latter in order to take advantage of the former. In Section 2, it starts with an overview of the state of the art in the field of data trustees and current data sharing approaches in the forest-based sector. Section 3 points out the opportunities and Section 4 the challenges for a data trustee in the forest-based sector. In Section 5, the study's considered use case for trustful sharing of harvester production data is presented. Section 6 shows the suggested current architectural design with a strong focus on the concepts of the International Data Spaces Association (IDSA) and their Reference Architecture Model (RAM) [5]. Finally, the work is concluded in Section 7.

2. State of the Art

Worldwide, digitalization is progressing rapidly. To fully exploit the related potential and optimize various processes, the sharing of data is essential. Different approaches to this have emerged in recent years, many of which are united by the desire for trustful and sovereign data sharing. Therefore, this section gives an overview

Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW'23) — Workshop on Data Ecosystems (DEco'23), August 28 - September 1, 2023, Vancouver, Canada
*Corresponding author.
schinke@mmi.rwth-aachen.de (L. Schinke); hoppen@mmi.rwth-aachen.de (M. Hoppen); atanasyan@mmi.rwth-aachen.de (A. Atanasyan); gong@mmi.rwth-aachen.de (X. Gong); frank.heinze@rt.rif-ev.de (F. Heinze); stollenwerk@comconsult.com (K. Stollenwerk); rossmann@mmi.rwth-aachen.de (J. Roßmann)
0009-0002-5632-5464 (L. Schinke); 0000-0002-9021-1551 (M. Hoppen); 0000-0002-7578-1820 (A. Atanasyan); 0009-0008-7372-6538 (X. Gong); 0000-0002-7792-8523 (F. Heinze); 0000-0002-8343-6939 (K. Stollenwerk); 0000-0002-8780-855X (J. Roßmann)

of data trustees and data spaces and shows how they can be distinguished from each other. It also provides a short overview of data trustees that are currently developed or already in use. Finally, aspects of data handling in the forest-based sector are shown.

2.1. What Is a Data Trustee?

In order to address the desire for trustful and sovereign data sharing, especially between unknown actors, data trustees have become increasingly important in recent years. Yet, despite their basic promise to simplify the conditions of data sharing, they lack a clear definition [6, 7]. Additionally, Lauf et al. [8] state that there are various other terms like "data cooperatives", "data stewardships", or "data brokers" whose definitions only partially overlap with concepts of data trustees and which yet are used as synonyms. Consequently, it is essential to ensure a common understanding of the term "data trustee".

Reiberg et al. [6] show that the focus in literature is either on managing or sharing data. The former is closer to common definitions of the general term "trustee", which sees trustees as entities that manage or control property for other entities [9]. This is consistent with what Lindner et al. [7] call the narrow understanding of a data trustee, where they compare a data trustee to a bank, which, instead of assets, manages data in a fiduciary relationship. On the other hand, in a wide understanding, data trustees are seen as an instance of trust in the sense of a "trusted third party" concept. According to this, a data trustee is an independent organization that offers an infrastructure and, optionally, services to enable trustful data sharing between data providers and data consumers. In this case, the final decision on whether to share the data is made by the actors themselves [7]. Consequently, this wide understanding focuses more on sharing data than managing data.

Comparing these two perspectives, it is easier to imagine a data trustee managing data without sharing it than thinking of examples where a data trustee shares data but does not manage it. Therefore, it is possible to argue that a definition with a focus on managing data is more all-encompassing than only focusing on data sharing [6]. To combine both perspectives, data trustees can generally be defined as "institutions that manage data or rights to data on behalf of and in the interest of others. In the course of their activities, trustees obtain control over data and then use it immediately or at a later point in time to enable the data provider or third parties access to it" (translated from [6], p. 7). This is also consistent with the authors' understanding of a data trustee.

The design of a data trustee, especially when using this general definition, still leaves a lot of room for interpretation, e.g., regarding functional scope or technical components. A generalized starting point for the design of a data trustee is mentioned in [10]. However, in order to meet the needs of the involved actors, the exact implementation can only be determined with direct reference to the field of application. For the healthcare sector, for example, Lauf et al. [8] propose four archetypes with different goals, "Data Brokerage Trustee", "Data Processing Trustee", "Data Aggregation Trustee", and "Data Custody Trustee".

In principle, though, it can be stated that a data trustee should act neutrally. This does not mean that a trustee cannot have any interests of its own, nor that any third-party interests must be ignored or disregarded. Instead, it means that a balance must be established between the interests of the different actors. Furthermore, every data trustee should establish an environment of trust. This includes ensuring data security and transparency. Here, the latter refers to both the use of data and the activities of the trustee itself [6, 7].

2.2. What Is a Data Space?

The idea of data spaces emerged when it became apparent that a central data storage solution cannot solve the problem of data handling any longer. Organizations had to handle ever increasing numbers of diverse data sources [11]. It was thus not possible to physically integrate all the data into a single data base, but instead chosen to leave it at the source and achieve integration on a semantic level.

Over the years, various initiatives have started to develop data space concepts with different focuses. To combine forces, twelve partners, including IDSA and Gaia-X AISBL, set up and operate the European Commission-funded Data Spaces Support Centre (DSSC) [12]. Their goal is to establish data spaces in multiple sectors, while enabling an interoperable data sharing environment. For this, the DSSC investigates the needs of various data space initiatives, develops guidelines for common data spaces, like security requirements or cross-sector standards for data sharing, and offers support for the deployment of data spaces [13]. In order to achieve a common understanding, the DSSC provides a glossary [14], in which a data space is defined as an "infrastructure that enables data transactions between different data ecosystem parties based on the governance framework of that data space" (p. 5).

When setting up a data space, it should be considered that this topic is extensive. For example, the IDSA structures their RAM [5] into business, functional, process, information and system layers, while Otto et al. [15] identify a business-oriented and a legal perspective in addition to a technical one. The business-oriented definition focuses on the role of data spaces as a collaboration format between different organizations. The legal perspective sees data spaces as intermediaries for data

sharing comparable to a data trustee in a wide sense.

2.3. What Is the Relationship Between Data Trustees and Data Spaces?

Section 2.1 and Section 2.2 show that there are various definitions and perspectives on data trustees and data spaces. Thus, there are overlaps of varying degrees, which is why it is currently not possible to clearly separate data trustees and data spaces. Nevertheless, a differentiation can be made based on the view of purposes. Data trustees are seen as institutions that manage data or the rights to data, while data spaces are seen as infrastructures that enable data transactions (cf. Section 2.1 and Section 2.2).

Thus, although it can be argued that data trustees form data ecosystems, it is not always possible to call them a "data space", due to their stronger focus on managing data. Yet, data trustees can be part of a data space or even enable it in the first place, since they represent a component that enables trustful data sharing. From this, it can be concluded that although it is not mandatory for data trustees to build on the principles of data spaces, it is an advantage with regard to future application possibilities (e.g., being part of a data space, enable sharing with data spaces) [6].

2.4. Existing Data Trustees

Lindner et al. [7] show that, despite the existing demand for such solutions, the realization of data trustees in practice is still low. Nevertheless, there are already data trustees in operation (in Germany eight in total), focusing on particular sectors. Other solutions are currently under development, so that there will likely be different realizations in the near future, for example, in healthcare, aviation, maritime industry, manufacturing, logistics, automotive and mobility [7]. Some concrete examples are presented below.

Regarding the B2B sector, there are mainly customized solutions for business processes. Providers of such centralized solutions include, for example, the German Bundesdruckerei with CenTrust [16], the Bochum-based company DATATRUSTEE [17] or Nortal (healthcare) [18]. A cross-domain solution is the Data Intelligence Hub [19] provided by T-Systems International GmbH [7]. Building on data space technologies, T-Systems promises that it is possible to connect to any data space and exploit the value of a provider's data by sharing it on the provider's terms and delivering it securely. As a pioneer of cross-domain data sovereignty, the Data Intelligence Hub is currently in (early) regular operation [19]. The Mobility Data Marketplace (MDM) is a neutral exchange platform that utilizes secure data and communication standards (e.g., brokers for data exchange, prevalent internet protocols, certificates, signatures, process logging) and acts as the National Access Point for mobility data in Germany. Although it is labeled as a "market place" and not as a data trustee, there are some parallels. Transparent conditions as well as sharing data reliably and securely are aspects that are in the focus of the MDM but also part of a data trustee [20]. In addition, both the MDM and a data trustee can be central components of data spaces (cf. [21] and Section 2.3). This shows that regardless of the naming, there is a need for data trustee functionalities.

2.5. How Does the Forest-Based Sector Handle Data Today?

The amount of data created in the forest-based sector is huge. Examples of large data sources are inventory data of forest stands, machine data produced by forest machines, and environmental data produced by different types of sensors. Data handling in the forest-based sector is closely connected to the use of geoinformation. The INSPIRE (Infrastructure for Spatial Information in the European Community) directive (2007/2/EC) [22] of the European parliament establishes a legal framework for the use of geoinformation in Europe. It basically obliges the member states of the EU to provide free access to different types of geodata.

Although some open data standards exist, they are not applied consistently along the value chain, and only small parts of the data are standardized. Internationally, the most important standards are StanForD or StanForD2010 [23] for forest machine data and papiNet [24] for the timber supply chain with a focus on the paper industry. For data sharing in timber handling, there are a number of national standards like FHPDAT in Austria, eFIDS [25] in Great Britain and ELDAT/ELDATsmart [26] in Germany. For forest data and forest management national standards are evolving, like Forestand [27] in Sweden, while in Germany, no such standard exists yet.

The relationship of the actors in the forest-based supply chain varies greatly among countries and regions. Whereas in Germany and Austria the supply chain often consists of many small independent companies, in Sweden and Finland larger companies usually control the whole supply chain from forests to the final wood product [3]. These differences have a high impact on the different mechanisms for data sharing. E.g., in Sweden, Biometria collects large parts of the forest machine data and provides a central platform for data providers to access their data. However, this model is more like a cloud storage for own data than a data trustee for data sharing.

The large Scandinavian machine manufacturers have built up manufacturer ecosystems used by stakeholders owning or using forest machines (e.g., John Deere TimberManager [28], Komatsu Maxifleet [29], PONSSE Opti [30]). Thus, sharing forest machine data is possible inside these ecosystems, but it is difficult to integrate

data from different manufacturers. Hartsch et al. [31] analyze legal, social and economic requirements for using forest machine data to improve timber supply chains in Germany. They identify legal issues and a missing technical infrastructure as main obstacles for sharing forest machine data but agree on the benefits of data sharing for the supply chains.

Digital solutions adopted in the forest-based sector often apply to an isolated process [3]. Examples for such isolated processes are described by Rönnqvist et al. [32]. Thus, data sharing is usually either implemented in a proprietary manner or uses outdated, insecure technologies such as attachments in emails. Attempts to find general solutions for open data sharing infrastructures are rare. Chen et al. [33] use open source software to develop an infrastructure for communication in the forest-based sector. Their work served as starting point for the presented approach.

3. Opportunities

In general, a data trustee can: enable new cooperation based on trustful sharing of highly sensitive business data, address concerns regarding legal certainty of data and its usage, reduce transaction costs by providing a single access point to several data sources, protect the data sovereignty of its users, counteract market distortions by making previously locked-in data available, or simplify data-driven research activities [7].

The concrete opportunities for a data trustee in the forest-based sector are manifold. The following examples and applications are based on the authors' experience from many research projects in the forest-based sector and numerous discussions with stakeholders. In general, digital data and its collection is already ubiquitous in this domain. As one form of application, a data trustee can serve as the trustful intermediary for targeted sharing among immediate, mainly known partners along the value chain. Here, the main opportunities for data providers and consumers are the simplification of data sharing in existing business relations as well as the opportunity to assert existing legal bindings on a technical level in terms of usage control.

An example case for this form of application deals with the promotion of consulting services for private forest owners. To open the market to independent tenderers, recent changes in German legislation made the previous approach of many state forest authorities illegal to offer consulting services at a subsidized rate. In North-Rhine Westphalia, this was encountered with a twelve-step process [34] to apply for public funding that incorporates extensive data sharing. In this scenario, a data trustee could provide the necessary trust among the business partners (forest owner and forest management association, service provider, and funding agency) to ensure the shared data (activity reports, reports on expenditure, invoices, funding decisions) is only used in the intended way.

Aside from targeted data sharing, a data trustee is particularly helpful to provide data to data consumers previously unknown to the provider. This is especially true for scenarios where data needs to be aggregated, anonymized, pseudonymized or otherwise preprocessed to hide original data owners from data users.

An exemplary use case for this scenario would be the provisioning of environmental data from sensors, e.g., on microclimate, soil condition, tree growth, sounds of the fauna etc. A data trustee could be used to aggregate and analyze data from various sensor (network) operators and offer it to various interested parties, e.g., policy makers ("What is the average tree growth rate and soil condition in the Sauerland region?"), forestry contractors ("Are skid trails passable without damage to the soil?"), environmentalists ("What region has a high biodiversity based on sounds from the fauna?"), or dam operators ("What is the influence of the soil moisture of catchment areas on the water reservoir?"). Examples for large installations of such sensor networks are the TERENO network [35] in Germany or the standard environmental monitoring program "ICP forests" [36, 37] in Europe and beyond. Data integration and data analysis are key in this kind of scenario and could be provided by a data trustee.

The case is similar for inventory data describing forest stands. Making this data available, especially on a large scale, would help many stakeholders to make better, informed decisions and to optimize their planning and their processes [38]. A data trustee can provide the necessary, easily accessible environment to aggregate and preprocess the heterogeneous data of large numbers of forest owners (e.g., about 60 % of forested land in the EU is privately owned [39]), while protecting their individual business interests. Sharing this data, a forest owner might be rewarded with better consulting, cheaper management, public subsidies, and the good conscience to support the green transformation.

As a final example, harvester production data holds a high potential for the improvement of many processes in forestry [40, 41, 42]. The standardized data format (cf. Subsection 2.5) makes it easy to handle and analyze. A data trustee can be used for targeted data sharing along the wood supply chain (forest owner placing the felling order, contractor performing the felling order etc.) in a trustful way, e.g., by trustfully handling sensitive business data like (GDPR-related) personal information of the machine operator, the forest owner's price matrix, or the contractor's exact performance and process data. Like sensor or inventory data, production data is also interesting for third parties, as it provides valuable ground truth data directly from the forest. For example, wood buyers

Table 1
Overview of opportunities in forest-based sector

	Opportunities
Targeted data sharing	Simplifying the data sharing between data providers and consumers in existing business relations, and asserting existing legal bindings on a technical level in terms of usage control
Potential data consuming	Ensuring trust between (previously unknown) actors and providing data to potential consumers by aggregating, anonymizing, pseudonymizing or pre-processing data (hiding sensitive data), e.g., sensor data aggregation, inventory data [38] and harvester production data sharing [46]
Digital participation	Open to all kinds of enterprises from the forest-based sector, especially for the large number of small enterprises

like sawmills are interested in large-scale production data to derive information about forest regions, e.g., in order to open up new markets. Harvester production data can be used to predict inventory attributes [43], estimate tree composition and volume [44], or model forest volume [45]. In this context, a data trustee can provide the necessary means to protect individual forest enterprises' data and, thus, their business interests. The value of the data is already discussed and quantified [46]. A data trustee can also provide the necessary means to monetize its provision.

Finally, a data trustee offers the opportunity to allow a digital participation for all kinds of enterprises from the forest-based sector, especially for the large number of small enterprises, by taking care of the challenges (see next section) of trustful data sharing. A short overview of the presented opportunities is shown in Table 1.

4. Challenges

Compared to other sectors such as supply chain or the automotive sector, a data trustee in the forest-based sector is confronted with general and domain-specific challenges, often rooted in the sector's heterogeneity (Figure 1). For a data trustee in the forest-based sector, three main perspectives are taken into account: infrastructural perspective (Section 4.1), user perspective (Section 4.2) and legal perspective (Section 4.3). To summarize this chapter, a short overview of challenges in the forest-based sector is shown in Table 2.

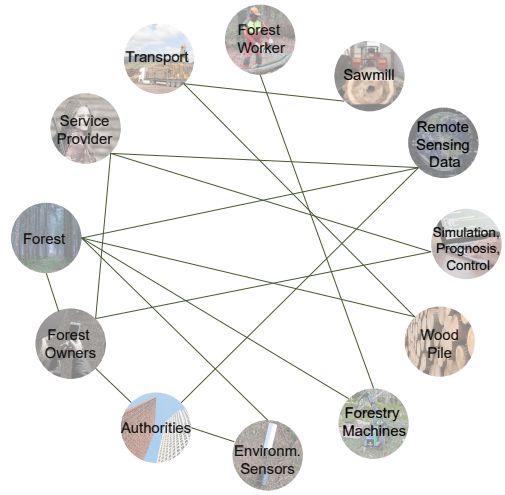

Figure 1: Illustration of the forest-based sector's heterogeneity leading to various intended data sharing relationships.

4.1. Infrastructural Perspective

From the infrastructural perspective, general challenges for a data trustee lie in its ability to serve as a trustful data sharing infrastructure that guarantees secure and reliable data sharing as well as sufficient data quality, enables fair collaboration among participants, and maintains neutrality and transparency [7].

In the context of the forest-based sector, data trustees encounter additional challenges specific to this domain. The first challenge for a data trustee in the forest-based sector pertains to unreliable internet connectivity [47]. Regularly, forest machines and individuals affiliated within the forest-based sector engage in work activities within geographically isolated regions, where a permanent internet connection is hard to guarantee [48]. Expecting the return of forest machines and subsequent data transmission inevitably leads to information delays and thus potential loss of benefits.

Secondly, the forest is a transient environment [49]. In contrast to environments like the factory floor, the conditions in the forest are constantly changing. Even with a reliable internet connection, this dynamic environment presents a challenge in facilitating real-time data sharing. For instance, dense trees and vegetation in the forest may lead to signal attenuation or blockage, wildlife may touch or damage sensors or devices, and severe weather conditions, including heavy rain or strong winds, may interrupt signal transmission, and spontaneous incidents such as fire accidents can cause extensive damage to equipment and further disrupt the signal transmission [50]. For a data trustee, therefore, developing strategies to address this ever-changing forest

environment becomes imperative.

Moreover, a distinguishing feature in the forest-based sector is its multitude of independent and heterogeneous systems (see Figure 1) [47]. As previously discussed in Section 3, data trustees provide the ability for various forest actors to aggregate and preprocess the large amount of heterogeneous data. However, this presents an inevitable challenge for a data trustee as well. Stakeholders and actors employ specialized machines, devices and software applications or services in different heterogeneous systems. In the past few years, a diverse array of digital technologies, including RFID, GPS-based tracking devices, and light detection and ranging (LIDAR), have been effectively utilized to gather data in the forest-based sector [3]. This data generated from heterogeneous systems often exhibits non-uniform data formats, which raises several issues that data trustees need to address, such as how to efficiently collect diverse data sets from various systems and make most of them usable, how to preprocess and analyze data sets with non-harmonized formats, as well as how to aggregate and integrate this data and deal with data compatibility issues.

Additionally, as elucidated in Section 3, the involvement of a data trustee provides a valuable opportunity for participants of various enterprise sizes. Especially for small- and medium-sized enterprises, their interests lie in using a data trustee for data sharing and analysis, thereby expanding their business and enhancing profitability in a cost-effective manner. In contrast, however, these enterprises are disinclined to invest excessive cost and effort in IT infrastructure, and they often lack deeper IT proficiency. Consequently, from a technical perspective, a critical requirement for a data trustee in the forest-based sector is to facilitate a streamlined and accessible process for these potential users to seamlessly join and leverage the infrastructure.

Last but not least, the forest-based sector poses a challenge for a data trustee in terms of organizing data sharing processes involving multiple actors, while ensuring transparency and traceability. In Germany, for example, the harvesting of timber is typically carried out by contractors on behalf of the forest owners. The contractor-owned harvesters collect a wealth of data on parameters such as length, thickness, and quality of the individual logs during felling and subsequent processing [51]. This scenario exemplifies the complexity faced by data trustees: when one of the users of the data trustee expresses interest in harvester production data, the forest owner, contractor and harvester all become associated with this data set. The involvement of multiple actors in a single data sharing process significantly increases the complexity for a data trustee, as it needs to organize the workflow in a proper manner and navigate transparency as well as traceability requirements at different levels. This presents a crucial requirement for a data trustee in the forest-based sector to effectively manage and coordinate data sharing processes.

4.2. User Perspective

From a user perspective, general challenges of a data trustee are whether data providing users are willing to delegate usage control to a third party, whether data consuming users can rely on quality and legal certainty of provided data, or whether the actual added value of the data trustee can be clarified to ensure its day-to-day operation on a permanent basis and to expand it as far as possible [7].

Focusing on potential users in the forest-based sector, there are specific challenges to overcome for the implementation of a data trustee, as well.

Many people employed in forestry are older (e.g., in 2019, more than a quarter were aged above 50 years [39]). Likewise, private forest owners are older (e.g., a study from 2010 showed a large proportion to be older than 60 years [52]). Although older adults more and more overcome the digital divide, they still lag behind [53]. Besides age, there's a digital inequality in terms of the size of companies. An estimated 80-90 % of forest enterprises are Micro-, Small- and Medium-sized Enterprises (MSME) [54]. Due to limited resources, they often do not have extensive IT knowledge or interest in dealing with legal framework conditions.

On the other hand, stakeholders in the forest-based sector - especially in Germany - have major concerns about sharing information with others. The authors' many years of experience from discussions with stakeholders revealed a wide variety of reasons. Contractors fear to reveal information that allows to derive exact performance figures from production data that might give clients an advantage in price negotiations. They also fear intense surveillance regarding adherence to nature conservation guidelines (harvesting measures at the wrong time or place). Forest owners want to protect their critical business data when negotiating with contractors (e.g., price matrix) or buyers like sawmills (e.g., exact locations of preferred wood qualities). Besides, there are conflicts of interest towards environmentalists regarding wood usage versus environmental protection (land set-aside) [4] as well as a general fear of intensive audits by state authorities.

The key requirements for a data sharing solution in the forest-based sector are therefore ease of access, ease of use, and trustworthiness. All these requirements can be met by a data trustee.

4.3. Legal Perspective

The data trustee will handle both personal data and machine generated data.

In the EU and the European Economic Area (EEA), handling personal data is strictly regulated by the General Data Protection Regulation (GDPR). This regulation is part of the EU privacy law and human rights law and focuses on data protection and privacy. It was an inspiration for several other laws concerning the protection of personal data. For instance, the California Consumer Privacy Act (CCPA) possesses many similarities with the GDPR.

The core element of the GDPR is to enable individuals to determine in which way their personal data may be saved, processed or transferred to third parties by data processors. Individuals – data subjects – possess full ownership of their personal data. The GDPR provides a very strict framework concerning, inter alia, rights of data subjects, duties of data controllers or processors, and transfers of personal data to third countries. Once a data trustee has to process personal data, obeying the GDPR is mandatory.

From the legal perspective, the major challenge for a data trustee will be handling machine generated data. In contrast to the strict regulations of the GDPR, the German legislator does not provide any concept of ownership of machine generated data. This lack of regulation causes massive uncertainty and confusion of anybody who is interested in sharing, processing and transferring machine generated data. Data sharing solutions in particular suffer from the nonexistent legal definition. The stakeholders' willingness to offer machine generated data is very low since they suspect the loss of sovereignty of data their machines generated (e.g., losing company secrets, legal implications). As mentioned in the previous section, this is a crucial problem in implementing a data sharing solution in the German forest-based sector.

Consequently, a data trustee that respects everyone's data sovereignty and enables adjusting usage rights for the data entrusted to it seems to be a possible solution for data sharing in the German forest-based sector. This causes two challenges.

Firstly, a definition of an ownership-like claim for machine generated data has to be defined and has to be mandatory within the data ecosystem of the data trustee. One possibility to establish this ownership-like claim is given by the General Terms and Conditions (GTC) that every participant in the data ecosystem of the data trustee has to accept [31]. Moreover, a catalog of usage rights templates needs to be developed and provided for the stakeholders.

Secondly, the data trustee should be able to transparently ensure the technical enforcement of the established usage rights as far as technically possible. This might be the major challenge.

Another legal challenge a data trustee has to deal with is the recently enacted EU Data Governance Act (EU DGA). It defines conditions for providing data intermediation services and will significantly influence the design and configuration of data trustees within the EU and EEA in the following years. It remains to be seen if the EU DGA will promote the acceptance of data sharing solutions or if its strict regulations for providing a data sharing solution will slow down the development of new data trustees [7].

Table 2
Overview of challenges in forest-based sector

	Challenges
Infrastructure perspective	Unreliable internet connectivity [48]
	The forest is an ever-changing environment [49]
	Numerous independent heterogeneous systems and non-uniform data formats [47]
	Providing users with a simple and convenient joining and usage process
	Organizing data sharing processes involving multiple actors while ensuring transparency and traceability
User perspective	Many forestry practitioners and private forest owners are older [39] [52]
	Requirements for low technological barriers and simplified onboarding process
	Stakeholders have major concerns about sharing information with others due to various reasons
Legal perspective	Lack of ownership concept for machine-generated data
	Mandatory definition of ownership-like claim for machine-generated data [31]
	Ensuring technical enforcement of the established usage rights transparently
	Compliance with EU Data Governance Act (EU DGA) [7]

5. Use Case

As mentioned in Section 2.1, developing a data trustee requires a reference to the field of application. For this, the authors consider the use case "trustful sharing of harvester production data". Its implementation with a data trustee is shown in Figure 2. It consists of five actors:

- A *Forest Owner*[1] who wants to sell wood after trees are felled and is therefore interested in production data. He is also willing to share data for research purposes.

[1] from here on, terms in *italics* refer to either elements of Figure 2 or Figure 3

- A *Contractor* who is hired by the *Forest Owner* to fell trees.
- A *Harvester* and its operator. The machine is owned by the *Contractor* and used to fell trees. While felling trees, harvester production data (*HPR*) is generated.
- A *Sawmill* that wants to buy wood and is therefore interested in *HPR* to check what stem sizes are available.
- A *Research Institute* that investigates forests in a certain area and wants to update forest stand data by using *HPR*.

The machine generated *HPR* conforms to the StanForD2010 standard mentioned in Section 2.5 and contains information that relates to different actors or their property. Consequently, different actors have rights to the machine generated data.

What is the advantage to use a data trustee in this use case? The machine-generated data contains information that is on the one hand valuable for the different actors and is on the other hand critical to share, because some of the actors are either not allowed to share the data with other actors (due to the GDPR), or do not want to share this information. Example of critical information includes the identity, working hours, and performance of the harvester operators, which can only be shared with the explicit consent of the harvester operator according to the GDPR. The forest owner can use the machine-generated data for operational control and to take account of how much timber is actually removed from his forest and how much timber should still be left. The *Sawmill* can use the data to improve the accuracy of the delivery forecasting and for organizing the timber logistics. The *Research Institute* might use the data to account for the harvested timber. Finally, the *Contractor* can be relieved of manually handling the data sharing, as it is often the case today.

The red dashed lines in Figure 2 show the communication outside the *Data Trustee*. In the first step, the *Forest Owner* sends a *Job* to the *Contractor* together with what he considers to be valid usage rights (*UR*) on the *HPR*, e.g., "filter out tree coordinates". The *Contractor* adds *UR* from his perspective, e.g., "filter out personal data of the harvester operator", and sends the package to the respective *Harvester*. Based on the *UR* defined by *Forest Owner* and *Contractor*, different offers (*O*) are created, one addressing the *Forest Owner* (left *O*, light blue), one addressing the *Contractor* (middle *O*, dark blue) and a general one, addressing a general audience (right *O*, green). After finishing the *Job*, the *Harvester* automatically combines the three offers with the *HPR* and sends the resulting package to the *Data Trustee*. Subsequently, other actors can view the respective offers. If an offer is accepted, an agreement (*A*) is reached. Depending on the agreement, the actors can use certain parts of the *HPR* in different ways. For example, the *Forest Owner* is not able to see the personal data of the harvester operator, while the *Contractor* is not able to see the tree positions. All other actors who agree to the "general offer" obtain identical rights to the *HPR*. Only the agreements differ in the respective assignee, e.g., *Sawmill* or *Research Institute*.

The focus of the initial development is supposed to be on trustful data sharing. The complexity is significantly increased if several actors can edit a data offer. For this reason, the first step assumes that *UR* are coordinated outside the *Data Trustee*, so that an offer is created on behalf of all actors who have rights to the data. However, the integration of the collaborative creation of a data offer into the *Data Trustee* is of crucial importance for both usability (e.g., tool for creating machine-readable *UR*) and security (e.g., ensuring that one's own *UR* are set exactly in the desired way) and should therefore be considered in the near future.

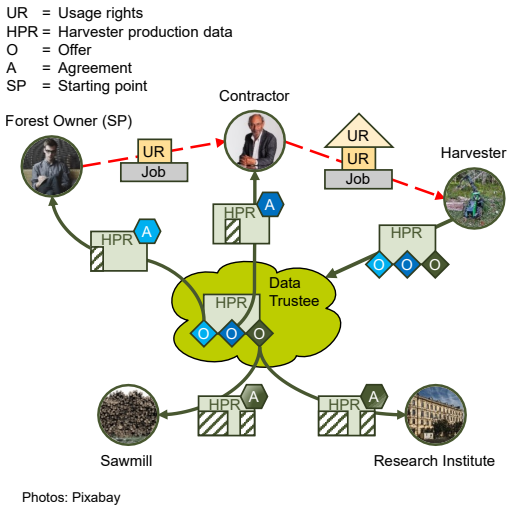

Figure 2: *Data Trustee* use case "trustful sharing of harvester production data".

6. Architecture

As shown in Section 4, a data trustee for the forest-based sector needs to address specific challenges that might not be present in other industries like an unreliable internet connection or a very heterogeneous user and developer base, especially in terms of IT infrastructure, IT expertise, or legal expertise. As far as known to the authors, there is no existing data trustee that already fulfills all these aspects. Therefore, this paper proposes a preliminary architectural design regarding the use case mentioned in Section 5. The intention is to leverage the opportu-

nities and tackle the challenges presented in Section 3 and Section 4 centering the architecture around the *Data Trustee* and making use of various components proposed by the IDSA in their RAM [5]. The following sections introduce requirements to support trusted data sharing at the technical and ecosystem levels (Section 6.1), the *Data Trustee's* architecture (Section 6.2) and outline the ongoing and planned implementation of it and the infrastructure around it (Section 6.3).

6.1. Requirements

At the technical level, relevant requirements can be derived from the aforementioned opportunities and challenges (cf. Section 3 and Section 4). The designed architecture needs to support reliable data sharing, provide low technical threshold for user onboarding and usage, and maintain transparency and traceability even in complex scenarios with multiple users. In addition, existing legal bindings must be enforced at a technical level with approaches of usage control.

From the ecosystem perspective, the most important aspect is creating trust, otherwise stakeholders will not be willing to share data. Consequently, an important requirement is providing a trustful environment that is built upon a foundation of mutual trust. Further requirements from the user and legal perspective can be derived from Table 2 but will not be presented in detail here.

6.2. Components and Data Flow

The relationship between data trustees and data spaces is the primary focus of Section 2.3. This section concludes that data trustees act as institutions managing data and their respective rights, while data spaces provide the infrastructure that facilitates data transactions. It further proposes that data trustees can be built effectively using data spaces as a foundation because data space components as defined by IDSA can carry out many of the expected functions of a data trustee. This approach potentially enables data sharing between a data trustee and data spaces. The first conceptual step to implement this approach is the identification of necessary data space components for a minimum viable data trustee.

To this end, Figure 3 shows an overview of the proposed *Data Trustee's* architecture with a focus on the data flow between important components. A part of the ongoing study will be to show, on the basis of an implementation, which data space components are indeed necessary for a minimal viable data trustee, and which are optional.

The basic path of data flow for a Usage Agreement-based data transaction[2] uses the *Data Trustee* as an intermediary: data sources (on the top in Figure 3) provide data and associated rights, which are then handled by the *Data Trustee* (middle) and finally obtained by data users (bottom).

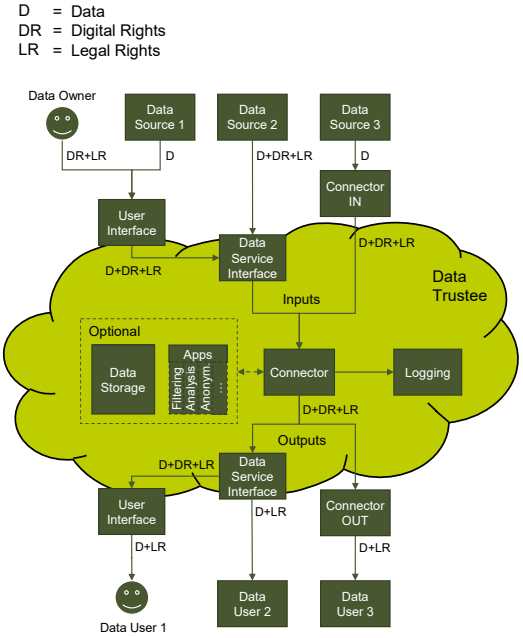

Figure 3: Architecture focusing on the data flow between important components of the proposed *Data Trustee*.

Based on the concepts of the IDSA, where connectors represent the major gateway for communication in data spaces and have the ability to perform usage control as a policy enforcement point [55], the central component of the proposed *Data Trustee* is the *Data Trustee's Connector*.

Even though the use case presented in Section 5 focuses on machine-generated *HPR* and the *Harvester* as the data source, the authors' goal is to consider transactions that can be both manually initiated by human users and automatically triggered by machines. This results in more diverse application possibilities, while at the same time facilitating a system rollout in practice. Following this, the *Data Trustee's* proposed architecture allows different types of data sources and users:

- Human users that require a *User Interface* (*Data Owner* with *Data Source 1* and *Data User 1*), e.g., *Forest Owner* or *Contractor* as an individual with minimal IT knowledge,

[2] In the context of data spaces, data transactions typically follow a process of negotiation or discovery, resulting in a Usage Agreement that outlines the terms of data use. This agreement is established either through active negotiation between the data provider and consumer, or by a consumer discovering the data offer and agreeing to the provider's pre-set terms. The authors refer to this process as a "Usage Agreement-based data transaction".

- Human users that do not require a *User Interface* or machines without data space connectors (*Data Source 2* and *Data User 2*), e.g., *Harvester* that automatically uploads *HPR* after executing a job or *Sawmill* using an own software to automatically analyze *HPR*, and
- Actors with own connectors (*Data Source 3* with its providing connector *Connector IN* and *Data User 3* with its consuming connector *Connector OUT*—named from the trustee's perspective), e.g., *Research Institute* with significant IT knowledge.

While the latter type can communicate via the *Connector* directly, the former two require a *Data Service Interface*, as a part of the *Data Trustee*, that supports protocols to directly communicate with the respective actor or can provide a backend for the *User Interface*. This, in turn, forms the *Data Trustee's* frontend to the human user, providing an interface to create, modify and delete data offers (including the definition of rights), negotiate agreements and to receive or view data. These two types provide seamless opportunities for potential participants to join and utilize the *Data Trustee*. They offer a user-friendly approach, eliminating the need for extensive IT expertise, enabling easy integration with the *Data Trustee*, securing data sharing and maintaining the data sovereignty, particularly for machine-generated data (cf. Section 4.3). The *User Interface* also acts as a singular marketplace for the envisioned application of a common trusted and sovereign data sharing ecosystem. Nevertheless, the vision of data federation suggests that data trustee functionalities might as well be distributed. In this case, the infrastructure needs to be designed for multiple trustees. This introduces additional complexity and challenges, particularly, around ensuring interoperability and managing the coordination and cooperation among different trustees, e.g., handling usage policies. While this will be considered as a perspective during implementation, the focus will lie on a single, scalable data trustee. As mentioned before, the proposed *Connector* forms the center of the *Data Trustee*. Nevertheless, the final implementation might require the inclusion of additional connectors for providers and consumers who do not have an own connector. The *User Interface's* backend is a potential location for these.

The last mandatory component of the *Data Trustee* is the *Logging* functionality. With it, all communication, access to and technically traceable usage of data is stored in the *Data Trustee* to provide transparency and enable accountability, thus, opening the possibility for the monetization of data provision. This allows *Forest Owner* and *Contractor* to track how the data to which they hold rights has been used and, if desired, to generate an invoice for the use of the data.

In addition, the proposed architecture features optional components that are particularly interesting for the realization of the opportunities in the forest-based sector as described in Section 3. An important component to ensure the availability of offered data is *Data Storage*, as many potential participants of a common forest-based data sharing infrastructure are not able to host data on-premise and guarantee continuous access to it. This can be due to missing IT knowledge, an unreliable internet connection (cf. Section 4.1) or sources like the *Harvester* being turned off after usage. Furthermore, many highly beneficial use cases are only possible by removing or modifying sensitive data (e.g., *Harvester* operator names or exact tree positions in *HPR*, cf. Section 4.2), achieved using techniques like filtering, pseudonymization, anonymization, or aggregation. Services to realize these functions or perform more domain-specific calculations like sensor data processing for wood harvest yield forecasting, climate impact analysis, or business process optimization need to be accessible via the *Data Trustee*, making its support for *Apps*[3] sensible. In the given use case, this allows the *Research Institute* to not only analyze data but provide their analysis algorithms, e.g., determination of wood harvest efficiency, as an *App*, e.g., to the *Contractor* for business optimization processes. While both, *Data Storage* and *Apps*, can be considered parts of the *Data Trustee* itself, the *Data Trustee* should at least provide interfaces to existing services with these functionalities. In the sense of decentralization and data sovereignty, this allows the user to choose where data is stored and whose apps are used. In addition, the integration of external apps leads to an increased number of potential uses.

6.3. Implementation Approach

The proposed *Data Trustee* is based on the concepts provided by the IDSA, as the initiative's technology is mature, and their components are already available for secure data sharing. While Gaia-X represents a broader and more modern approach to federated data and service infrastructure, its relative immaturity means that IDSA currently provides a more practical foundation for a data trustee. Importantly, the current choice does not compromise future interoperability with Gaia-X, as both initiatives are member of the DSSC and committed to compatibility.

Connector, *Logging*, and *Apps* are functional blocks of the proposed *Data Trustee* that can be realized using International Data Spaces (IDS)[4] components—connectors, clearing houses, and apps, respectively. Many implementations exist, such as the Eclipse Dataspace Connector [56] and the IDS Clearing House prototype implemen-

[3] In the IDSA sense of code-to-data

[4] IDSA and IDS are related as IDS is the concept for a secure and trusted data sharing environment, while IDSA is the association that works to develop, implement and promote the IDS concept [5].

tation [57]. IDS Apps, due to their domain specificity, are components that have to be developed individually and can be distributed using IDS App Stores (see [58] for an implementation) after a defined certification process. The usage policies to be enforced by the connectors are based on the Open Digital Rights Language (ODRL)[5].

As the proposed architecture focuses on the development of an IDS-compliant data trustee, it does not feature all components necessary to establish a data space but is designed to allow seamless integration into them and relies on further components for its proper functioning. Obligatory components include, in particular, an identity provider to enable authentication and (user- or role-dependent) authorization for data usage. Additionally, a metadata broker is required to facilitate querying the metadata of services (including, e.g., filtering, anonymization, and value-added analysis services) and data offers[6]. Finally, vocabularies serve to ensure a common understanding of the various terms used to describe the data and services provided. This facilitates the analysis and integration of data in different formats.

For the findability of relevant data by human users, another important component is some kind of data marketplace (not a defined component of the IDS infrastructure). From the frontend perspective, as mentioned before, this relies on the *User Interface* that is already part of the *Data Trustee*. The corresponding backend relies on the metadata broker which allows appropriate searches for (meta)data, and, potentially, on vocabularies to prevent misunderstandings by offering user-specific data and service descriptions. This can be particularly useful in the forestry-based sector, as the terms employed there can vary depending on the region and the role of a given participant of the envisioned data sharing ecosystem enabled by the *Data Trustee* (cf. Section 4.1). The authors' approach is to provide the *User Interface* as a web application, as this does not require the user to install any specific software and allows access from different devices, like computers, smartphones or tablets. Furthermore, it can be assumed that users with little IT knowledge, e.g., *Forest Owner* of advanced age (cf. Section 4.2), are more likely to use a web application, if it is self-explanatory, instead of performing a possibly complicated software installation. Consequently, this implementation is intended to keep the entrance barrier for the user low.

7. Conclusion

The forest-based sector plays an important role in the green transformation of our economy. Yet, its digitalization is still restricted due to limited communication and data sharing. The opportunities for the sector by using a data trustee seem manifold—from sharing environmental, process and production data between immediate business partners to optimizing the supply chain to providing added value through secondary use by third parties (scientists, policy makers, environmentalists, dam operators ...), data refinement and monetization. However, the identified challenges, ranging from infrastructural to user-related to legal issues, cannot be ignored. The authors assume the proposed architecture to be a viable approach to resolve these challenges, while unlocking the outlined opportunities for the forest-based sector and enable the evaluation of viable business solutions for a trustful data sharing ecosystem in a domain that is highly diverse with respect to roles and the economical as well as ecological scale.

In the context of their ongoing study, the authors will focus on the presented use case of harvester production data and its sharing between forest owner, contractor, harvester, sawmill and research institute. Their next steps will be the prototypical implementation of the architecture to assess its suitability using practical examples from this use case. This prototype, within a realistic testing environment, will help to understand the practicability of the presented architecture with respect to requirements and expectations.

The authors are aware that there are other issues relevant for the future with regard to data rights that are not addressed in this paper. Due to the focus of the ongoing study on machine-generated data, data generated by other techniques, like machine learning models, is not discussed within the scope of this study. Identifying the differences involved could be the focus of a future study.

Acknowledgments

This work was supported by the Federal Ministry for Education and Research (BMBF), Germany [grant number 16DTM102A..D], and funded by the European Union - NextGenerationEU. The views and opinions expressed are solely those of the authors and do not necessarily reflect the views of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

[5] ODRL is a policy expression language and an endorsed W3C Recommendation since 2018 [59].

[6] Identity provider and metadata broker are not featured in Figure 3 due to its focus on data flow.

References

[1] Advisory Committee on Sustainable Forest-based Industries (ACSFI), Building back a better post-COVID-19 world with sustainable forest products, 2020. doi:10.4060/cb1556en.

[2] European Commission, New eu forest strategy for 2030, 2021. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021DC0572.

[3] J. Scholz, A. D. Meyer, A. S. Marques, T. M. Pinho, J. Boaventura-Cunha, J. V. Orshoven, C. Rosset, J. Künzi, J. Kaarle, K. Nummila, Digital technologies for forest supply chain optimization: Existing solutions and future trends, Environmental Management 62 (2018) 1108–1133. doi:10.1007/s00267-018-1095-5.

[4] T. Baycheva-Merger, Forest policy information networks and the role of trust: Cooperative and competitive orientations and underlying causes, Forests 10 (2019) 359. doi:10.3390/f10040359.

[5] B. Otto, S. Steinbuss, A. Teuscher, S. Lohmann, Ids reference architecture model, 2019. URL: https://doi.org/10.5281/zenodo.5105529. doi:10.5281/zenodo.5105529.

[6] Dr. Abel Reiberg, Dennis Appelt, Peter Kraemer, Datentreuhänder, Datenvermittlungsdienste und Gaia-X, Whitepaper, Gaia-X Hub Deutschland, 2023. URL: https://gaia-x-hub.de/wp-content/uploads/2023/03/WP23.2_Datentreuhaender_DE.pdf.

[7] Maximilian Lindner, Sebastian Straub, Datentreuhänderschaft Status Quo und Entwicklungsperspektiven, Whitepaper, Begleitforschung Smarte Datenwirtschaft, 2023. URL: https://www.iit-berlin.de/wp-content/uploads/2023/02/SDW_Datentreuhand.pdf.

[8] F. Lauf, S. Scheider, J. Friese, S. Kilz, M. Radic, A. Burmann, Exploring design characteristics of data trustees in healthcare — taxonomy and archetypes, 2023.

[9] Dictionary.cambridge.org, Trustee | definition in the cambridge english dictionary, 2023. URL: https://dictionary.cambridge.org/us/dictionary/english/trustee.

[10] M. Stachon, F. Möller, T. Guggenberger, M. Tomczyk, J.-L. Henning, Understanding data trusts, in: 2023 European Conference on Information Systems, 2023.

[11] M. Franklin, A. Halevy, D. Maier, From databases to dataspaces: A new abstraction for information management, SIGMOD Rec. 34 (2005) 27–33. URL: https://doi.org/10.1145/1107499.1107502. doi:10.1145/1107499.1107502.

[12] D. S. S. Centre, Dssc, 2023. URL: https://dssc.eu/.

[13] F. ISST, Facilitating interoperable data sharing, the new data spaces support centre is now launched, 2022. URL: https://www.isst.fraunhofer.de/de/news/pressemitteilungen/2022/PM-12102022_data-spaces-support-centre.html.

[14] Antti 'Jogi' Poikola, Bert Verdonck, Rieks Joosten, DSSC Glossary, Glossary, Data Spaces Support Centre (DSSC) c/o Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V., Hansastr. 27c, 80686 Munich, Germany, 2023. URL: https://dssc.eu/wp-content/uploads/2023/03/DSSC-Data-Spaces-Glossary-v1.0.pdf.

[15] B. Otto, M. Hompel, S. Wrobel, Designing Data Spaces, 2022. URL: https://library.oapen.org/handle/20.500.12657/57901. doi:10.1007/978-3-030-93975-5.

[16] B. GmbH, Data trustee: An independent trust authority, 2023. URL: https://www.bundesdruckerei-gmbh.de/en/solutions/data-trustee.

[17] D. GmbH, Datatrustee, 2023. URL: https://www.datatrustee.org/, available in German only.

[18] Nortal, Trusted third party, 2023. URL: https://nortal.com/digital-healthcare/secure-authentication-in-healthcare/.

[19] T.-S. I. GmbH, Data intelligence hub, 2023. URL: https://dih.telekom.com/en.

[20] B. für Straßenwesen, The mdm, 2023. URL: https://www.mdm-portal.de/about-mdm/?lang=en.

[21] M. D. Space, The mdm in the mobility data space, 2023. URL: https://www.mobility-data-space.de/en/the-mdm-in-the-mobility-data-space.html.

[22] EUR-Lex, Inspire, 2023. URL: https://eur-lex.europa.eu/eli/dir/2007/2/2019-06-26.

[23] J. Arlinger, Stanford, 2023. URL: https://www.skogforsk.se/english/projects/stanford/.

[24] papiNet User Group, papinet, 2023. URL: http://www.papinet.org/.

[25] O. F. I. T. Committee, efids, 2023. URL: https://www.oasis-open.org/committees/download.php/29052/eFIDS-Description.html.

[26] D. Forstwirtschaftsrat, Eldatsmart, 2023. URL: https://www.dfwr.de/projekte/eldat-smart/.

[27] Skogforst, Forestand, 2023. URL: http://www.forestand.org/.

[28] J. Deere, Timbermanager, 2023. URL: https://www.deere.de/de/forstmaschinen/timbermatic-karten/timbermanager/.

[29] K. Forest, Maxifleet, 2023. URL: https://www.komatsuforest.de/services/maxifleet.

[30] PONSSE, Ponsse opti, 2023. URL: https://www.ponsse.com/en/products/information-systems#/.

[31] F. Hartsch, J. Kemmerer, E. R. Labelle, D. Jaeger, T. Wagner, Integration of harvester production data in german wood supply chains: Legal, social and economic requirements, Forests 12 (2021) 460. URL:

http://dx.doi.org/10.3390/f12040460. doi:10.3390/f12040460.

[32] M. Rönnqvist, S. D'Amours, A. Weintraub, A. Jofre, E. Gunn, R. G. Haight, D. Martell, A. T. Murray, C. Romero, Operations research challenges in forestry: 33 open problems, Annals of Operations Research 232 (2015) 11–40. doi:10.1007/s10479-015-1907-4.

[33] J. Chen, M. Schluse, H.-J. Roßmann, Enabling a Secured Communication in Distributed IoT Using the Smart Systems Service Infrastructure, in: 2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), SPT-IoT, 5. IEEE PERCOM Workshop on Security Privacy and Trust in the Internet of Things, Kassel (Germany), 22 Mar 2021 - 26 Mar 2021, IEEE, 2021, pp. 674–679. URL: https://publications.rwth-aachen.de/record/818947. doi:10.1109/PerComWorkshops51409.2021.9431124.

[34] UNIQUE forestry and land use GmbH, Waldbauernlotse, 2023. URL: https://www.waldbauernlotse.nrw/, available in German only.

[35] H. R. Bogena, TERENO: German network of terrestrial environmental observatories, Journal of large-scale research facilities JLSRF 2 (2016). doi:10.17815/jlsrf-2-98.

[36] M. Lorenz, International co-operative programme on assessment and monitoring of air pollution effects on forests-icp forests, Water, Air, and Soil Pollution 85 (1995) 1221–1226.

[37] T. I. of Forest Ecosystems, Icp forests, 2023. URL: http://icp-forests.net/.

[38] A. S. Kangas, Value of forest information, European Journal of Forest Research 129 (2009) 863–874. doi:10.1007/s10342-009-0281-7.

[39] European Commission. Statistical Office of the European Union., Agriculture, forestry and fishery statistics: 2020 edition., Publications Office, 2020. URL: https://ec.europa.eu/eurostat/en/web/products-statistical-books/-/ks-fk-20-001. doi:10.2785/143455.

[40] M. Maltamo, M. Hauglin, E. Naesset, T. Gobakken, Estimating stand level stem diameter distribution utilizing harvester data and airborne laser scanning, Silva Fennica 53 (2019). doi:10.14214/sf.10075.

[41] J. Kemmerer, E. R. Labelle, Using harvester data from on-board computers: a review of key findings, opportunities and challenges, European Journal of Forest Research 140 (2020) 1–17. doi:10.1007/s10342-020-01313-4.

[42] J. Söderberg, J. Wallerman, A. Almäng, J. J. Möller, E. Willén, Operational prediction of forest attributes using standardised harvester data and airborne laser scanning data in sweden, Scandinavian Journal of Forest Research 36 (2021) 306–314. doi:10.1080/02827581.2021.1919751.

[43] A. Saukkola, T. Melkas, K. Riekki, S. Sirparanta, J. Peuhkurinen, M. Holopainen, J. Hyyppä, M. Vastaranta, Predicting forest inventory attributes using airborne laser scanning, aerial imagery, and harvester data, Remote Sensing 11 (2019) 797. doi:10.3390/rs11070797.

[44] J. Rasinmäki, T. Melkas, A method for estimating tree composition and volume using harvester data, Scandinavian Journal of Forest Research 20 (2005) 85–95. doi:10.1080/02827580510008185.

[45] M. Hauglin, E. Hansen, E. Sørngård, E. Næsset, T. Gobakken, Utilizing accurately positioned harvester data: modelling forest volume with airborne laser scanning, Canadian Journal of Forest Research 48 (2018) 913–922. doi:10.1139/cjfr-2017-0467.

[46] A. Kaulen, D. Mayer, Runder Tisch "Datenschatz Forstmaschine", FTI - Forsttechnische Informationen (2022) 18–19. Available in German only.

[47] R. Sahal, S. H. Alsamhi, J. G. Breslin, M. I. Ali, Industry 4.0 towards forestry 4.0: Fire detection use case, Sensors 21 (2021) 694. URL: https://doi.org/10.3390/s21030694.

[48] R. Singh, A. Gehlot, S. V. Akram, A. K. Thakur, D. Buddhi, P. K. Das, Forest 4.0: Digitalization of forest using the internet of things (iot), Journal of King Saud University-Computer and Information Sciences 34 (2022) 5587–5601. URL: https://doi.org/10.1016/j.jksuci.2021.02.009.

[49] J. Reitz, M. Schluse, J. Roßmann, Industry 4.0 beyond the factory: An application to forestry (2019) 107–116. doi:10.1007/978-3-662-59317-2_11.

[50] A. Salam, Internet of things for sustainable forestry, Internet of Things for sustainable community development: Wireless communications, sensing, and systems (2020) 147–181. URL: https://doi.org/10.1007/978-3-030-35291-2_5.

[51] T. Purfürst, J. Erler, The human influence on productivity in harvester operations, International Journal of Forest Engineering 22 (2011) 15–22. URL: https://doi.org/10.1080/14942119.2011.10702606.

[52] F. Schmithüsen, F. Hirsch, Private forest ownership in europe (2010) vi, 110 p. :. URL: http://digitallibrary.un.org/record/697427, at head of title: United Nations Economic Commission for Europe/-Food and Agriculture Organization of the United Nations, Forestry and Timber Section, Geneva, Switzerland.

[53] J. Francis, C. Ball, T. Kadylak, S. R. Cotten, Aging in the digital age: Conceptualizing technology adoption and digital inequalities, in: Ageing and Digital Technology, Springer Singapore, 2019, pp. 35–49.

doi:10.1007/978-981-13-3693-5_3.

[54] E. Pohnan, B. Cammaert, T. Cavanagh, Enabling micro, small and medium-sized enterprises to participate in legal timber production and trade – Transformational changes generated by the FAO-EU FLEGT Programme, number 189 in FAO Forestry Paper, FAO, Rome, Italy, 2022. doi:10.4060/cc3107en.

[55] L. Zhi, W. Jing, C. Xiao-su, J. Lian-xing, Research on policy-based access control model, in: 2009 International Conference on Networks Security, Wireless Communications and Trusted Computing, volume 2, IEEE, 2009, pp. 164–167. doi:10.1109/NSWCTC.2009.313.

[56] E. Foundation, Eclipse dataspace components, 2023. URL: https://projects.eclipse.org/projects/technology.edc.

[57] F. AISEC, Ids clearing house github repository, 2023. URL: https://github.com/Fraunhofer-AISEC/ids-clearing-house-service.

[58] F. FIT, Ids appstore github repository, 2023. URL: https://github.com/International-Data-Spaces-Association/IDS-AppStore.

[59] S. V. Renato Iannella, Odrl information model 2.2, 2018. URL: https://www.w3.org/TR/odrl-model/.