

AGRICULTURAL DATA SPACE (ADS)

A PUBLICATION OF THE FRAUNHOFER LIGHTHOUSE PROJECT "COGNITIVE AGRICULTURE"

WHITEPAPER





AGRICULTURAL DATA SPACE (ADS)

WHITEPAPER

Digital Transformation in Agriculture

At the technical level, the digital transformation in agriculture requires digitally available data from the environment, from farms, machines, and processes to enable software-supported products and services to work smoothly.

In this whitepaper, we motivate the need for and show the added values of an "Agricultural Data Space" (ADS for short). We outline an ADS concept and describe the necessary prerequisites and technical solution approaches. Complemented by the possibilities of a transparent and open marketplace for data, digital products, and software services, such a data space would address many of the existing obstacles to widespread acceptance and take-up of digital technologies. Overall, an ADS as part of an extended digital ecosystem will significantly advance digitalization in agriculture. The foundations required for this are being developed in the Fraunhofer lighthouse project "Cognitive Agriculture" (see info box).

Fraunhofer Lighthouse Project Cognitive Agriculture (COGNAC)

In the Fraunhofer lighthouse project COGNAC, eight Fraunhofer Institutes are researching and developing innovations for the digital transformation in agriculture. Our vision is to create a living digital ecosystem for the agriculture of the future, the Agricultural Data Space (ADS). To this end, advanced sensor solutions, field robotics applications, and cognitive data services will be designed, piloted, and evaluated in this project. A central digital platform will link these solutions in the context of the ADS.

A cross-system agricultural data space offers direct added value for farmers, who will get support for their operational decisions on the basis of high-resolution measurement data from airborne or ground-based systems and their refinement by means of cognitive services. At the same time, autonomous field robotics will benefit from the connection to the ADS, as it will offer services for customized control of plant-specific field work on the farmer's fields.



Automation and optimization by means of digitalization and software

Nowadays, the optimization potential of individual agricultural machines is severely limited. Improvement potential mainly exists in the work processes and the higher-level planning and decision processes. This requires comprehensive contextual information from the past, the present, and the predicted future, which requires automation (i.e., through software) in order to be recorded and processed in the necessary quantity.

To enable optimal operational management supported by software services, all data required for decision-making should therefore be available in digital form.

This is why investments in digitalization are already being made today, especially in large companies. Detailed monitoring of the condition of agriculturally cultivated areas (e.g., plant growth, plant health), farm animals (e.g., movement profiles, body temperature), and environmental parameters (e.g., amount of precipitation, ambient temperature, amount of feed in the troughs) enables the farmer, for example, to detect problems quickly and react to them. Even the application of measures can already be automated today. In addition to the detailed monitoring of fields and farm animals, complete work steps are already carried out in an automated manner in some cases, for instance by means of precision farming with subplot-specific fertilization or feed troughs with cow-specific feed mix and quantity. This development increasingly leads to a reduction of the farmer's workload while simultaneously increasing efficiency and minimizing risks. At the same time, it supports the trend away from area-focused performance and towards individual optimization in line with a farm's own strategy.

However, many isolated solutions have emerged as a result of individual products from different manufacturers. Their interoperability and integration pose great challenges for farmers. Changing suppliers is also a difficult task, so in light of limited benefits, unclear implications of manufacturer loyalty, and rapid technological change, there is great uncertainty and reluctance regarding investments.

Currently, agricultural data and service platforms that bundle several products and services are gaining in importance. Software-based services can provide farmers with helpful information for decision-making in challenging situations. Services on the platform process the farmer's own data as well as data from many other farmers with comparable conditions into a decision-making aid.

It must be easy for the farmers to network sensors, actuators, and the platform in order to optimize digitalization for their own farms and enable them to participate in creating added value with the data. As soon as a farmer provides their own data for evaluations in digital services, they currently lose sovereignty over this data. The platform providers are called upon to create solutions with which farmers can control and monitor data sovereignty easily and in a self-determined manner. Interoperability for universal data usage should not end with the farmer, however; rather, it should be enabled along the entire value chain from processing operations to the consumer. This is the only way that all stakeholders can benefit from data analyses and decision support based on them, and the only way in which comprehensive transparency can be achieved.

AGRICULTURAL DATA SPACE (ADS)

Who should get access to the data?

In the context of access to data, the terms “data sovereignty” and “data ownership” are often used. We define **data sovereignty** in agriculture as the ability to make self-determined decisions about the use of data collected in the context of one’s own farm, to understand the impact of these decisions, and to change these decisions at any time. According to our understanding, **data ownership** comprises the legal legitimation and technical ability to be allowed to and able to authorize data processing in an operational context and revoke it at any time. Data processing here also includes the collection, storage, transmission, and deletion of data.

In addition, data access can also be enabled, of course, through individual contracts. This is essential for automation through

software-based services and for digitalization, and must therefore be easy and flexible.

Data categories regarding data privacy

Currently, there is no legal concept of “data ownership”. According to the law, data privacy only refers to personal data, respectively data that can be associated with a person. This protection is stipulated in the EU’s General Data Protection Regulation (GDPR), respectively in Germany’s new Federal Data Protection Act (BDSG-new). The sensitivity of other data and the resulting need for data privacy are strongly dependent on the subjective value and can be subdivided into several categories for an agricultural enterprise, examples of which are listed in the table below:

Data Sensitivity Class	Example	Access
Person-related	Working hours, sick days	Individual, farmer
Enterprise-related	Customer relations, price offers	Internal
Resource-related and enterprise-relevant	Fertilizer consumption, fuel consumption, costs	Internal, contractor
Resource-related and product-relevant	Tractor engine data, feed composition in mixing wagon	Tractor manufacturer, equipment manufacturer
Product-related	Protein content, humidity	Customer in supply chain
Process-related	Field borders, driving tracks, yield map	Subcontractors, consultants, other contractors
Public-interest data, non-competitive data	Diseases, invasive plants, green areas (legal duty to report, public interest, e.g., environmental protection, safety relevance)	Government bodies, public authorities
Open data	Maps, aerial images, ground reliefs, (data collected in the public interest)	Everyone



Need for a Cross-Platform Agricultural Data Space

A free and open competition of data-based services requires an infrastructure that supports the digital economy accordingly. With regard to the data, some aspects that we consider essential are listed below:

Enable added value creation from data

We live in an age in which data volumes are growing rapidly and the extraction of information from data is becoming increasingly relevant. The reasons for this include increasing networking, digitalization, and digital transformation in a wide variety of domains, including smart farming. In this area, data can be aggregated and refined into helpful information and recommendations for farmers and other stakeholders in the agricultural sector. For example, farmers can be offered recommendations for fertilization and harvesting as a service by aggregating and analyzing weather data, soil moisture and nutrient data, as well as data about plant growth.

For companies offering such services, such data is “the new gold” or “the new oil” of the 21st century. They must therefore use and share comprehensive data for their data-centric business models, which requires easy and fast access to high-quality data. This data, in turn, comes from different data sources of a variety of customers or business partners. Such a customer may be a farmer who has collected this data on his farm. The willingness to share this data for the purpose of gaining information does exist, provided there is a concrete added value for the farmer sharing the data. One of the aims of the ADS is therefore to provide a central marketplace for data as part of a digital agricultural ecosystem. In this marketplace, suppliers of data should find incentives for sharing their data. Cognitive services can then refine this data and thus offer added value to data consumers.

Data as a tradeable asset while maintaining data sovereignty?

A digital ecosystem should always allow sovereign handling of data. Data-driven companies are partly motivated intrinsically, partly legally obligated not to collect, process, and pass on data in an uncontrolled manner. In the legal sense, there is no general ownership of data. That is, whoever possesses data may process it. Exceptions are, for example, personal data, for which companies must comply with the legal requirements of the new Federal Data Protection Act and the underlying General Data Protection Regulation, or data containing artistic works protected by copyright. Here, the affected persons usually control the processing of their own data. Most other data is not subject to any legal coverage and must be protected by the producer or owner from undesired processing. Data suppliers must therefore be empowered to control their own shared data. In order to achieve and maintain data sovereignty while enabling data trading at the same time, organizational or technical measures are required. In addition to contractual agreements, which are often time-consuming to draw up and inflexible, technical concepts such as data usage control and smart contracts can make data sovereignty possible for data owners and enable flexible data usage.

Data quality despite Big Data

The possibilities available today to analyze large amounts of data in an automated manner within a short period of time and to resolve even complex interrelationships represent a major usage driver on the road towards smart farming solutions. However, an issue that is often neglected are the high requirements regarding data quality. Particularly with regard to the use of supervised Machine Learning methods (a subarea of AI methods), data must be labeled descriptively in order to be suitable for the training and validation of decision models such as neural networks. In an ADS, data must therefore be characterized, respectively get qualified, with regard to its quality in

AGRICULTURAL DATA SPACE (ADS)

WHITEPAPER

order to be used for cognitive services. The quality of the data (e.g., missing values, outliers) also has a great influence on the conclusions that can and may be drawn from the derived decision models. Not least for reasons of acceptance, the quality of the data collected in the agricultural context must therefore be taken into account when interpreting the decision models used by humans or machines. The aim of the ADS is to provide support for the implementation of corresponding standards and to define quality standards. Furthermore, an appropriate data platform can partly also complement characteristic quality data such as authenticity, collection context, or completeness regarding a process context on its own. Such data quality properties are also important for determining the value of the data in a data marketplace.

No end-to-end data interoperability

In the past thirty years, countless IT solutions have been developed in the agricultural sector that independently support farmers and agricultural enterprises in their tasks. Only in a few areas do standards exist that enable easy and smooth communication across interface boundaries. This lack of interoperability is one of the main obstacles to digital transformation in agriculture. However, it is an illusion to believe that all subject-specific elements can be covered by standards. In an agricultural data space, we must create the possibility to translate between different representations as automatically as possible and to interpret these. This requires machine-readable interface specifications that enable digital services to – if necessary – interpret these at runtime and integrate them dynamically. The solution approach of flexibly expandable ontologies enables innovation and evolution without imposing restrictions. Although such an approach may initially appear cumbersome and an obstacle to performance, it is precisely here that platform approaches can leverage their scaling effects and offer corresponding data consistency through translation services.

Digital Ecosystems in Agriculture

Digitalization in almost all industries and domains no longer affects only existing business processes; rather, complex digital ecosystems are emerging with increasing speed and degree of penetration: Processes are first digitalized and then networked; the Internet of Things (IoT) integrates machines and peripherals; customers, service providers, and other stakeholders interact via systems in the Cloud; novel types of business models emerge along digitalized value chains. According to our definition, all stakeholders and technical systems in a domain, such as agriculture, form a **digital domain ecosystem**. In these, individual, independent, and partly overlapping **digital ecosystems** can exist. Examples can be found, for instance, in agriculture, where manufacturers of agricultural machines network their digitally-enabled products or services. In order to network stakeholders and technical systems in such a digital ecosystem and thereby enable the ecosystem in the first place, **digital platforms** are typically used. They form the basic infrastructure for the development of complex digital ecosystems.

The agricultural domain is a high-technology and widely digitalized domain. Machines with digital interfaces and capabilities are state of the art. In recent years, digital ecosystems have increasingly evolved, although they have been developed mostly for the products and services of individual manufacturers and service providers. To date, there has been hardly any exchange of data or functions between the ecosystems themselves. Attempts to make use of so-called data hubs or bilateral connections to establish comprehensive connectivity – so far, with a rather modest range of functionality or only limited coverage of the agricultural domain ecosystem – are quite recent. From today's perspective, it is not yet clear whether digital ecosystems will integrate into a single cross-manufacturer ecosystem, or whether the market will consolidate into a few dominant ecosystems.

AGRICULTURAL DATA SPACE (ADS)

Figure 1 shows an example of the digital ecosystems of two manufacturers. A farm uses digitally-enabled agricultural machines from both manufacturers and the Farm Management Information System (FMIS) of manufacturer B. The machines transmit farm-related data (e.g., amount of pesticides applied and crop yields) into their respective ecosystems. In both systems, digital platforms implement the connection of the machines and the communication infrastructure, the storage of the data collected by the machines, and access to software solutions for the farmer.

The data collected by the machines and stored on the respective platform is now stored in separate places. Whereas the farmer can see and use part of the data using the FMIS of manufacturer B, the data of the machine from manufacturer A cannot be accessed there. As mentioned above, there are initiatives

for connecting both data sets; however, cross-manufacturer, complete integration is not yet foreseeable. Today, platforms of individual manufacturers and individual services of solution providers usually only integrate their own product portfolio.

Providers of standalone solutions also face additional challenges in the context outlined above: In most cases, they depend on data sets or machine connectivity that they do not have themselves. In order to offer their own solution, they would need to develop their own digital ecosystems or build on existing ones. The implementation of their own ecosystems is a challenge that is difficult to master even for large companies, while the integration into existing environments always entails a dilemma: making multiple efforts to cover all digital ecosystems of a domain or opting for one ecosystem and thus risking limited market access.

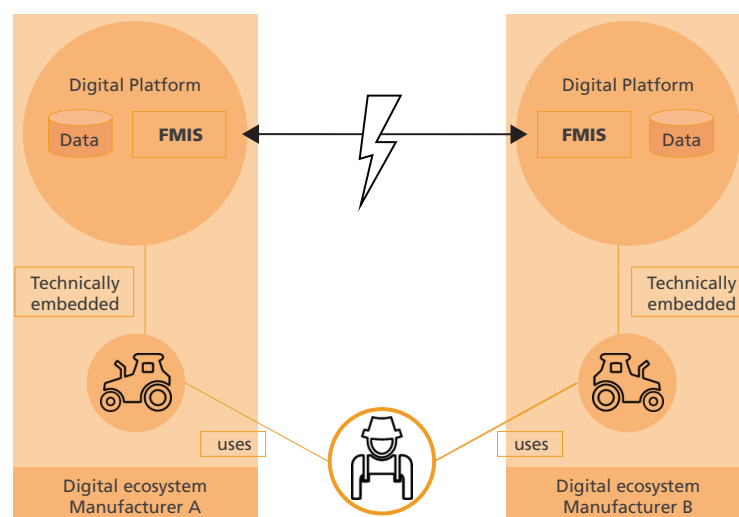


Figure 1 – Participants in a digital ecosystem connected via a digital platform





Concepts of the Agricultural Data Space

The Fraunhofer lighthouse project “COGNAC” addresses the issues and problems of the status quo in the world of digital ecosystems in agriculture. In the project, concepts for solving the above-mentioned challenges and for consolidation in an overall, comprehensive architecture of a digital domain ecosystem for agriculture are being researched and tested. The result characterizes our vision of an Agricultural Data Space (ADS).

Fraunhofer already outlined the basic idea for such data spaces in 2014 in the context of an “Industrial Data Space”¹. The concepts created there are currently being discussed and refined for various application areas in the context of the International Data Spaces Association. The Agricultural Data Space represents an industry-specific adaptation and extension of the concepts that takes into account the specific requirements and the market situation. Following the concepts of the International Data Space (IDS)², we understand the ADS as the totality of all components of a digital ecosystem that generate, store, manage, or consume data. Just as in the digital ecosystems of the manufacturers outlined above, the ADS requires a digital platform as enabler of a thriving digital ecosystem. One of the goals is the best and greatest possible integration of all components of the digital agricultural domain ecosystem. Existing or future offers by manufacturers should not be forced out, however, but should be connected and supported. This is not only about connectivity and interoperability between previously unconnected ecosystems, but also about offering independent third-party solutions and services. The ADS platform is intended to enable these providers to make their products and services available on as many digital agricultural markets as possible on a manufacturer-neutral basis.

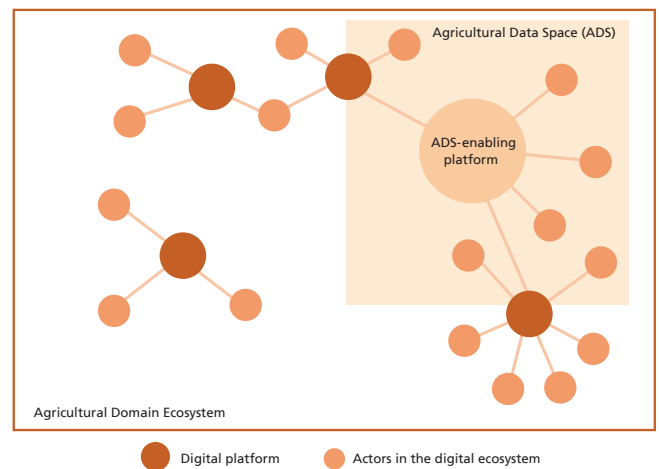


Figure 2 – ADS-enabled platform opens up the data space as an intermediary between ecosystems

Figure 2 outlines our vision of the ADS as a comprehensive whole of all stakeholders in the digital agricultural domain ecosystem with a digital ADS enabler platform as intermediary and integrator of previous individual solutions. In addition, a manufacturer-neutral marketplace for data could be offered on an ADS enabler platform that would allow all stakeholders non-discriminatory access while ensuring data sovereignty at the same time. Figure 3 provides an overview of the major functions and components of the ADS enabler platform, including:

- Central data storage: This is where the agronomic and operational data is stored and managed as digital twins (the concept of digital twins is explained in detail on page 14). These include components for access and usage control as well as cognitive interfaces for smart access.
- Access to the management of access and usage rights: An environment for farmers to individually configure access and usage rights for their data.

1 https://www.fraunhofer.de/content/dam/zw/de/Forschungsfelder/industrial-data-space/Industrial-Data-Space_whitepaper.pdf
 2 <https://www.internationaldataspaces.org/>

AGRICULTURAL DATA SPACE (ADS)

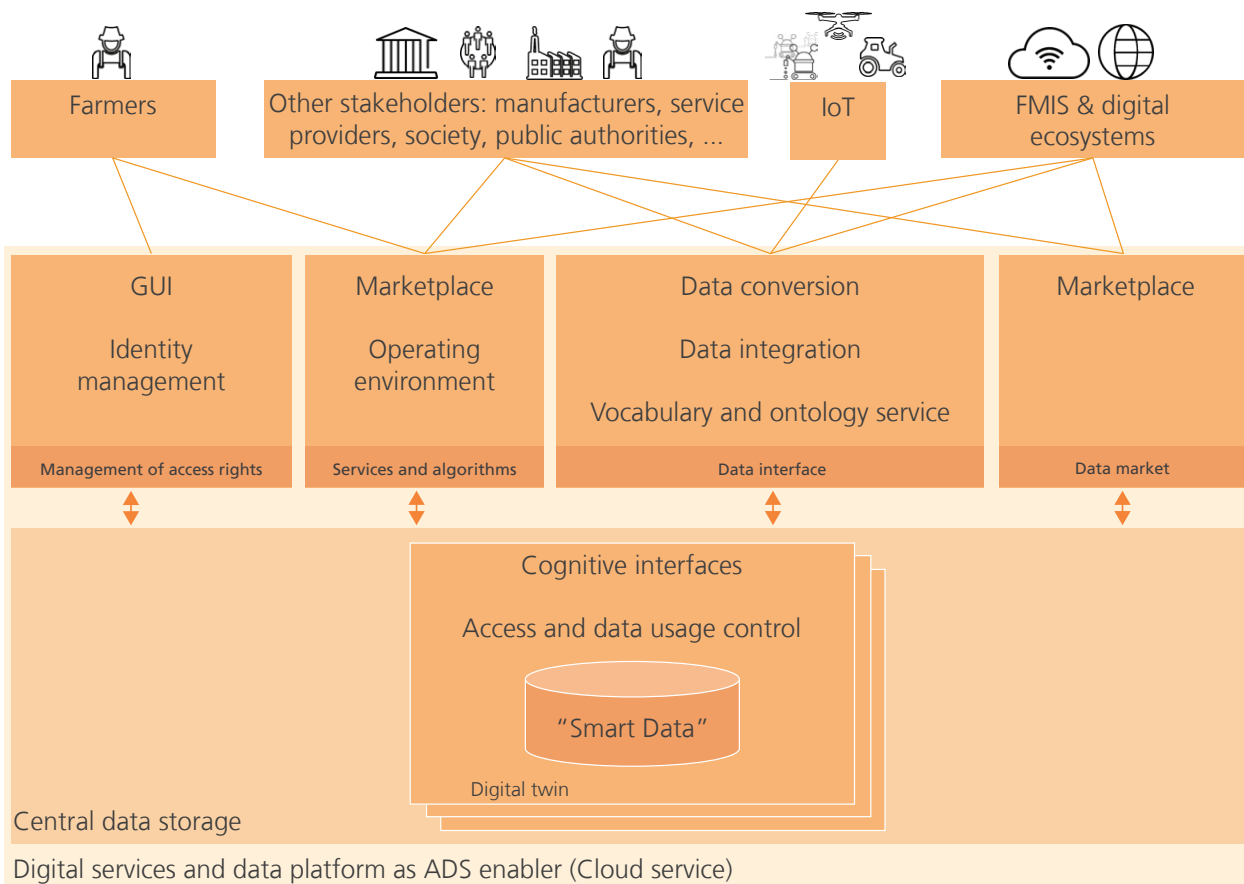


Figure 3 – Overview of the functions of the ADS enabler platform

- Environment for services and algorithms:
An operating environment for digital services, algorithms, or apps that can be used in the context of the digital platform. Via a marketplace, algorithms (e.g., growth models for plants) can be offered individually and can be integrated into services.
- Data interface:
Interface for data access from external systems. It should also be possible to make data available outside the ADS en-

abler platform so that there is no obligation to use software solutions in the context of a specific platform.

- Data marketplace:
Data-driven business models are becoming increasingly important in agriculture. The ADS makes it possible to offer data for trading via a marketplace. It is planned to integrate possibilities for aggregation, pseudonymization, and anonymization.



In the current conceptual design of the ADS and the presentation of the platform functions, the focus is on the farmer. As a rule, it is the farmer's data that is used and traded in the ADS, and this is what we want to focus on in the first iteration of this research project. In the long term, the concept can be extended to any type of stakeholders. Thus, other stakeholders in the agricultural value chain (e.g., food producers) should also be able to participate in the ADS with their respective data as securely as farmers.

ADS connector

Access to the agricultural data space can be mapped via connectors such as those specified by the IDS. The goals are, among others, secure and trustworthy access to data and services and management of identities. For us, the focus is on data usage control, which ensures that data only leaves the area of the ADS if this is authorized. In the context of COGNAC, we must take into account the status quo in digital agriculture as well as the demands of other stakeholders and the state of the art when we develop the conceptualization.

Data usage control for flexible data access

The concept of data usage control does not only include controlling access to data (i.e., access control), but also its usage after it has been passed on to corresponding data-processing services. Usage policies are enforced at interfaces between systems (e.g., between the farmer and the connector to the ADS) and in target systems, so that violations of data sovereignty or privacy are avoided. This also includes who may use which data for which processing, when, for how long, and how often in which context. The policies can be realized through technical implementation of data usage control, for example through anonymization, pseudonymization, partial encryption, data filtering, data masking, data deletion, contextual data usage, or time-limited data usage, either at interfaces or in target systems. In Figure 4, example control possibilities are sketched that data owners can specify prior to passing on their data.

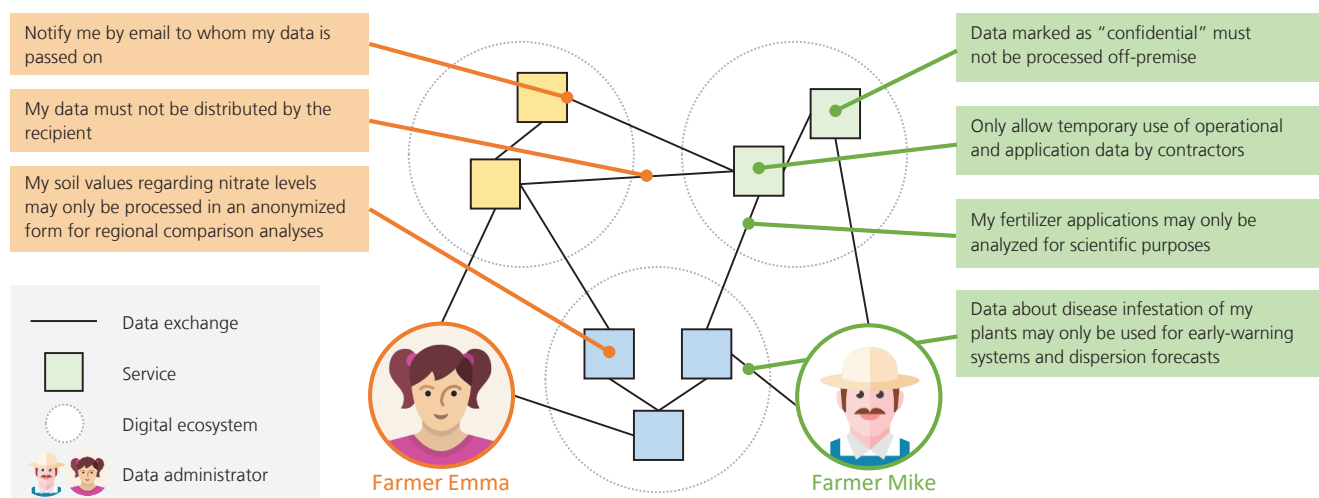


Figure 4 – Example of control possibilities with data usage control

AGRICULTURAL DATA SPACE (ADS)

Marketplace for data

In the goals of the ADS, we demand that it shall promote the trading of data in order to enable comprehensive value creation from the data. Companies with data-centric business models require easy and fast access to high-quality data. Data providers, such as farmers, need a simple and attractive way to make the data available.

The following basic principles apply to an ADS data marketplace:

- Data sovereignty must always be guaranteed for the data owners.
- The marketplace enables fair trading of data with added value for all stakeholders.
- Interoperable interfaces make it easy to provide and use data.

Marketplace for services

By itself, data does not provide any added value. In order to make data usable for the participants of the digital ecosystem for the agricultural domain, end-to-end integration of data and services is a prerequisite for value creation. Only a marketplace for services and service provision in which providers and users come together and where a choice in supply meets demand enables an open ecosystem. For this to work in the virtual world, measures must be taken to make the quality of a service transparent and promote trust between business partners.

But what are these services anyway? In digital ecosystems, the focus is on digital services. In agriculture, this includes services such as the analysis of data sets to get information about the regional yield situation. However, such purely digital services without any reference to a concrete work activity are rather rare even in digitalized agriculture. More often we will encounter

services that involve a service provided downstream. One example of this is the hiring of a contractor via a digital service that optimizes the planning over several harvest campaigns. Another example is the data-based optimization of nutrient cycles via an app that farmers can use to accurately plan the application of nutrients when they are out and about or when they are at the farm. Such services require a large amount of data, often across farms.

The providers of services in digital agriculture can be as diverse as the users. The constellations can be simple; e.g., if a pesticide manufacturer offers an app that enables farmers to reliably detect plants infested with a disease and recommends the appropriate pesticide produced by said manufacturer. On the other hand, complex chains are also conceivable; for example, if a digital service provider offers various possible solutions once a disease has been identified or directly commissions other service providers – similar to a consultant.

We see the optimal linking of services and data as well as service providers and service users as a basic prerequisite for a thriving ecosystem. In addition to technical issues such as compatibility of data and systems, we believe that the simple and flexible use of services is essential for enabling a digital agricultural data space such as the ADS. Our focus is on the farmers, who should be enabled to find and use services easily and reliably. Our approaches comprise the following concepts:

- A marketplace in which the services being sought can be found quickly and where their quality and reliability can be checked reliably. Analogies are the app stores of Google or Apple, for example.
- Non-discriminatory access of services to data – for providers and users alike.



- Easy use of services – Users should not have to deal with background tasks such as connectivity and should be optimally supported in their work by uniform access to services.
- End-to-end data privacy and data usage control for services.

Uniform and flexible data access via digital twins

For uniform and flexible access to data, we propose using the concept of the “digital twin”, which is also used in factory automation in the form of asset administration shell for plants and products in Industrie 4.0. A uniform asset administration shell provides an interface for calling up (data) services of a company’s tangible or economic assets. In agriculture, this may be the digital twin of a dairy cow, a tractor, or a field. All digital twins together thus represent a digital image of a farm, which can be used by software services. It makes sense to perform data usage control already at the level of the digital twin. The standardized interface elements only cover basic elements, such as a directory service regarding possible data services or specific functions of the digital twin. This enables flexible expansion.

Intelligent functions create interoperability

Simple, non-technical interfaces on the digital twins and services create flexibility and compatibility. In order to be compatible at this meta-level, however, software functions must have a certain “intelligence” to understand what a specified function means. This requires vocabularies that describe and characterize elements, as well as ontologies that create relations between corresponding semantic structures. Ontologies have so far been developed and used mainly in research projects, such as the plant-related ontologies listed at cropOntology.org. However, if these semantic directories are managed and quality-assured appropriately, they enable software functions to achieve a certain cognitive understanding, which also learns by adapting and extending the ontologies.

With the help of appropriate vocabularies, related and similar terms as well as translations can be managed. The AGROVOC directory³, which is supported by the Food and Agriculture Organization (FAO), already lists more than 36,000 concepts and covers 33 languages. This approach may appear cumbersome and effort-intensive at first glance, but it ensures independent extensibility and enables proprietary standards to co-exist. Using formalized domain knowledge, cognitive functions can behave in an application-oriented and context-sensitive manner.

The example in Figure 5 shows how an agronomic service “talks” to a digital field twin to start the correct query for data on nitrogen content.

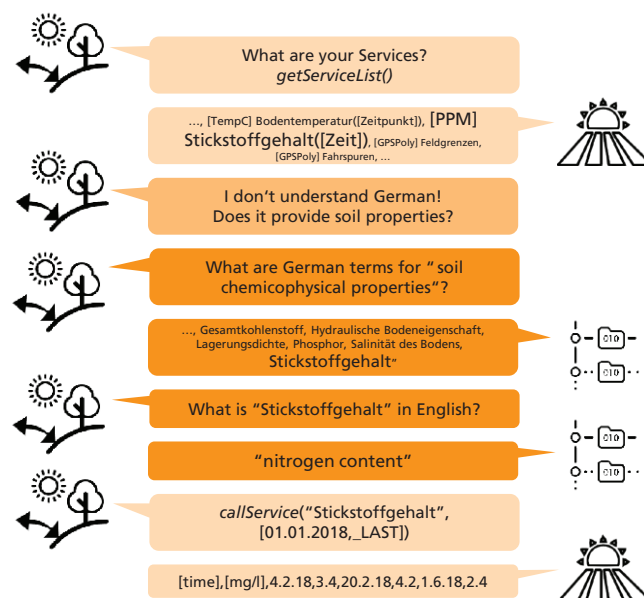


Figure 5 – Example of simplified communication between an agronomic service for nitrogen balance (left) and a digital field twin with the aid of a vocabulary service (right).

3 <http://aims.fao.org/standards/agrovoc>

AGRICULTURAL DATA SPACE (ADS)

WHITEPAPER

Quality assurance for data and services

Marketplaces only work well if the quality of the offered goods can be evaluated (ideally prior to a purchase). This is also true for data and services. There is the possibility to share subjective experiences – an example are the user reviews known from online shops. This may work for apps and software solutions, but it is not feasible for abstract and large amounts of data. Automated services, however, depend on high-quality data. It would only be logical to check the quality of the data centrally in a platform, respectively to define quality standards there. Poor data quality and incompleteness of data are indeed the main reasons why a smart farming application is not successful. To solve such problems, quality-enhancing services can be offered on the platform, such as testing of pre- and post-conditions (sensor calibration), plausibility checks in the concrete agricultural context (soil dry despite rain), value range tests, or basic statistical tests.

Furthermore, whether a service behaves in compliance with the current legislation (for instance, whether the fertilizer recommendation is in conformity with the current legal situation) could also be a verification service offered by a platform. However, this would require that the service can be virtually integrated into a corresponding simulation environment. In the age of digital ecosystems with dynamically changing networked products and services, however, this step becomes a necessity in order to be able to guarantee the safety and reliability of the entire digital ecosystem once a certain level of complexity is reached.

Cognitive services

In many cases, the optimization and automation of work and business processes requires the inclusion of extensive contextual knowledge and a decision based on complex patterns. We call software-based services that claim to perform information processing similar to human cognitive services. This information processing is based solely on the transformation of sensory patterns by a processing system that has been parametrized as much as possible on the basis of sample data. Thus, a system is not built using, e.g., physical rules (goal of cybernetics), but rather generic mathematical constructs whose parameters are adjusted by processing observational data. This area of Artificial Intelligence known as Machine Learning is currently experiencing an enormous development boost, in particular because the progress in digitalization and networking is making machine-readable information available.

In the context of the task definition of the Agricultural Data Space, we consider the following basic central tasks of a platform necessary to support or use cognitive services:

- **Support for the training of decision models (e.g., neural networks)**, e.g., by providing context data to support corresponding labeling/semantic segmentation, or the semantic integration of parameterizable simulation models to supplement training data for neural networks.
- **Support for the validation of decision models (e.g., neural networks)**, incl. versioning of the respective data.
- **Quality assurance of cognitive services**, e.g., through a digital test bed (virtual validation) or simulation games with historical data / public test data.
- **Monitoring of cognitive services**, such as compliance with guidelines and threshold values during application and use in the approved application context.



Basic services could also be offered on data itself, which the platform users could then integrate into their own offers. Some examples are:

- Provision of basic services for Automated Machine Learning (AutoML), monitoring and validation checks of input and output data or sensor-specific data evaluation, e.g., calibration steps or sensor simulation
- Compression and processing of information from various data sources that cannot be interpreted directly by the farmer, through model-based transformation
- Detection of recurring information patterns and derivation of associated proposals for action

The challenge of methods that work on a cognitive basis is that their models are not universally valid, but are always only valid in the context of the data set used for parameterization and represent a type of black box. This fact must be taken into account when setting up a service by ensuring that the implementation of quality standards by the service is obligatory (Sculley et al. 2015⁴, Kläs & Sembach 2019⁵, Jöckel & Kläs 2019⁶).

Overall, the implementation of cognitive services requires a comprehensive understanding of interrelationships and the inclusion of the corresponding information. Analogous to common maturity models in digitalization, Figure 6 shows which level of decision supports requires which prerequisites.

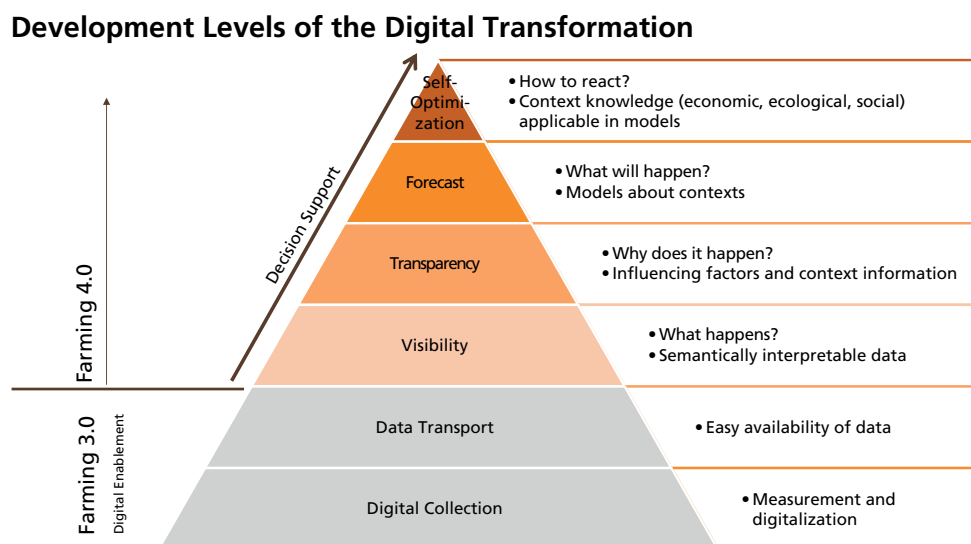


Figure 6 – Maturity model for decision support

4 Sculley, D.; Hold, G.; Golovin, D.; Davydov, E.; Phillips, Todd; Ebner, D.; Chaudhary, V.; Yong, M; Crespo, JF; Dennison, D. (2015): Hidden Technical Debt in Machine Learning Systems, NIPS' 15 Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2, pp. 2503-2511

5 Michael Kläs, Lena Sembach: Uncertainty Wrappers for Data-Driven Models - Increase the Transparency of AI/ML-Based Models Through Enrichment with Dependable Situation-Aware Uncertainty Estimates. SAFECOMP Workshops 2019, pp. 358-364

6 Lisa Jöckel, Michael Kläs: Increasing Trust in Data-Driven Model Validation - A Framework for Probabilistic Augmentation of Images and Meta-data Generation Using Application Scope Characteristics. SAFECOMP 2019, pp. 155-164

AGRICULTURAL DATA SPACE (ADS)

Application example

Agriculture is a diverse sector, and farms are part of a complex network with various interest groups. A large amount of different data is generated on a farm, which in the future will be increasingly collected, stored, evaluated, and documented by many different systems. For this reason, the approaches of the ADS with its possibilities for integrated data access and data usage control are of essential importance. One application example for the ADS and corresponding services is the evaluation of ecological and economic sustainability via the nutrient cycle. In agriculture, a balanced and appropriate nutrient cycle forms the core of the efficient, productive, and sustainable production of plant as well as animal products. In addition to documentation, the focus is increasingly on the optimization of the nutrient cycle. In our example, we consider a dairy farm with crop cultivation and grassland farming. Optimization can only be achieved by linking different specialist areas. On the one hand, there is agricultural machinery data, which records both the quantities of fertilizer applied and, indirectly via yield mapping, the nutrients applied. Various suppliers of sensor systems record soil, plant, and weather data. In the dairy farm sector, conclusions can be drawn about the nutrients entering and leaving the cycle by looking at data about the feeding together with the milk yield. Figure 7 depicts the general steps in which nutrients are generated. However, in almost all places (air, soil, groundwater, ...), nutrients are leaked, and these losses can be reduced to a certain level (although practical implementability must be given and feasible).

In this example, many different stakeholder groups are represented, which either elicit, collect, or evaluate data, or record and store data for legal documentation. In this case, the ADS approach can offer supporting services and at the same time protect the farmer's data, as not all data necessarily needs to be and should be accessed by everyone. In this example, data sovereignty plays an important role. Data is generated from dif-

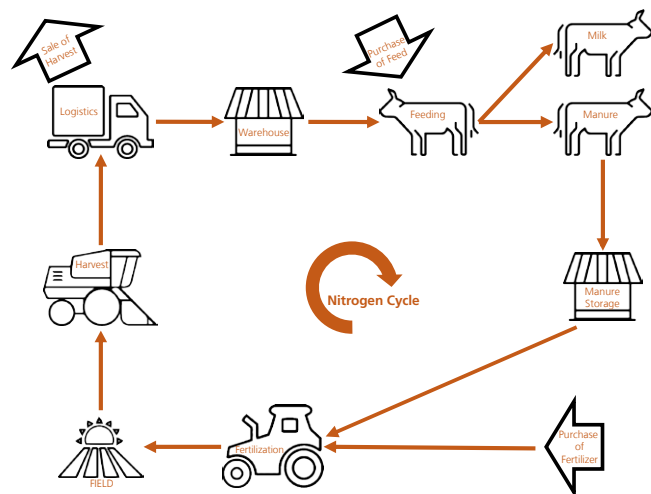


Figure 7 – Overview of nutrient cycle for crop cultivation and dairy animals

ferent categories of data (cf. section on data categories with regard to data privacy), and different data privacy regulations apply accordingly. Mainly farm-related data is collected that is not subject to legal data privacy. In order to optimize the nutrient cycle, the relevant data must be collected and must be available in sufficient quality. Important components here are interoperability, uniform ontologies, and cognitive processing of the data. Missing data, for example, must be interpolated or modeled accordingly. By representing the nutrient cycle in the form of a digital twin, the farmer can get information about their current nutrient balance and thus identify possible problem areas. Based on the digital twin, services offering the farmer appropriate decision-making aids can be purchased on the service marketplace for a fee. Since the nutrient cycle is a highly complex representation of various parameters, there are several sub-areas in which services can provide support. Examples include optimal feeding aimed at reducing the amount of nitrogen in the liquid manure, subplot-specific fertilization, or improved utilization of the nutrients in the liquid manure.

Summary and Outlook

In this whitepaper, we have presented the concept of an Agricultural Data Space, which can greatly advance digitalization in agriculture. To do so, the Agricultural Data Space takes up the concepts of the International Data Spaces Association and extends them with solutions for the agricultural sector.

This data space integrates data and services from different platforms without restricting them. An enabler platform is required for this, into which further data and services can be integrated successively, provided that other platforms implement a corresponding connector and describe data access via a service directory.

Many of the elements outlined above address current challenges, but even greater investments are required on the part of providers and users to realize the vision of a common data space for the agricultural sector.

In the context of the Fraunhofer lighthouse project “Cognitive Agriculture”, the concrete implementation and evaluation of the individual concepts presented here is planned. Companies are invited to participate in shaping the vision of an overarching data space in agriculture and to help drive its development.

About “Cognitive Agriculture”

In the lighthouse project “**Cognitive Agriculture**”, eight Fraunhofer Institutes are jointly researching the basic principles for producing agricultural products in an environmentally friendly, resource-saving, and highly efficient manner. Solution approaches include sensor technology for data collection as well as the digitalization and automation of agricultural processes. The analysis of highly complex interactions between the biosphere and production is to be used in an ecosystem of networked data and services (“**Agricultural Data Space**”) and support decision-making.

COGNAC Consortium

Fraunhofer IESE (project lead), Fraunhofer IFF, Fraunhofer IKTS, Fraunhofer IOSB, Fraunhofer IPA, Fraunhofer IPM, Fraunhofer ITWM, Fraunhofer IVI

Do you need further information or are you interested in collaborating with us? Please feel free to contact us!

Contact Person

Ralf Kalmar
ralf.kalmar@iese.fraunhofer.de
Phone: +49 631 6800-1603
www.iese.fraunhofer.de

Project Lead

Prof. Dr.-Ing. Peter Liggesmeyer

Fraunhofer Institute for Experimental Software Engineering IESE

Fraunhofer-Platz 1
67663 Kaiserslautern
Germany
www.cognitive-agriculture.de

AGRICULTURAL DATA SPACE (ADS)

WHITEPAPER



