Forestry Data Trustee

A data trustee to enhance data sharing in the forest-based sector

Lennart Schinke, M.Sc. and Prof. Dr.-Ing. Jürgen Roßmann, Institute for Man-Machine Interaction, RWTH Aachen University, schinke@mmi.rwth-aachen.de

The Forestry Data Trustee (FDT) aims at increasing data sharing in the forest-based sector. Different entry points are provided to address the varying IT capabilities of the stakeholders and enable comprehensive, sovereign and trustworthy data sharing in the short and long term. Therefore, we consider the innovation potential as high. The FDT currently focuses on harvester production data as a first application scenario. Additional use cases are being discussed with stakeholders from the forest-based sector. The presentation is intended to provide an insight into FDT's conceptual approach to discuss design decisions and implementation aspects.

In the forest-based sector, various types of data are generated. Large data sources include for example (1) machine data generated by forest machines, (2) environmental data collected by sensors and (3) inventory data of forest stands. Most of this data is collected but never shared. The resulting untapped potential is huge [1]. To address this issue and handle data sharing, not only in the forest-based sector, several initiatives and regulations have emerged in recent years. Some regulations restrict, e.g., the General Data Protection Regulation (GDPR) [2], and others require, e.g., the Regulation on Deforestation-free products (EUDR) [3], the sharing of data. To increase data sharing and handle the resulting challenges, two concepts are currently popular in the literature: (1) data spaces and (2) data trustees [1], [4].

A data space is defined as a "distributed system defined by a governance framework that enables secure and trustworthy data transactions between participants while supporting trust and data sovereignty" (Chap. 2 in [5]). Two examples from the forest-based sector, which are currently under development, are Dataspace Forestry 4.0 [6] and Forest Data Space [7]. Data trustees can 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 [...]" (translated from [8], p. 7). Their aim is to reduce the users' costs associated with both searching and sharing data. Additionally, there is potential to create trust by monitoring data sharing and removing sensitive information to prevent its dissemination, while ensuring compliance with relevant laws. Therefore, they should handle data neutrally, securely and transparently [4].

The Forestry Data Trustee (FDT) focuses on the needs of the forest-based sector. The aim is to help users to overcome the uncertainties of data sharing on a legal (data protection, contractual arrangements etc.) and technical (compatibility, infrastructure, transparency, encryption etc.) level. We have identified forest machine data, especially harvester production data (HPR) from the StanForD2010 standard [9], as an ideal starting point, as various stakeholders can benefit from increased sharing of this data. For example, processes and supply chains can be monitored and optimized and new business models can be developed [1]. The basic idea is

shown in Figure 1. The forest owner orders a felling measurement by sending a job and usage rights (that define what can be done with data related to himself) to the contractor. The contractor forwards the job to a harvester's digital twin[1], creates an associated data package in the FDT and attaches the forest owner's as well as his own usage rights to the data package. While felling trees, the harvester generates HPR data and its digital twin sends it to the data package inside the FDT. Based on the contractor's configuration, the FDT combines the HPR data from the data package and the usage rights and processes them to create individual data offers for different stakeholders. After agreeing on the data offers, especially accepting the usage rights, consumers can use the data in different formats and with different content as defined by the contractor. The forest owner can use a generated PDF report that shows the status of the felling job. A research institute can use an anonymized HPR data set for research purposes. A digital twin of the stem sections is generated and sent to a sawmill that is a participant of Dataspace Forestry 4.0. For this purpose, FDT offers four different entry points: (1) a web-based GUI, (2) a REST API, (3) digital twins of forest machines, e.g., a harvester, and (4) a data space connector. For more information on opportunities and challenges for a data trustee in the forest-based sector, please refer to [1].

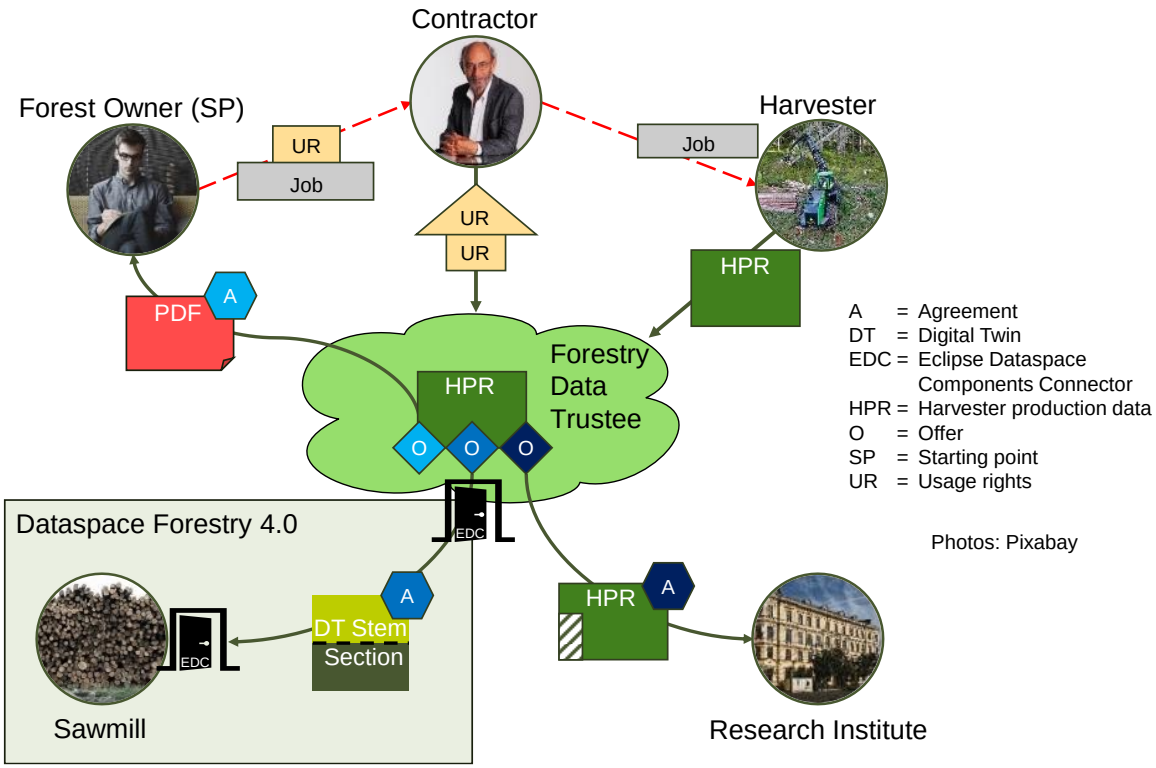

Figure 1: Basic idea of the Forestry Data Trustee

Technically, the FDT can be separated into two parts, the core system and the data space extension. The core system focuses on all aspects that are relevant to ensure that data can be shared, e.g., data processing, usage rights, data storage, or authorization. The data space extension includes all relevant components to connect to Dataspace Forestry 4.0, e.g., an

[1] A digital twin is a virtual representation of a physical component.

Eclipse Dataspace Components-based connector. The aim is to pilot this connection and create a blueprint for using a data trustee as a simple entry point to a data space. The core system will offer a short term and easy-to-use solution for increasing data sharing in the forest-based sector, while the data space connection will provide a long term solution for participants with limited IT capabilities to join data spaces which are innovative but complex technologies.

The main parts of this work are conducted in the context of the research project DTMForst that is supported by the German Federal Ministry for Education and Research (BMBF) under grant 16DTM308 A-G 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.

References

[1] L. Schinke *et al.*, “Trustful Data Sharing in the Forest-based Sector - Opportunities and Challenges for a Data Trustee,” *VLDBW 2023 : workshops at VLDB 2023 : joint proceedings of workshops at the 49th International Conference on Very Large Data Bases (VLDB 2023)*, p. 14, 2023, doi: 10.18154/RWTH-2023-08214.

[2] Proton AG, “General Data Protection Regulation (GDPR) Compliance Guidelines,” GDPR.eu. Accessed: Dec. 11, 2024. [Online]. Available: https://gdpr.eu/

[3] European Commission, “Regulation on Deforestation-free products.” Accessed: Dec. 11, 2024. [Online]. Available: https://environment.ec.europa.eu/topics/forests/deforestation/regulation-deforestation-free-products_en

[4] L. Schinke and J. Roßmann, “Enabling Trustful Data Sharing in Industry 4.0 - A Comparison between Data Spaces, Data Markets and Other Related Concepts,” Sep. 16, 2024. doi: 10.36227/techrxiv.172651024.46145536/v1.

[5] Data Spaces Support Centre (DSSC), “DSSC Glossary | Version 2.0 | September 2023 - Glossary - Data Spaces Support Centre.” Accessed: Jul. 24, 2024. [Online]. Available: https://dssc.eu/space/Glossary/176553985/DSSC+Glossary+%7C+Version+2.0+%7C+September+2023

[7] M. Hoppen *et al.*, “Digital technologies for precise carbon balancing in timber procurement,” *EJFR,* revision submitted after review.

[7] wetransform GmbH, “Home,” Forest Data Space. Accessed: Nov. 21, 2024. [Online]. Available: https://forestdataspace.com/

[8] A. Reiberg, D. Appelt, A. Smoleń, and P. Kraemer, “Datentreuhänder, Datenvermittlungsdienste und Gaia-X,” Gaia-X Hub Deutschland, Whitepaper 2/2023 Version 2.0, Jun. 2023. Accessed: Mar. 07, 2024. [Online]. Available: https://gaia-x-hub.de/wp-content/uploads/2024/02/WP-GX-Datentreuhaender.pdf

[9] Skogforst, “StanForD,” StanForD. Accessed: Sep. 05, 2024. [Online]. Available: https://www.skogforsk.se:443/english/projects/stanford/