

Multifunctional Data in a Multifunctional Landscape

A Concept Paper for Cross-Community Research Data Management for Near-Surface Geophysical Data

U. Werban¹, C. Hoffmann², J. Seegert³, C. Keller⁴, C. Schütze¹

¹ Helmholtz Center for Environmental Research - UFZ

² Leibniz Centre for Agricultural Landscape Research

³ TUD Dresden University of Technology

⁴ German Archaeological Institute

Near-surface geophysical data play a crucial role in multiple scientific domains, including precision agriculture, archaeological prospecting, and Earth System Sciences. Although the methodology and sensor technology used is often similar or even the same, data processing methodologies, (meta)data handling and long-term archiving practices remain highly fragmented. This limits cross-domain research data interoperability, comparability, accessibility, and reusability. To address these challenges, this approach seeks to integrate efforts across three key NFDI consortia: FAIRagro, **NFDI4Objects**, and **NFDI4Earth**.

The primary objective of this initiative is to develop a harmonized framework for geophysical data workflows, ensuring FAIR (Findable, Accessible, Interoperable, and Reusable) principles are consistently applied across disciplines. By leveraging expertise from agricultural research (FAIRagro), archaeology and cultural heritage protection (NFDI4Objects) and environmental and Earth System Sciences (NFDI4Earth), this initiative will help to establish workflows for standardized data formats and metadata schemas, and best practices for geophysical data handling. Additionally, it will foster interoperability between existing research infrastructures and promote sustainable data management. Hence, this approach aligns with the overarching goals of the NFDI by enhancing data accessibility and reusability, fostering interdisciplinary collaboration, and supporting data-driven research. By bridging domain-specific methodologies and establishing common standards, this initiative will significantly improve the integration of geophysical datasets across scientific fields, ultimately benefiting researchers, policymakers, and stakeholders engaged in environmental monitoring, heritage conservation, and precision agriculture.

The above-mentioned multiple scientific domains share significant overlap in their use of near-surface geophysical investigations to describe and analyze subsurface structures and processes. Despite their distinct primary objectives, the bridging elements among these disciplines include methods, instruments, and research goals, which allow for cross-disciplinary data reuse and methodological integration. The considered research fields rely on the use of geophysical methods as a proxy for soil properties, functional characteristics and state variables. Geophysical datasets can help to describe the targeted subsurface phenomena, albeit with different interpretations for each field. The same geophysical instruments are applied in the different communities and generate datasets relevant to multiple disciplines (e.g. Doolittle & Brevik, 2014; Boaga, 2017; De Smedt et al., 2014). Hence, sharing datasets maximizes research efficiency and broadens interdisciplinary insights (Webber et al., 2017; Cuenca-Garcia et al., 2018).

Recently, four primary challenges hinder to achieve effective cross-domain geophysical data reuse.

(1) Data discovery and accessibility (finding and getting the data):

Geophysical data are often collected for project-specific purposes and remain stored in local or inaccessible repositories. Even when funding agencies mandate data availability, the lack of standardized publication workflows means datasets are often only accessible through informal

networks. Access restrictions due to proprietary concerns (e.g., industry-funded agricultural projects) or data sensitivity (e.g., cultural heritage protection) further complicate data sharing.

(2) Inconsistent use of public repositories:

While open-data repositories (like Zenodo, PANGAEA, ...) exist, there is no widely adopted, discipline-specific protocol for archiving geophysical data in a way that supports cross-disciplinary discovery and reuse. Different research groups use varied metadata standards and file formats, making it difficult for potential users from other fields to find, interpret, and integrate the data into their own research.

(3) Lacking standards in metadata and documentation:

Even when data are uploaded to open-access repositories, they often lack sufficient contextual metadata (e.g., instrument settings, georeferencing information, acquisition context) and information about (environmental) boundary conditions (e.g., soil type, land-use, meteorological situation) making it impossible to assess the validity of the data or even the fit-for-purpose evaluation for new research questions. Different scientific communities use varying terminologies and classification schemes, making cross-domain data interpretation difficult. Without structured and standardized metadata and vocabularies, datasets become difficult to interpret or integrate with other research domains, reducing their reuse potential.

(4) Differences in data processing and interoperability:

Data formats, covered scales, depths, spatial and temporal resolution, and applied processing pipelines differ between disciplines, creating challenges in direct data exchange. For some research questions (e.g. in soil sciences) additional environmental data (such as soil properties, meteorological conditions, vegetation coverage, ...) might be required in terms of an increased data and process understanding.

To start a process to overcome the mentioned challenges, we propose a **structured three-phase concept**, demonstrated at a characteristic pilot site. This site will serve as a testbed for acquiring and systematically analyzing comprehensive near-surface geophysical data across three disciplines. The selected pilot site is of particular interest to all relevant fields, as it features characteristics suitable for agricultural, archaeological, and Earth System Sciences investigations. The focus will be on developing harmonized workflows from data acquisition to archiving, establishing standardized data and metadata structures, and defining cross-disciplinary best practices to ensure multi-purpose dataset usability. This could follow a hierarchical approach integrating different geophysical methods in a stepwise manner. Each phase introduces new geophysical methods based on the specific needs of the considered disciplines, while leveraging existing datasets from previous steps. This ensures that data acquired in a specific domain remains useful for other user groups, enhancing the collateral benefit of multi-purpose usability.

Phase 1 - Agricultural Prospection: The objective of agricultural prospection ("agrogeophysics") is to assess e.g., soil variability, soil moisture distribution, and nutrient status, which are essential for precision farming (Garré et al., 2021; Rhymes et al., 2023). The great advantage of "agrogeophysical" methods is their potential rapidity, low cost, and spatial coverage and thus can easily be integrated in precision farming practices both in terms of prospection and monitoring. In this phase, Electromagnetic Induction (EMI) will be used to map soil conductivity that is related to soil texture, soil water content and salinity, while Gamma-Ray Spectroscopy (GRS) will provide additional insights into soil texture (clay content) and can thus support management decisions. Another promising approach is represented by multiple proximal sensing technologies with integrating multi-sensor platforms, enabling simultaneous measurements and sensor data fusion. Thus, a broader range of soil properties can be rapidly assessed, enhancing the size of the covered areas and the soil property estimation accuracy (Tavakoli et al., 2024). These geophysical measurements also reveal distributed subsurface anomalies indicating past human activity or cultural remains, providing a valuable foundation for subsequent archaeological investigations. At this step the implementation of a harmonized workflow is required, ensuring that the data and metadata generated in this phase are structured, standardized, and properly documented for their integration into archaeological prospection. This includes the availability of standardized protocols essential for the structured acquisition, storage, and processing of the EMI, GRS or multi-sensor platform data and metadata as well as the consideration and further development of common

standardized vocabularies for sensor metadata and soil property description. The activities in Phase 1 are particularly relevant to the *FAIRagro consortium*.

Phase 2 - Archaeological Prospection: Archaeogeophysics will be applied to enhance structural analysis by high-resolution methods and to identify subsurface archaeological structures (walls, roads, and ditches). The data from EMI, GRS and multi-sensor platforms from agricultural prospection derived in phase 1 will be utilized to pre-identify areas with potential anthropogenic soil disturbances in order to design Ground Penetrating Radar (GPR) surveys and to optimize their efficiency and precision. GPR surveys are the preferred method to provide high-resolution 3D-imaging of buried features (Wunderlich et al., 2024). In this phase the further elaboration of standardized protocols and workflow harmonization is required and the degree of cross-disciplinary FAIRness (mainly re-usability) will be tested and further improved. Existent vocabularies will be enhanced concerning relevant terminologies concerning method and discipline to ensure consistency in data description. The activities in Phase 2 are particularly relevant to the *NFDI4Objects consortium*.

Phase 3 - Observation of Soil Dynamics: Monitoring the dynamic subsurface is essential for understanding processes in soil science, hydrology, hydrogeology, including anthropogenic influences over time. Geophysical methods - especially the time-lapse concepts - can capture temporal variations in soil moisture, soil stability, soil temperature, and soil integrity; enabling a more comprehensive approach to long-term environmental assessments beyond agricultural requirements. A widely used method is Electrical Resistivity Tomography (ERT) providing 3D subsurface imaging of e.g., water infiltration, root penetration, and soil degradation (Garré et al., 2011; Perrone et al., 2014; Dimeche et al., 2022). In this phase the workflow harmonization will be further elaborated with special focus on the time-lapse methods including the expansion to incorporate the perspectives of soil processes. Moreover, the comprehensive integration of all applied method data holds immense potential for generating deeper insights, improving data reliability, and uncovering previously undetectable subsurface patterns. In the final phase all pilot data will be provided resulting in a comprehensive subsurface model that unites agricultural, archaeological, and soil science perspectives, while also generating an AI-ready dataset for multi-purpose geophysical modeling. The activities in Phase 3 are particularly relevant to the *NFDI4Earth consortium*.

The three-phase concept is designed to **enable bidirectional data reuse**, ensuring that datasets generated in later phases also provide added value to earlier research fields. For example, high-resolution GPR data from archaeological investigations (Phase 2) can contribute to agricultural soil assessments or model validation if properly described and archived. Similarly, ERT datasets from Earth System Sciences (Phase 3) can offer critical insights for both archaeology and precision farming, supporting a more comprehensive understanding of subsurface structures and processes across all disciplines. In addition, if a geophysical dataset acquired for agricultural purposes are structured and archived in a way that makes them accessible to archaeological research and Earth System Science, benefits will emerge in (1) increased data availability and cost efficiency, (2) enhanced knowledge generation through multi-purpose data, (3) improved AI-driven predictive modeling by leveraging diverse datasets across disciplines, enabling automated pattern recognition, adaptive learning, and cross-domain insights for applications such as soil property mapping, archaeological feature detection, and land-use change analysis, and (4) interdisciplinary insights and holistic landscape understanding.

However, the aim to provide multifunctional near-surface geophysical data also means raising awareness, capacity building and fostering a willingness among communities to embrace broad usability of this data. Achieving this requires effective communication and tailored training efforts across the different user groups. This can be achieved, for example, through presentations at specialized conferences or through targeted workshops where different user groups can share knowledge, gain hands-on experience with cross-domain geophysical data, and develop a common understanding of best practices for data management and joint interpretation. These efforts will be integrated into all three phases of the proposed concept, ensuring that awareness-raising, communication, and training accompany each phase.

The harmonization of geophysical data across the three domains of agriculture, archaeology, and Earth System Sciences presents a major opportunity for efficient data reuse and interdisciplinary research. With the proposed initiative we will foster harmonization and standardization along the geophysical data life-cycle across the three disciplines, enhancing multi-purpose usability through FAIR and TRUST principles, while bridging natural and social sciences for greater interdisciplinary impact. Furthermore, **cross-disciplinary data is a unique selling point for the development of data-driven models and the application of AI approaches**. AI models, particularly in geosciences, agriculture, archaeology, and soil science, benefit significantly from diverse datasets that capture a wider range of environmental and subsurface features. Providing harmonized geophysical datasets will ensure the embedding of AI-readiness into our standardized geophysical data framework as well as a valuable integration into future AI-driven applications.

This showcase is connecting three branches of science and addresses the research data domains: Life Science (FAIRagro), Humanities (NFDI4Objects) and Natural Sciences (NFDI4Earth). It is intended to serve as a role model for future cross-disciplinary cooperation and esp. data handling and re-usability approaches and is an initial cross-consortia application-related cornerstone towards "oneNFDI".

References

- Boaga, J.: The use of FDEM in hydrogeophysics: A review, *Journal of Applied Geophysics*, 139, 36-46, <https://doi.org/10.1016/j.jappgeo.2017.02.011>, 2017.
- Cuenca-Garcia, C., Armstrong, K., Aidona, E., De Smedt, P., Rosveare, A., Rosveare, M., Schneidhofer, P., Wilson, C., Faßbinder, J., Moffat, I., Sarris, A., Scheiblecker, M., Jrad, A., van Leusen, M., and Lowe, K.: THE SOIL SCIENCE & ARCHAEO-GEOPHYSICS ALLIANCE (SAGA): going beyond prospection, *Research Ideas and Outcomes*, 4, e31648, 2018. <https://doi.org/10.3897/rio.4.e31648>
- De Smedt, P., Van Meirvenne, M., Saey, T., Baldwin, E., Gaffney, C., and Gaffney, V.: Unveiling the prehistoric landscape at Stonehenge through multi-receiver EMI, *Journal of Archaeological Science*, 50, 16-23, <https://doi.org/10.1016/j.jas.2014.06.020>, 2014.
- Dimech, A., Cheng, L., Chouteau, M., Chambers, J., Uhlemann, S., Wilkinson, P., Meldrum, P., Mary, B., Fabien-Ouellet, G., and Isabelle, A.: A Review on Applications of Time-Lapse Electrical Resistivity Tomography Over the Last 30 Years: Perspectives for Mining Waste Monitoring, *Surveys in Geophysics*, 43, 1699-1759, <https://doi.org/10.1007/s10712-022-09731-2>, 2022.
- Doolittle, J. A. and Brevik, E. C.: The use of electromagnetic induction techniques in soils studies, *Geoderma*, 223-225, 33-45, <https://doi.org/10.1016/j.geoderma.2014.01.027>, 2014.
- Garré, S., Javaux, M., Vanderborght, J., Pagès, L., and Vereecken, H.: Three-Dimensional Electrical Resistivity Tomography to Monitor Root Zone Water Dynamics, *Vadose Zone Journal*, 10, 412-424, <https://doi.org/10.2136/vzj2010.0079>, 2011.
- Garré, S., Hyndman, D., Mary, B., and Werban, U.: Geophysics conquering new territories: The rise of "agrogeophysics", *Vadose Zone Journal*, 20, e20115, <https://doi.org/10.1002/vzj2.20115>, 2021.
- Perrone, A., Lapenna, V., and Piscitelli, S.: Electrical resistivity tomography technique for landslide investigation: A review, *Earth-Science Reviews*, 135, 65-82, <https://doi.org/10.1016/j.earscirev.2014.04.002>, 2014.
- Rhymes, J., Chadwick, D. R., Williams, A. P., Harris, I. M., Lark, R. M., and Jones, D. L.: Evaluating the accuracy and usefulness of commercially-available proximal soil mapping services for grassland nutrient management planning and soil health monitoring, *Precision Agriculture*, 24, 898-920, <https://doi.org/10.1007/s11119-022-09979-z>, 2023.
- Tavakoli, H., Correa, J., Vogel, S., Oertel, M., Zimne, M., Heisig, M., Harder, A., Wruck, R., Pätzold, S., Leenen, M., and Gebbers, R.: The RapidMapper: State-of-the-art in mobile proximal soil sensing based on a novel multi-sensor platform, *Computers and Electronics in Agriculture*, 226, 109443, <https://doi.org/10.1016/j.compag.2024.109443>, 2024.
- Webber, H., Heyd, V., Horton, M., Bell, M., Matthews, W., and Chadburn, A.: Precision farming and archaeology, *Archaeological and Anthropological Sciences*, 11, 727-734, <https://doi.org/10.1007/s12520-017-0564-8>, 2017.
- Wunderlich, T., Majchczack, B. S., Wilken, D., Segschneider, M., and Rabbel, W.: What Is Beyond Hyperbola Detection and Characterization in Ground-Penetrating Radar Data?—Implications from the Archaeological Site of Goting, Germany, <https://doi.org/10.3390/rs16214080>, 2024.