Agris on-line Papers in Economics and Informatics

Volume XVIII Number 3, 2026

Data Management in Agricultural Enterprises: A Framework for Data-Driven Decision Support

Michal Stočes[1], Jan Jarolímek[1], Irena Benešová[2], Pavel Šimek[1], Zdeňka Žáková Kroupová[2], Ludmila Pánková [2], Jiří Vaněk[1], Eva Kánská[1], Martin Havránek[1], Michael Anderle[1]

[1] Department of Information Technologies, Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

[2] Department of Economics, Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

Abstract

The growing digitalisation of agriculture is leading to a rapid increase in the volume of data produced by agricultural enterprises; however, its practical use is often limited by fragmented data sources, varying data quality, insufficient metadata, and inconsistent storage and management practices. The aim of this article is to propose a practical framework for the systematic management of data in agricultural enterprises to support its long-term availability, reuse, and integration into decision-making processes. The proposed approach is based on four interlinked steps: identifying data needs and issues, creating and recording a data inventory, drawing up a Data Management Plan (DMP), and implementing it in practice. The framework links technical aspects of data management – such as data formats, metadata, quality, integrity, security, backup and archiving – with the definition of responsibilities for individual staff members and farm (agricultural enterprise) management. For practical implementation, two basic approaches to data storage are distinguished: file systems and data platforms or repositories. The more advanced platform-based approach enables the centralisation of heterogeneous data sources, the automation of their transfer and processing, and subsequently the integration of the acquired data with analytical tools and decision-support systems. The proposed framework thus represents a path from the isolated collection of operational data to its systematic management and utilisation as a long-term information source for both operational and strategic decision-making within an agricultural enterprise.

Keywords

Agricultural data, Data Management Plan, data quality, metadata, precision agriculture, Institutional Data Repositories, Decision Support Systems, FAIR data principles .

Stočes, M., Jarolímek, J., Benešová, I., Šimek, P., Žáková Kroupová, Z., Pánková, L., Vaněk, J., Kánská, E., Havránek, M. and Anderle, M. (2026) "Data Management in Agricultural Enterprises: A Framework for Data-Driven Decision Support", *AGRIS on-line Papers in Economics and Informatics*, Vol. 18, No. 3, pp.51-60. ISSN 1804-1930. DOI 10.7160/aol.2026.180305.

Introduction

The digitalisation of agriculture is leading to a rapid increase in the volume of data generated during routine operational, production and administrative processes. Data is generated by agricultural machinery, sensors, information systems, agricultural enterprise staff and external organisations. Its importance is growing further with the development of precision farming, automation, Business Intelligence (BI), Decision Support Systems (DSS) and artificial intelligence methods. The main sources of data for agricultural enterprises are the enterprises themselves (internal data), technology suppliers, public administration, the market environment and research organisations. The growing availability of these sources creates significant potential for data-driven decision-making (Oral et al., 2023), whilst also placing high demands on systematic data management.

However, the sheer volume of available data does not automatically guarantee its practical usability. Previous research shows that a significant proportion of agricultural data remains in formats unsuitable for subsequent analytical processing. Records often remain isolated in separate applications, on local personal computers or in suppliers' accounts. Retrieving data then depends on a specific employee's personal knowledge, and linking data requires repeated exports and manual adjustments. The most common problems include low consistency, incompleteness, a lack of context and metadata (Stočes et al., 2018), or errors arising from manual entry. Operational and analytical uses of data, however, have different requirements: whilst incomplete data may suffice for immediate machine control or operational intervention, a lack of context is a major obstacle for long-term evaluation, prediction and comparison of production cycles. Without precise knowledge of the location, time, sensors used and data acquisition methodology, it is not possible to reliably compare individual production periods.

Metadata is a key element in preserving the meaning and traceability of data. (Michener et al., 1997) Metadata describes the original source, structure, format, temporal and spatial coverage, and currency of the data, and supports their correct interpretation and integration across different information systems. In an agricultural context, metadata includes, for example, geographical location, time of collection, equipment identification, meteorological conditions, and information on soil and crops. Data that is inadequately described loses its informational value over time, and its further processing becomes unfeasible (Šimek et al., 2013).

Another significant problem is fragmented storage and limited interoperability. Data formats that are dependent on the original software mean that a file that has been technically preserved may not be practically usable in the future (proprietary data format). Another potential problem is that the data is owned by the service provider. In this context, the FAIR principles (Findable, Accessible, Interoperable, Reusable) (Kumar et. al., 2024) provide a relevant general framework not only for scientific data but also for the business environment of agricultural organisations.

This paper builds directly on the authors' previous research. The study "*Information sources in agriculture*" (Jarolímek et al., 2024) identified the main sources of agricultural data and highlighted the importance of a Data Management Plan (DMP) (Michener, 2015). for their effective and secure management. The follow-up study, "*Agriculture Data Platform – Institutional Data Repository – Selected Aspects*" (Stočes et al., 2023), focused on the technological requirements of a data platform, including repository architecture, data types, metadata, sharing, interoperability and the FAIR principles. The results show that the organisational and technological aspects of data management are closely interrelated.

The aim of this article is to propose a comprehensive and practical methodological framework for the systematic management of data in agricultural enterprises, integrating the identification of data sources, their recording, the creation of a Data Management Plan, and its subsequent implementation. The framework covers the entire data lifecycle, from acquisition and storage through to analytical processing and use for both operational and strategic decision-making within the enterprise, whilst assessing two basic options for technological implementation: an organised file system and a data platform.

Materials and methods

The framework was based on the methodology "*Data management in the agricultural sector*" (Stočes et al., 2026) and on related analyses of agricultural information sources. These materials are the result of long-term collaboration and structured interviews with representatives of agricultural, forestry and food businesses, technology suppliers and public administration bodies. The results of the authors' previous research, focusing on the architecture of data repositories and institutional data platforms, were also utilised. To design the framework, the available data sets, methods of their storage and the practical needs of business management were analytically assessed. Legal and contractual frameworks governing the protection, access and use of data were also considered, in particular the General Data Protection Regulation (GDPR) (European Parliament and Council of the European Union, 2016). and the European Data Act (European Parliament and Council of the European Union, 2023).

The methodological approach combined content analysis, qualitative comparison and design synthesis. Content analysis was used to identify and categorise data management requirements across the phases of the data life cycle: acquisition, description, storage, protection, and processing. For each requirement, the operational problem being addressed and a verifiable output were defined (e.g. assigning a metadata record to missing context, or a requirement for verifiable export to avoid dependence on a supplier).

The proposed synthesis subsequently integrates the identified requirements into a comprehensive four-step methodological framework for data management: Identification of data needs and issues – assessment of data origin, traceability and portability. Data registration and inventory creation – compilation of a structured overview of datasets, their location, formats and responsibilities. Preparation of a Data Management Plan – formalisation of rules for description, storage, backup, archiving, access and ethical and legal aspects. Implementation and roll-out – application of DMP rules in day-to-day practice, setting up processes and staff training.

The FAIR principles (Findable, Accessible, Interoperable, Reusable) have been integrated into the methodological framework.

The qualitative comparison assessed two basic architectural options for data storage and management: a controlled file system and a data platform or repository. The evaluation criteria included data organisation, metadata management, processing automation, interoperability, traceability of changes, security and overall operational complexity (Lin et al., 2020).

Metadata requirements were assessed according to their ability to ensure long-term traceability, correct interpretation and repeated processing of data without reliance on the creator's personal memory. The basic scope of a dataset description includes the dataset identifier, source, custodian, purpose of use, and temporal and spatial boundaries. At the attribute level, the framework requires the development of a data dictionary that defines the meaning of items, units of measurement, code lists, known sensor calibration conditions, and the method for recording missing or estimated values.

Drawing on the data governance roles discussed by Abraham et al. (2019), the procedure adapts this framework into four basic organisational roles to support operational discipline and enable staff to cover one another's responsibilities. To ensure operational discipline and substitutability, the procedure defines four basic organisational roles: *Originator*: creates the primary record and provides details of the circumstances of its creation. *Administrator*: checks the structure, quality, storage and documentation of the dataset. *User*: assesses the suitability of the source data for a specific decision-making task. *Company management*: provides the necessary resources, approves the DMP rules and contractual terms.

The continuity of processing was assessed through the controlled separation of three data layers: *Raw data layer*: represents data in its as-is or unmodified form from devices, applications or external services. *Cleansed data layer*: includes the results of data cleansing, the removal of duplicates, the harmonisation of units of measurement and the standardisation of identifiers. *Analytical layer*: contains aggregated datasets, spatial layers, calculated indicators and data for visualisation in analytical and decision-making tools (e.g. Microsoft Excel, Power BI, QGIS, R, Python, Grafana) (Grafana Labs, 2024; Negi et al., 2025).

A key validation criterion was the ability to unambiguously link the derived analytical output to a specific version of the input data and the calculation method used. Routine transformations and data cleansing must not overwrite the original records, whilst the retention of both the originals and the processed datasets is subject to specified retention rules and disposal periods.

This work is of a conceptual and design nature. The result is a methodological data management model adaptable to the technical, personnel and economic capabilities of a specific agricultural enterprise. As part of the assessment, a model data process was created, linking both automated (e.g. IoT data sources) and manual data collection, through the storage of primary data, to analytical models and support for both operational and strategic management decision-making.

The internal consistency of the proposal was assessed based on the logical sequence of individual steps – from the identification and recording of sources, through metadata description and secure storage, to the final analytical reports. Economic return on investment and direct operational benefits were not empirically measured at this stage and represent a direction for follow-up research in the real-world operations of agricultural enterprises. Saiz-Rubio and Rovira-Más, 2020).

Results and discussion

Classification of sources and data formats

The data overview for an agricultural holding distinguishes between three main categories of sources: agricultural enterprise data e.g. operational records, agrotechnical operations, sensor measurements, financial and economic data, public sources e.g. LPIS (Land Parcel Identification System), land registry, meteorological data, and open statistical data) and external non-public sources (paid consultancy services, supplier databases, measurements from contracted machinery).

The origin of the data, access regime and technical format are distinct characteristics. The choice of format has a significant impact on portability. Open formats with publicly available specifications protect against dependence on specific software, whilst proprietary formats require licences or conversion tools. (Falcão et al., 2023).

Data representations and requirements for their preservation

In practice, the condition for further processing means that it is not enough simply to store data physically on a disk; it is also necessary to attach a basic technical and content-related description (metadata) to it, without which the file will lose its meaning over time. If this condition is not met, the file remains technically preserved but is practically unusable for analysis. Put simply, this is a minimum set of instructions that enables both humans and software tools to correctly retrieve and accurately interpret the data in the future. Table 1 provides an overview of the most common data representations found on agricultural holdings, together with their typical uses and specific conditions for further processing (Nurseitov et al., 2009; Butler et al., 2016).

Practical implementation of the framework:

- *Identification and registration of data sources:* The organisation shall compile a register of all available datasets. For each dataset, the source, format, location, intended use, historical scope and dependence on an external supplier shall be recorded. The register will reliably identify duplicates and gaps in coverage.
- *Creation and updating of the Data Management Plan:* The DMP formalises file-naming conventions, object identifiers (e.g., unique land parcel code), data dictionaries, backup rules, retention periods, and contingency procedures in the event of a supplier service outage.
- *Defining organisational roles and responsibilities:* Effective governance requires a clear division of responsibilities:
    - *Data originator:* Responsible for the initial collection of the record and for providing details of the circumstances of its creation (e.g. machine operator, agronomist).
    - *Data manager:* Checks the structure, validity and completeness of the metadata description and the storage of the dataset.

Data Format	Typical Uses	Conditions for further processing
CSV	Spreadsheet exports, sensor logs, time series	Describe the delimiter, character encoding (UTF-8), units of measurement and column data types.
XLSX, ODS	Spreadsheet records, operational calculations and working analyses	Carefully separate the input data from the formulas; describe the structure of the worksheets and the meaning of the data cells.
GeoPackage, GeoTIFF	Vector (block boundaries) and raster (yield maps, orthophotos) spatial data	Keep the reference coordinate system (CRS, e.g. S-JTSK, WGS84), extent and resolution.
JSON, XML	Structured exchange of records between systems, e.g. REST API	Keep the data schema (XSD/JSON Schema), the meanings of the attributes and the interface version number.
Proprietary formats	Exports of machine terminals and specialised software	To provide technical documentation, a verified conversion to an open format and access to the original application.
Image formats and images (JPEG, PNG, TIFF)	Photographic documentation of vegetation, images from unmanned aerial vehicles (UAVs), sensors, remote sensing data	Retain EXIF metadata, GPS location, capture time, spectral bands, resolution and sensor calibration.
IoT data streams and time series (Parquet, InfluxDB, MQTT)	Continuous monitoring of the microclimate, soil moisture and stable operations	Define the sampling frequency, calibration curves, UTC timestamp and rules for measurement failures.

Source: Own processing

Table 1: Overview of data representations in agricultural practice.

- *Data user:* Uses the data for analytical evaluation and assesses its suitability for a given decision.
- *Company management*: Ensures the necessary capacity and investment in IT, and approves the company's data strategy.

Metadata and the separation of data layers

A key architectural principle of the proposed framework is the strict separation of original records from processed layers: *Raw Data:* This is stored in an unaltered form as received from telemetry, sensors or imports. The originals are never overwritten, which guarantees reproducibility. *Cleansed Layer* (cleansed data): Created by applying validations, standardising units, removing errors and adding identifiers. The cleansing process generates a log of the changes made. *Analytical Layer* (Processed Data): Contains aggregated datasets, yield maps, cost calculations and key performance indicators (KPIs) prepared for visualisation and as input for DSS (Wieder and Nolte, 2022).

A comparison of file-based and platform-based solutions

In the practical implementation of data management on an agricultural enterprise, a distinction is made between two basic technological levels: a managed file system and a data platform or repository. The choice of a specific architecture depends on the volume and diversity of the data generated, the size of the agricultural enterprise, the technical infrastructure, and the organisation's human and financial resources. A managed file system represents a technically less complex solution, suitable primarily for a manageable range of sources, administrative tasks and less frequent processing. Its reliability relies on strict adherence to a uniform directory structure, file-naming conventions, the maintenance of a set register and the use of accompanying metadata files or a data dictionary, which requires a high degree of discipline. In contrast, a data platform (repository) enables the centralised storage of heterogeneous data sources from machines, sensors and applications, the automation of data flows via APIs, and direct integration with analytical and decision-making tools. The platform-based approach replaces manual steps with controlled formation processes; upon ingestion, it automatically checks the structure and mandatory attributes and ensures precise traceability of data provenance across layers. However, it requires a higher level of technical management, continuous monitoring and dedicated administrative support. The transition to a data platform alone will not resolve incomplete or erroneous records; its successful implementation is contingent upon the prior harmonisation of identifiers and the standardisation of metadata descriptions as early as the file stage (Fountas et al. 2015; Wilson et al. 2017).

When selecting the technological infrastructure, the organisation is considering two main options. Table 2 presents a comparative analysis based on five defined qualitative criteria (Sawadogo and Darmont, 2021).

The criterion	Managed File System	Data Platform / Repository
Data organisation	Structured directories, naming conventions and manual set registration	Data catalogue, unique identifiers and relationships between objects
Metadata Management	Accompanying text and file metadata records and dictionaries	A centralised description integrated directly with the dataset and checks
Data Processing	Manual imports/transfers or batch scripts	Managed automated ETL/ELT (Extract, Transform, Load or Extract, Load, Transform) processes and integration channels
Traceability	File versioning and manual change logs	Automatic tracking of runs, data lineage and transformation history
Operational Complexity	Low technical complexity, high demands on discipline at work	It needs professional administration, monitoring and technical support

Source: Own processing

Table 2: Comparison of file-based and platform-based solutions.

Integration with analytical software

Cleaned and structured data from a data platform or managed repository can be used in various types of analysis and decision-support tools, depending on users' needs (Wolfert et al., 2017), for example:

- *Spreadsheet programmes* (e.g. Microsoft Excel): These are used for quick data checks, creating simple reports and pivot tables, or analysing smaller data sets.
- *Reporting and decision-support tools* (e.g. Microsoft Power BI): These enable data from different areas of the business – such as agronomic and financial data – to be linked together to create interactive reports for business management.
- *Geographic Information Systems* (e.g. QGIS): These enable the analysis of spatial data and the combined display of, for example, soil analysis results, yield maps and plot boundaries. To ensure these data are correctly linked, their coordinate systems must be aligned.
- *Programming languages and data analysis tools* (e.g., R and Python): These are used for advanced statistical analysis, data processing, machine learning, and the automation of repetitive procedures.
- *Tools for real-time monitoring and data visualisation* (e.g. Grafana): These are used to create dashboards that display trends in measured values over time and the current status of equipment, for example, based on data from sensor networks. They also allow automatic alerts to be set up when specified thresholds are exceeded or abnormal conditions occur.

The legal assessment of data management emphasises that the requirement to "own the data" necessitates a contractual definition of specific user rights. Contracts with technology suppliers should guarantee the availability of complete and unaltered data exports, including the necessary metadata. Attention must also be paid to proprietary formats, whose use may be restricted to a specific supplier's software. To mitigate the risk of vendor lock-in (dependence on a single supplier), it is advisable to contractually ensure the ability to export data to open or widely supported formats that enable transfer to other systems and further use. The framework also takes into account compliance with the GDPR when handling employees' personal data, as well as the provisions of the European Data Act, which, among other things, regulates users' rights to access data generated during the use of connected agricultural machinery and equipment and to share such data with third parties. (Wiseman et al., 2019).

It is recommended that the proposed framework be implemented gradually, considering the agricultural enterprise size, available technical equipment, and staff capacity. A suitable starting point is to select a single high-value-added area, such as monitoring costs and yields for individual land parcels. This task can be used to verify the availability and quality of the necessary data, the possibilities for linking it, and the usefulness of the resulting reports for decision-making. Experience gained from the initial implementation can then be used to refine data management rules and to gradually extend the framework to other areas of the agricultural enterprise. The framework is designed as a generally applicable methodological guide, the specific form of which must be tailored to the needs and capabilities of the individual agricultural enterprise. The framework itself does not constitute proof of economic return or an increase in operational efficiency. Verification of these benefits will be the subject of follow-up research based on the practical implementation of the framework and an evaluation of the results achieved. The evaluation should include, for example, the costs of implementation and operation, the time required to prepare data and reports, the error rate of the data, and its availability for decision-making. A comparison of the baseline situation with the situation following the implementation of the framework will enable an assessment of its actual benefits and limitations under the specific conditions of the agricultural enterprise (Žáková Kroupová et al., 2025).

Conclusion

The proposed approach demonstrates that the effective use of data on an agricultural enterprise requires more than just its collection and technical storage. A key prerequisite is the systematic management of data throughout its entire lifecycle, linking the identification of data sources, the creation of a data inventory, the definition of operational rules within the Data Management Plan, and the subsequent implementation of these rules into the agricultural enterprise's day-to-day operations. A process setup in this way creates the conditions for the long-term availability, traceability, quality, and reusability of data in accordance with the international FAIR principles.

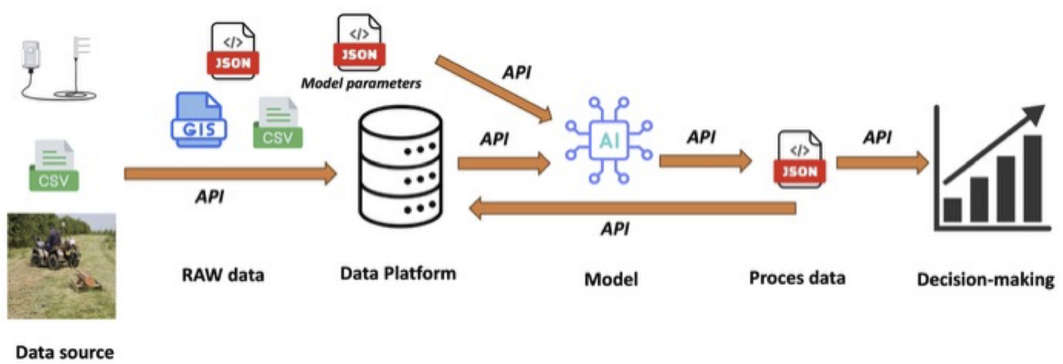
The results confirm the particular importance of data quality and standardised data description. Data that is inconsistent, insufficiently complete or lacking appropriate metadata has limited value for advanced analytics, decision support systems (DSS) and the application of machine learning and artificial intelligence methods. Data quality is not merely a technical issue but is fundamentally influenced by human factors and established working procedures (Milleret el., 2024). For this reason, the Data Management Plan should be viewed as a continuously updated management tool, rather than a one-off document. Regular reviews, updates, staff training and adaptation to new technologies and legislative requirements (e.g. the GDPR and the Data Act) are essential for its long-term effectiveness.

A key benefit of the framework is the comprehensive integration of organisational data management with its technological storage and automated analytical processing. As illustrated by the model data process (Figure 1), the entire chain links together the individual stages from primary collection through to decision support:

- *Acquisition of primary data* (Data Source → RAW Data): Data from automated IoT sensors, mobile measuring devices or structured files (CSV, GIS) is transferred via an application programming interface (API) to the raw data layer.
- *Centralised storage* (Data Platform): The data platform stores raw records in their unprocessed form (RAW data), manages metadata structures and model parameters, and strictly separates the original data from processed datasets.
- *Model processing and AI* (Model / AI → Processed Data): The platform communicates via a bidirectional API with analytical and predictive models utilising artificial intelligence algorithms, which transform the raw data into validated and structured derived data (Processed Data).
- *Decision-making support:* The processed data is delivered to management dashboards, GIS systems or analytical tools for both operational and strategic business management.

Practical implementation also confirms that no single technological solution is suitable for all agricultural businesses. The choice between a managed file system and an advanced data platform must be tailored to the agricultural enterprise's size, the nature and volume of the data being processed, the available infrastructure, and the agricultural enterprise’s human and financial resources. In both cases, however, contractual control over the data, format portability and the avoidance of vendor lock-in are essential.

The main benefit of the proposed framework is therefore the possibility of a gradual transition for an agricultural enterprise from a situation in which data is used only once for a specific operational task to one in which it constitutes a long-term managed information resource. Further research should focus on the practical validation of the entire framework in enterprises of various sizes and on quantifying its impact on data quality, data management costs and the effectiveness of decision-making processes.



Source: Own processing

Figure 1: Automated model data process.

Acknowledgements

The results and knowledge included herein have been obtained owing to support from the following institutional grant: Internal grant agency of the Faculty of Economics and Management, Czech University of Life Sciences Prague, grant no. 2026A007.

Corresponding author:
doc. Ing. Pavel Šimek, Ph.D.
Department of Information Technologies, Faculty of Economics and Management
Czech University of Life Sciences Prague
Kamýcká 129, 165 00 Prague-Suchdol, Czech Republic
Phone: +420 224386680, E-mail: simek@pef.czu.cz

References

[1] Abraham, R., Schneider, J. and vom Brocke, J. (2019) "Data governance: A conceptual framework, structured review, and research agenda", *International Journal of Information Management*, Vol. 49, pp. 424-438. ISSN 0268-4012. DOI 10.1016/j.ijinfomgt.2019.07.008.

[2] Butler, H., Daly, M., Doyle, A., Gillies, S., Hagen, S. and Schaub, T. (2016) "*The GeoJSON Format*", RFC 7946, Internet Engineering Task Force (IETF), Aug. 2016. ISSN 2070-1721. DOI 10.17487/RFC7946.

[3] European Parliament and Council of the European Union. (2016) "Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)", *Official Journal of the European Union*, L 119, 4 May 2016, pp. 1-88. ISSN 1977-0677 [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj. [Accessed: 15 Jun. 2026].

[4] European Parliament and Council of the European Union. (2023) "Regulation (EU) 2023/2854 of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act)", *Official Journal of the European Union*, L, 2023/2854, 22 Dec. 2023. E-ISSN 1977-0677 [Online]. Available: https://eur-lex.europa.eu/eli/reg/2023/2854/oj. [Accessed: 15 Jun. 2026].

[5] Falcão, R., Matar, R., Rauch, B., Elberzhager, F. and Koch, M. (2023) "A Reference Architecture for Enabling Interoperability and Data Sovereignty in the Agricultural Data Space", *Information*, Vol. 14, No. 3, p. 197. ISSN 2078-2489. DOI 10.3390/info14030197.

[6] Fountas, S., Carli, G., Sørensen, C. G., Tsiropoulos, Z., Cavalaris, C., Vatsanidou, A., Liakos, B., Canavari, M., Wiebensohn, J. and Tisserye, B. (2015) "Farm management information systems: Current situation and future perspectives", *Computers and Electronics in Agriculture*, Vol. 115, pp. 40-50. E-ISSN 1872-7107. DOI 10.1016/j.compag.2015.05.011.

[7] Grafana Labs. (2024) "*Grafana Alerting*", Grafana documentation. [Online]. Available: https://grafana.com/docs/grafana/latest/alerting/. [Accessed: 17 Jun. 2025].

[8] Jarolímek, J., Samek, J., Šimek, P., Stočes, M., Vaněk, J. and Pavlík, J. (2024) "Information sources in agriculture", *Plant, Soil and Environment*, Vol. 70, No. 11, pp. 712-718. E-ISSN 1805-9368. DOI 10.17221/361/2024-PSE.

[9] Kumar, P., Hendriks, T., Panoutsopoulos, H. and Brewster, C. (2024) "Investigating FAIR data principles compliance in Horizon 2020 funded agri-food and rural development multi-actor projects", *Agricultural Systems*, Vol. 214, p. 103822. E-ISSN 1873-2267. ISSN 0308-521X. DOI 10.1016/j.agsy.2023.103822.

[10] Lin, D., Crabtree, J., Dillo, I., Downs, R. R., Edmunds, R., Giaretta, D., De Giusti, M., L'Hours, H., Hugo, W., Jenkyns, R., Khodiyar, V., Martone, M. E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D. V., Stockhause, M. and Westbrook, J. (2020) "The TRUST Principles for digital repositories", *Scientific Data*, Vol. 7, p. 144. ISSN 2052-4463. DOI 10.1038/s41597-020-0486-7.

[11] Michener, W. K., Brunt, J. W., Helly, J. J., Kirchner, T. B. and Stafford, S. G. (1997) "Nongeospatial Metadata for the Ecological Sciences", *Ecological Applications*, Vol. 7, No. 1, pp. 330-342. E-ISSN 1939-5582. DOI 10.1890/1051-0761(1997)007[0330:NMFTES]2.0.CO;2.

[12] Michener, W. K. (2015) "Ten Simple Rules for Creating a Good Data Management Plan", *PLOS Computational Biology*, Vol. 11, No. 10, p. e1004525. ISSN 1553-7358. DOI 10.1371/journal.pcbi.1004525.

[13] Miller, R., Whelan, H., Chrubasik, M., Whittaker, D., Duncan, P. and Gregório, J. (2024) "A Framework for Current and New Data Quality Dimensions: An Overview", *Data*, Vol. 9, No. 12, p. 151. ISSN 2306-5729. DOI 10.3390/data9120151.

[14] Negi, A., Tripathi, L., Prabha, S. and Azad, C. (2025) "Charting the Future: An In-Depth Review of Cuttingedge Big Data Visualization Tools for 2025", *2025 12th International Conference on Computing for Sustainable Global Development (INDIACom), Delhi, India, IEEE,* pp. 1-6. ISSN 0973-7529. DOI 10.23919/INDIACom66777.2025.11115697.

[15] Nurseitov, N., Paulson, M., Reynolds, R. and Izurieta, C. (2009) "Comparison of JSON and XML Data Interchange Formats: A Case Study", *Proceedings of the ISCA 22nd International Conference on Computer Applications in Industry and Engineering (CAINE), San Francisco, USA, ISCA,* pp. 157-162. ISBN 9781615676668. [Online]. Available: https://www.cs.montana.edu/izurieta/pubs/caine2009.pdf. [Accessed: 20 Jun. 2025].

[16] Oral, E., Chawla, R., Wijkstra, M., Mahyar, N. and Dimara, E. (2023) "From Information to Choice: A Critical Inquiry Into Visualization Tools for Decision Making", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 30, No. 1, pp. 359-369. ISSN 1077-2626. DOI 10.1109/TVCG.2023.3326593.

[17] Saiz-Rubio, V. and Rovira-Más, F. (2020) "From Smart Farming towards Agriculture 5.0: A Review on Crop Data Management", *Agronomy*, Vol. 10, No. 2, p. 207. ISSN 2073-4395. DOI 10.3390/agronomy10020207.

[18] Sawadogo, P. and Darmont, J. (2021) "On data lake architectures and metadata management", *Journal of Intelligent Information Systems*, Vol. 56, pp. 97-120. E-ISSN 1573-7675. DOI 10.1007/s10844-020-00608-7.

[19] Stočes, M., Jarolímek, J., Benešová, I., Cihelka, P., Čechura, L., Havránek, M., Žáková Kroupová, Z., Pánková, L. and Vaněk, J. (2026) "*Správa dat v zemědělském sektoru*", Prague, Faculty of Economics and Management, Czech University of Life Sciences Prague. DOI 10.83124/7yfg4-bqs05. (In Czech).

[20] Stočes, M., Šilerová, E., Vaněk, J., Jarolímek, J. and Šimek, P. (2018) "Možnosti využití otevřených dat v sektoru cukr-cukrová řepa", *Listy cukrovarnické a řepařské*, Vol. 134, No. 3, pp. 117-121. ISSN 1210-3306. (In Czech).

[21] Stočes, M., Vaněk, J., Jarolímek, J., Novák, V., Masner, J., Šimek, P., Kánská, E., Havránek, M., Kubata, K. and Voral, V. (2023) "Agriculture Data Platform-Institutional Data Repository-Selected Aspects", *AGRIS on-line Papers in Economics and Informatics*, Vol. 15, No. 4, pp. 127-133. ISSN 1804-1930. DOI 10.7160/aol.2023.150409.

[22] Šimek, P., Vaněk, J., Jarolímek, J., Stočes, M. and Vogeltanzová, T. (2013) "Using metadata formats and AGROVOC thesaurus for data description in the agrarian sector", *Plant, Soil and Environment*, Vol. 59, No. 8, pp. 378-384. E-ISSN 1805-9368. DOI 10.17221/261/2013-PSE.

[23] Wieder, P. and Nolte, H. (2022) "Toward data lakes as central building blocks for data management and analysis", *Frontiers in Big Data*, Vol. 5, p. 945720. E-ISSN 2624-909X. DOI 10.3389/fdata.2022.945720.

[24] Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L. and Teal, T. K. (2017) "Good enough practices in scientific computing", *PLOS Computational Biology*, Vol. 13, No. 6, p. e1005510. ISSN 1553-7358. DOI 10.1371/journal.pcbi.1005510.

[25] Wiseman, L., Sanderson, J., Zhang, A. and Jakku, E. (2019) "Farmers and their data: An examination of farmers' reluctance to share their data through the lens of the laws impacting smart farming", *NJAS-Wageningen Journal of Life Sciences*, Vol. 90-91, p. 100301. E-ISSN 2212-1307. DOI 10.1016/j.njas.2019.04.007.

[26] Wolfert, S., Ge, L., Verdouw, C. and Bogaardt, M.-J. (2017) "Big Data in Smart Farming-A review", *Agricultural Systems*, Vol. 153, pp. 69-80. ISSN 0308-521X. DOI 10.1016/j.agsy.2017.01.023.

[27] Žáková Kroupová, Z., Aulová, R., Rumánková, L., Bajan, B., Čechura, L., Šimek, P. and Jarolímek, J. (2025) "Drivers and barriers to precision agriculture technology and digitalisation adoption: Meta-analysis of decision choice models", *Precision Agriculture*, Vol. 26, p. 17. E-ISSN 1573-1618. DOI 10.1007/s11119-024-10213-1.