An Approach to Develop a Forestry Data Space Using Geoinformation and Remote Sensing Data for Forest Management and Protection

NINA KRUEGER[1] & STEPHAN PLABST[1]

Abstract: The digital transformation strongly affects the field of forest management and protection and provides a lot of potential for improvement in many decision-making processes. However, in the actual forestry ecosystem, despite of numerous solutions, often most of them are not compatible. In this study, we outline a preliminary forestry data space concept, which resulted mainly from the FutureForest project "KI-Einsatz bei Waldzustandsanalyse, der Bewertung zukünftiger Waldentwicklung sowie Entscheidungsvorbereitung zum klimaangepassten Waldumbae" (FF.ai) and describe the technical components. The implemented prototype would try to address the main existing obstacles such as data restriction, complex access and data gaps to extend acceptance of digital technologies in forestry and remote sensing community. Furthermore, a forestry data space as part of an extended digital ecosystem with example scenarios will help to advance digitalization in forestry.

1 Introduction

In our surrounding environment, transformation in digital form is gaining speed and data plays an important role in this transformation, in terms of functionality and sharing. Moreover, the amount of acquired data with high value is rapidly growing. However, make an optimal use of this huge available data is still lacking due to the restrictions and complicated access rules, limited or nonexistence trust between data providers and data users, and data gaps (CURRY 2016). In this regard, to address the mentioned issues, new developments in the area of data infrastructures and their future is happening under the name so-called data spaces in Germany and worldwide. Initiatives such as International Data Space (IDS) and GAIA-X community are leading the design principles of a data space model under important criterions such as data sovereignty, governance, interoperability, portability and trust between participants. Data Space (DS) stands as a decentralised infrastructure for trustworthy, qualified, transparent data sharing and exchange based on commonly agreed principles and rules. It enables many parties to access, use and share their data with each other efficiently and securely. In contrast, to the current data infrastructures with mostly uncontrolled access/use of data, a DS attempts to guarantee confidentiality and security (FRANKLIN et al. 2005). In case of removing restriction and making data open, however, the owners of the data are often exposed to the danger of losing control of their data and thus the strategic value of their data resources without gaining anything as exchange. In consequence, the sharing of data is still only taking place in a limited extent and the potential of DSs remains not explored.

[1] M.O.S.S. Computer Grafik Systeme GmbH, Hohenbrunner Weg 13, D-82024 Taufkirchen,
E-Mail: nkrueger@moss.de

 DOI: 10.24407/KXP:1841077496

Nowadays, DSs developments are focusing on the implementation of the European data strategy in different sectors. While the first use cases are already available in industry, a deeper understanding about the design and properties of a sector specific DS such as forestry with its own specific remote sensing data requirements and regulations is necessary. In this regard, a DS for forest management purposes is aiming at creating and maintaining a single environment for forest relevant data and processes to shared and exchanged across different participants efficiently and securely. This is a challenging goal due to missing reference data and lacking or varying quality of the input data. Therefore, it is necessary to know and understand the concept and architecture model of a DS to adapt the requirements of it to the key aspects of the users and applications in geoinformation and remote sensing community for forest management as well.

The FutureForest[2] (FF.ai) as an ongoing project (until 2025), aims to develop a DS concept in the field of forestry using remote sensing data and Artificial Intelligence (AI) techniques in order to support the climate-adapted decision making for forest management and planning in Germany. The project consortium members are FU Berlin (FUB), Chair of Ecosystem Dynamics and Forest Management at the Technical University of Munich (TUM), M.O.S.S. Computer Grafik Systeme GmbH (MOSS) and wetransform GmbH (WE). Furthermore, the conceptual prototype DS based on forestry-relevant data explores harmonization and simple access regulations including a certification and authorisation stage for data and participants. The available technologies and scientific approaches in the community used as base line and extended to the forestry relevant user scenarios. Based on the conducted experiments during the project, the developed and later implemented prototype will create a road map aiding future forest management by enabling data providers to sharing data while preserving sovereignty on the secure exchange of data in the private and public sector.

2 Problem Statement and Motivation

Forests are changing continuously and playing an essential role in our environment. In the recent years, however, different natural hazards such as storms, fires and bark beetles had great impacts on the actual and future stands of forests. The decision of natural resources managers regarding the future of forest directly connected to the geospatial and remote sensing data in aids of monitoring forests in large scales. Therefore, an accurate, transparent, secure, and shared data ecosystem of forests including biophysical parameters are required to support sustainable forest management. Next to the rapidly rising remote sensing data, other issues regarding a secure shared data environment, which can be accessible and profitable for data provider or data user, are gaining importance. Additionally, it is clear that the remotely sensed data is a dynamic matter in order to support the long-term forest monitoring purposes. Therefore, a DS for forestry management purposes that has access to geoinformation and remote sensing data from various open or private resources, regularly updated, secured, based on the simple rules, providing products and services to the participants is necessary.

[2] https://future-forest.eu/

3 Concept

After a short introduction to DS initials and the goal of FF.ai project, this section describes the main components of a DS known as the soft infrastructure (3.1), connectors and their role as a gateway in this project (3.2) and digital identity (3.3).

3.1 Components, Information Model and Governance

In this section, the structure and main components of a DS individually explained (Fig. 1). Additionally, a detailed description provided to understand the phenomenon of data sovereignty and the architecture as a federated approach for the sovereign exchange of data (OTTO 2022; BADER et al. 2020; JARKE et al. 2014). In following, the connectors and their roles in DS according to the current IDS reference architecture and finally an example scenario, which specially defined for forestry application, described.

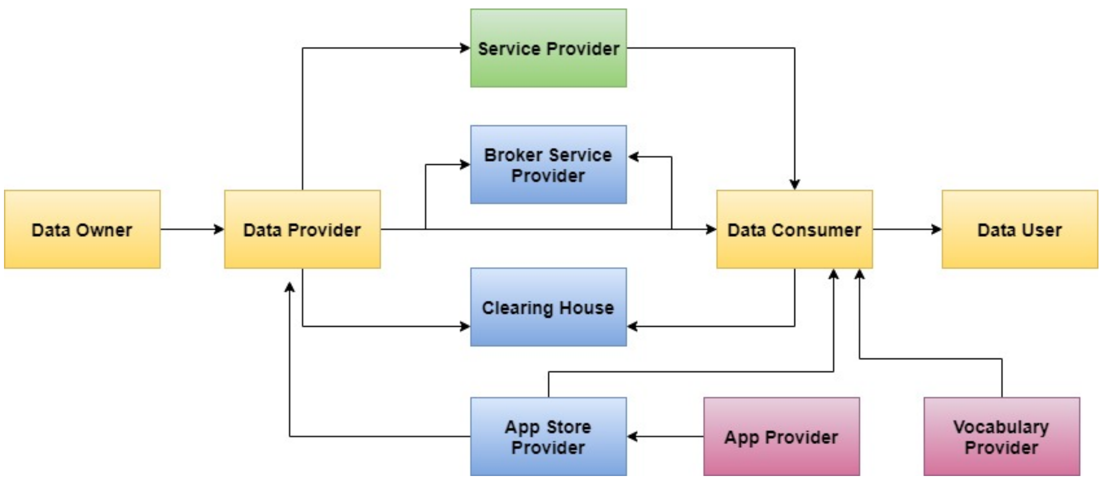

Fig. 1: The soft infrastructure of a DS and the main components

3.1.1 Data Sovereignty

The key factor data sovereignty stands as a precondition for trustworthy data exchange and supports aspects such as identity, trustworthiness of participants. Moreover, it controls data access and usage as well.

3.1.2 Data Governance

In DS fundament, governance plays an important role as a rule-based mechanism (terms and conditions), which is responsible to put data visibility, transparency, and sovereignty in order. As an example, a data can be open for all participants or limited to a specific group based on the allowance of provider.

3.1.3 Data Owner/Provider

The data owner stands as a party who defines and/or exercises control over the data by defining policies and allowing access to data. It should have the technical capability and the responsibility in case of usage contracts and policies. In the other hand, the data provider is responsible for providing the data for exchange between an owner and a consumer including software components that are compliant with the IDS Reference Architecture Model for this purpose. The data provider and data owner not necessarily, but in most scenarios are identical.

3.1.4 Data Consumer/User

It receives data from a data provider and before initializing a request for data, searches for existing datasets by making an inquiry at a Broker Service Provider (see 3.1.5). The data user is the entity that has the legal right to use the data of a data owner as specified by the usage policy. In most cases, the data user is identical with the data consumer.

3.1.5 Broker Service Provider

This component acts as a mediator between data providers and users requesting for data. A Broker provides functions, for data owners to publish their data sources, data users to search through the data sources and to make agreements on the control and use of specific data.

3.1.6 Clearing House

This component keeps/logs all activities performed in the course of a data exchange. After a data exchange completed, or parts of it, both the data provider and the consumer confirm the data transfer by logging the details of the transaction at the Clearing House. Next to keeping record of exchange transactions, the clearing house reports on the search for data sources and on exchange transactions as well.

3.1.7 Service Provider/APP Provider

The IDS promotes the development of a business ecosystem in which participants may develop software (especially services) and make this software available via the App Store. Apps development direction should be compliant with the system architecture of the IDS to be used in the IDS relevant use cases including certifications (MENZ et al. 2019) in order to increase trust in these applications. In this project, a web GIS service further extends based on the WEGA© software technology from MOSS (M.O.S.S. 2019) and a decision support app develops by WE to support the forest management tasks.

3.1.8 Vocabulary Provider

In order to better define how one party owns data, specific vocabularies can be defined and made available for all the DS participants. This component of DS manages and offers vocabularies (i.e., ontologies, reference data models, or metadata elements) in order to annotate and describe datasets.

3.2 Data Space Connectors

The Connector ensures that participants maintain sovereignty over the data and at the same time, functions as an interface between the internal systems of the DS participants and the DS itself. Depending on the configuration, the Connector hosts a variety of system services ensuring, for example, a secure communication, enforcement of usage policies upon exchanged content, monitoring, and logging of content transactions for clearing purposes. Here, a detailed evaluation of available connectors and their performance is out of scope.

3.3 Digital Identity

Certification is an important element in the DS to establish a trustworthy communication. The stage of certification, authorisation based on user and role type in DS defined as digital identity to provide controlled and restricted access via rules and pre-defined regulations. The Identity Provider shall offer a service to create, maintain, manage, monitor, and validate identity information of and for participants in the DS. This is to avoid unauthorized access to data. Further investigations on security and identity management is out of focus of this paper.

4 Application Scenario

In forestry, there are variety of potentials for improvement in case of planning and decision making with respect to the climate adaptation. To benefit from an optimal operational management system supported by software services, all data acquired for decision-making processes must be existed in digital form (LENZERINI 2002). However, for digitalization, the complex models, with lots of parameters and a high amount of required data is challenging. Moreover, there are many different, partly isolated services and systems with redundant or missing data, processes, and software solutions, which cannot be connected with each other due to the lack of interoperability of data, different interfaces and regulations. Furthermore, specific challenges of the forestry such as offline capabilities, use of data locally, or lack of minimum IT infrastructure in private sector, have to be considered for any further integration.

In the first step, a conceptual forestry DS should provide an infrastructure for efficiently supporting forestry interactions and work processes with services and information, provide boundaries and specifications for flexible adaptation of needs from different participants (UL HASSAN et al. 2020). The FF.ai project team combines the expertise of science and research with the interests of the private sector to develop a concept for such a DS while placing challenges like data availability, sovereignty and interoperability in the focus. The ongoing development on a DS prototype in FF.ai project using IDS reference architecture model and connectors at the current state for use cases presented in Fig.2. Each digital DS initially formed by two or more participants. In this project, these participants are foresters, GIS consultants, and research organizations. Three actors involved in a scenario that focuses on use of AI techniques to provide information that can support climate-adapted forest management are:

- The foresters as the data owners,
- The research organisations as the data users and at the same time providers,
- Third-party system or service providers (GIS services, DSS App etc.).

In the proposed scenario, three different participant groups defined to involve and have access via connectors to the DS, which either provide or collect data, use or evaluate or process data for AI based analysis and provide services/apps. Here, the DS approach can offer supporting services via connectors and at the same time protect the forester’s data, as not all data necessarily needs to be free for everyone. Furthermore, data sovereignty plays an important role while besides open data; most of the data in forestry acquired from different governmental sources and private owners. In order to optimize the access cycle, the relevant data must be complete and shared in agreed quality by participants. Thus, the participant as forester can get information about his/her current forest status and thus identify possible problem areas and ask for consultant. Alternatively, a GIS consultant who wants to evaluate a catastrophic situation in forest and provide the owner solutions can use the advantage of connecting to the DS. The connector as a gate provides a secure and authorised access to the DS in order to meet the requirements such as collect the necessary data to do the GIS analysis and the use the results of a trained AI model. As another application example for the DS and corresponding services, is the evaluation of ecological and economic sustainability via the climate change adaption. In forestry, a balanced and appropriate climate adapted cycle forms the core of the efficient, productive, and sustainable planning. In our project, a participant from scientific organisation as an AI expert uses the DS potential to access remote sensing data such as Sentinel 2, aerial imagery etc. for training and optimization of an AI process in order to provide tree species map. The aim, which realistically achievable is by linking different forest areas, remote sensing data and sensor systems records together.

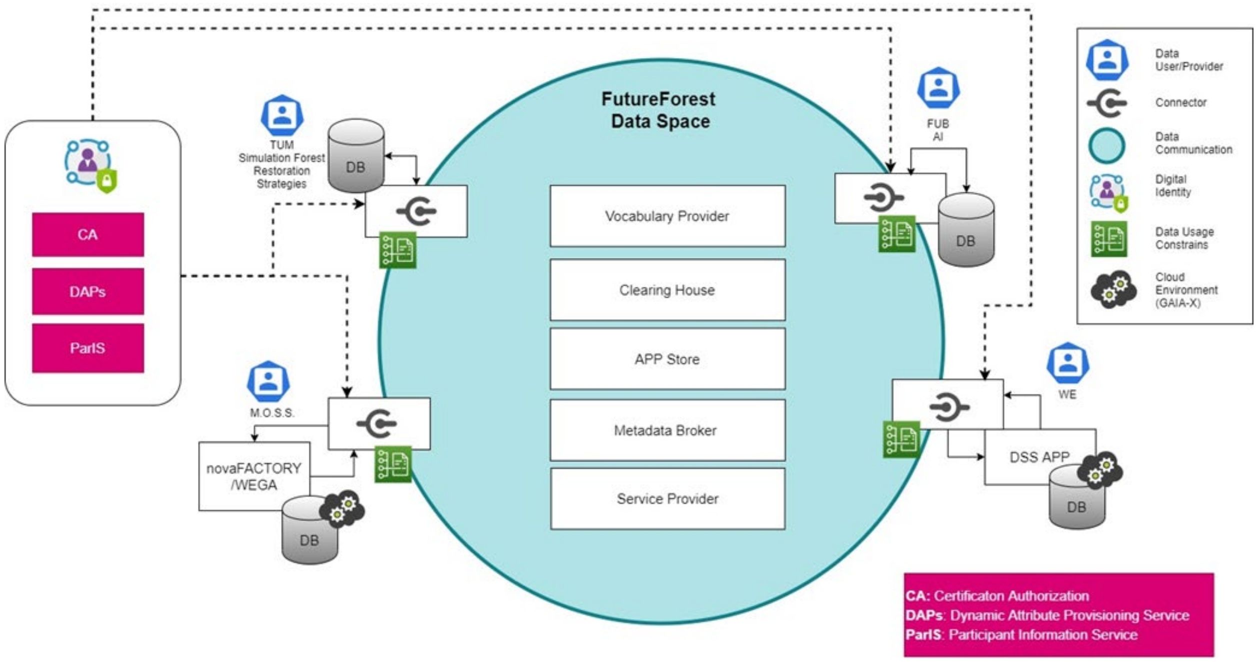

Fig. 2: The developed DS concept in FF.ai project using IDS reference architecture model (for detailed explanation of components see section 3).

Finally, the DS services offer the participants for decision-making aids thorough an integrated web GIS platform based on the software technology of *novaFACTORY*© and *WEGA*© by MOSS (M.O.S.S. 2019) and a decision support app in development phase by WE.

5 Summary and Outlook

In this study, we have motivated and presented the concept of a prototypic forestry DS, which can greatly advance digitalization in forestry. To do so, the proposed concept, takes the concepts of the IDS standardisations, adapts, and extends them with solutions for the forestry applications. However, there are still uncertainties especially in case of certification and authorisations possibilities with respect to the remote sensing data, which requires further investigations. Therefore, the association and participation from remote sensing and forestry community during this project is very important in order to develop the appropriate concept.

6 Acknowledgement

The three-year research project is funded by the German Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection (BMUV) under the funding program " KI-Leuchttürme für Umwelt, Klima, Natur und Ressourcen " (funding code: 67KI21002A). The BMUV is supported by Zukunft - Umwelt - Gesellschaft (ZUG) GmbH as project executing organization.

7 References

BADER, S. R., PULLMANN, J., MADER, C., TRAMP, S., QUIX, C., MÜLLER, A. W., AKYÜREK, H., BÖCKMANN, M., IMBUSCH, B. T., LIPP, J., GEISLER, S. & LANGE, C., 2020: The international data spaces information model - an ontology for sovereign exchange of digital content. Int. Proceedings of International Semantic Web Conference (ISWC), Athens, Greece, 176-192, https://doi.org/10.1007/978-3-030-62466-8_12.

CURRY, E., 2016: The big data value chain: Definitions, concepts, and theoretical approaches. In J. M. Cavanillas, E. Curry, & W. Wahlster (Eds.), New horizons for a data-driven economy, Springer International Publishing, 29-37, https://doi.org/10.1007/978-3-319-21569-3_3.

EUROPEAN COMMISSION, 2021: Results of the new European Data Market study 2021-2023 online at https://ww.digital-strategy.ec.europa.eu/en/library/results-new-european-data-market-study-2021-2023, last access 11.09.2022.

FRANKLIN, M., HALEVY, A. & MAIER, D., 2005: From databases to dataspaces. ACM SIGMOD Record, **34**(4), 27-33, https://doi.org/10.1145/1107499.1107502.

GAIA-X, 2021: IDSA Reference Architecture Model, Version 4.0. https://internationaldataspaces.org/ids-ram-4-0/, last access 10.12.2022, https://www.bmwi.de/Redaktion/EN/Dossier/gaia-x.html, last access 20.08.2022.

JARKE, M., JEUSFELD, M. & QUIX, C., 2014: Data-centric intelligent information integration -from concepts to automation. Journal of Intelligent Information Systems, **43**(3), 437-462.

LENZERINI, M., 2002: Data integration: A theoretical perspective. In Proceedings ACM-PODS, 233-246.

MENZ, N., RESETKO, A. & WINKEL, J., 2019: IDS certification explained. https://internationaldataspaces.org/download/16456/, last access 13.01.2023.

M.O.S.S., 2019: Umsetzung eines Digitalen Lagers für Geodaten mit novaFACTORY. M.O.S.S. Computer Grafik Systeme GmbH, https://t1p.de/nFdigitaleslager20, last access 20.01.2023.

NAGEL, L., 2021: Design principles for data spaces. https://design-principles-for-data-spaces.org/, last access 11.09.2022.

OTTO, B., 2022: A Federated Infrastructure for European Data Spaces. Communications of the ACM, **65**(4), 44-45, https://doi.org/10.1080/10580530.2015.1044344.

UL HASSAN, U., OJO, A. & CURRY, E., 2020: Catalog and entity management service for internet of things-based smart environments. In Real-time Linked Dataspaces, Springer, 89-103, https://doi.org/10.1007/978-3-030-29665-0_6.