AI4SoilHealth

AI4SoilHealth.eu

Accelerating collection and use of soil health information using AI technology to support the Soil Deal for Europe and EU Soil Observatory

Project title: AI4SoilHealth — an open access European-wide digital infrastructure for monitoring and forecasting soil health indicators
Horizon Europe call: HORIZON-MISS-2021-SOIL-02-02 (Research and Innovation and other actions to support the implementation of a mission in the area of Soil health and Food)
Type of Action: HORIZON-RIA
Summary: The objective of AI4SoilHealth is to co-design, create and maintain an open access European-wide digital infrastructure, compiled using state-of-the-art Artificial Intelligence (AI) methods combined with new and deep soil health understanding and measures. The AI-based data infrastructure functions as a Digital Twin to the real-World biophysical system, forming a Soil Digital Twin. This can be used for assessing and continuously monitoring Soil Health metrics by land use and/or management parcel, supporting the Commission's objective of transitioning towards healthy soils by 2030. The project is divided into seven (7) work-packages including: (WP2) Policy and stakeholder engagement — networking and synchronizing with EU and national programs, (WP3) Soil health methodology and standards - developing/testing methodology to be used by WPs 4-6, (WP4) Soil health in-situ monitoring tools and data - developing field and laboratory solutions for Observations & Measurements, (WP5) Harmonised EU-wide soil monitoring services - developing the final suite of tools, data and services, (WP6) Multi-actor engagement pilots — organizing field-works and collect users' feedback, (WP7) Soil literacy, capacity building and communication — organizing public campaigns and producing educational materials. Key deliverables include: 1) Coherent Soil Health Index methodology, 2) Rapid Soil Health Assessment Toolbox, 3) AI4SoilHealth Data Cube for Europe, 4) Soil-Health-Soil-Degradation-Monitor, and 5) AI4SoilHealth API and Mobile phone App. Produced tools will be exposed to target-users (including farmer associations in >10 countries), so their feedback is used to improve design/functionality. Produced high-resolution pan-European datasets will be distributed under an Open Data license, allowing easy access by development communities. AI4SoilHealth will provide an effective Soil Health Index certification system to support landowners and policy makers under the new Green Deal for Europe.
Keywords: Biogeochemistry, biogeochemical cycles, environmental chemistry, Earth observations from space/remote sensing, Environment, resources and sustainability, Environmental monitoring systems, Terrestrial ecology, land cover change.
Project duration: Jan 2023 – Dec 2026
Total requested budget: 13,789,328 EUR

Soil health monitoring: why is it important?

Why AI4SoilHealth and why now? Unsustainable land use and population pressure are major drivers of soil degradation (Borrelli *et al.*, 2017; Hengl *et al.*, 2021a). The threats to soil health include loss of soil organic matter, biodiversity, and the soil itself through erosion. We are at a crossroads in history in our attempt to preserve European and global soil resources; the overexploitation causing the US dust bowl is still within living memory (Heslin *et al.*, 2020). It is a striking paradox that on the one hand soils are one of the solutions to mitigating greenhouse gas emissions, while on the other hand **60–70% of EU soils are unhealthy**, mainly because of unsustainable management practices.

To prevent soil degradation and help combat climate change, the European Commission has set an ambitious goal with several concrete targets such as **net zero emissions** and **planting 3 billion trees by 2030**. By 2028, every land manager should have access to **verified [CO_2] emission and removal data** at parcel level, and **carbon farming** should support the achievement of the proposed 2030 net removal target of 310 Mt CO_2eq in the land sector across EU[1]. The ability to measure and assess progress toward healthy soils will be critical to the success of improving soil health across Europe for both stakeholders and policy-makers. Currently eight (8) soil health indicators have been proposed to assess the European Mission Board (MB) for Soil Health and Food objectives, these are: (1) Presence of soil pollutants, excess nutrients and salts; (2) Soil organic carbon stock; (3) Soil structure including soil bulk density and absence of soil sealing and erosion; (4) Soil biodiversity; (5) Soil nutrients and acidity (pH); (6) Vegetation cover; (7) Landscape heterogeneity, and (8) Forest cover.

AI (Artificial Intelligence) encompasses software and big data solutions that enhance and automate decision systems with a potential to replace human labor and automate and improve a diverse range of processes. AI is proven to deliver services such as Earth Observation (EO) processing and geosciences and in many more applications[2]. In soil science, especially through **pedometrics**, AI is forging a path for a diverse range of applications including soil measurements, soil image recognition and soil mapping. Now is the moment to systematically apply cutting edge AI methods to soil datasets and measurements and build an automated data-driven decision support for soils. We propose to develop a robust indicator framework by harnessing the advances in AI and digital technologies that present new and untapped opportunities for monitoring the pressures and condition of soils and land. For this, we will utilize a pan-European soil monitoring system (LUCAS) in conjunction with EO and ancillary data to produce **AI4SoilHealth — an open access European-wide digital infrastructure for monitoring and forecasting soil health indicators**.

Project objectives

The goal of this proposal is to support the EU 'The Soil Health & Food Mission'[3] mission towards reaching its mission goals and targets set by the **EU Soil Strategy 2030[4]**. We plan to co-design, create, and maintain an open access **European-wide digital infrastructure**, termed ***"AI4SoilHealth"*.** The infrastructure will be used for assessing, and continuously monitoring, soil health metrics by land use and/or management. We will build AI4SoilHealth in logical steps:

(1) Identify robust & realistic **Soil Health Proxies** that can detect state and change in soil health relative to the desired soil ecosystem functions/services;
(2) Test new proxies based on emerging methods such as soil spectroscopy and molecular arrays that show high promise for characterizing both state and change in soil health;
(3) Create a digital infrastructure named a ***"Soil Digital Twin"*** (representing the physical environment in a big-data digital environment; Fig. 1) that can collect and integrate enormous amounts of soil data. Soil health information can then be effectively shared by stakeholders, including policy-makers and soil professionals in an open access environment; mapped & analyzed at scales ranging from single farms to all of Europe. The Soil Digital Twin will be based on the substantial developments in **Artificial Intelligence** (AI) methods such as machine learning and artificial neural networks.

We will test predictions on independent national and thematic datasets, and use these tests to guide directed sampling field campaigns, including in our, and other, living labs. This will also enable us to test emerging methods such as spectral measurements and genomics to validate and further develop a next-

[1] https://ec.europa.eu/clima/eu-action/forests-and-agriculture/sustainable-carbon-cycles_de
[2] https://unchartedterritories.tomaspueyo.com/p/ai-does-it-best
[3] https://data.europa.eu/doi/10.2777/918775
[4] https://ec.europa.eu/environment/publications/eu-soil-strategy-2030_en

generation indicator framework for soil health. We will further harness expertise in the consortium to develop a robust indicator framework sitting within the concept of a '*Soil Digital Twin*' (Grieves, 2017).

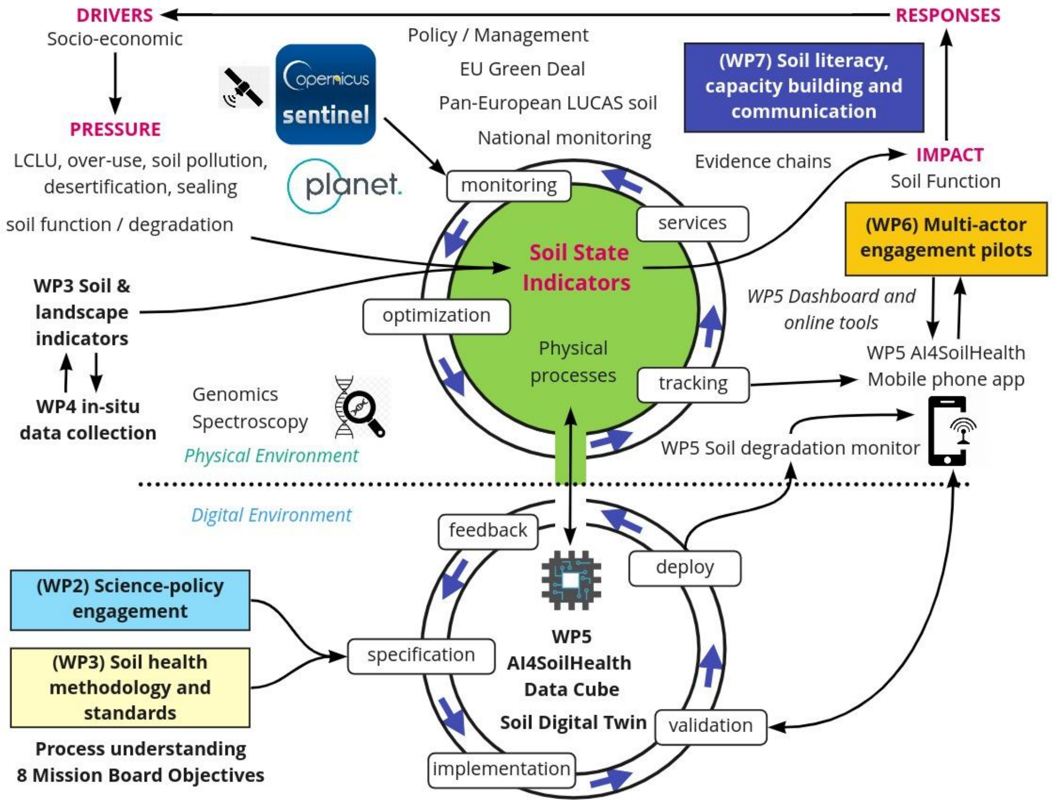

Fig. 1: AI4SoilHealth ultimately aims to prototype a Soil Digital Twin-type system for Europe.

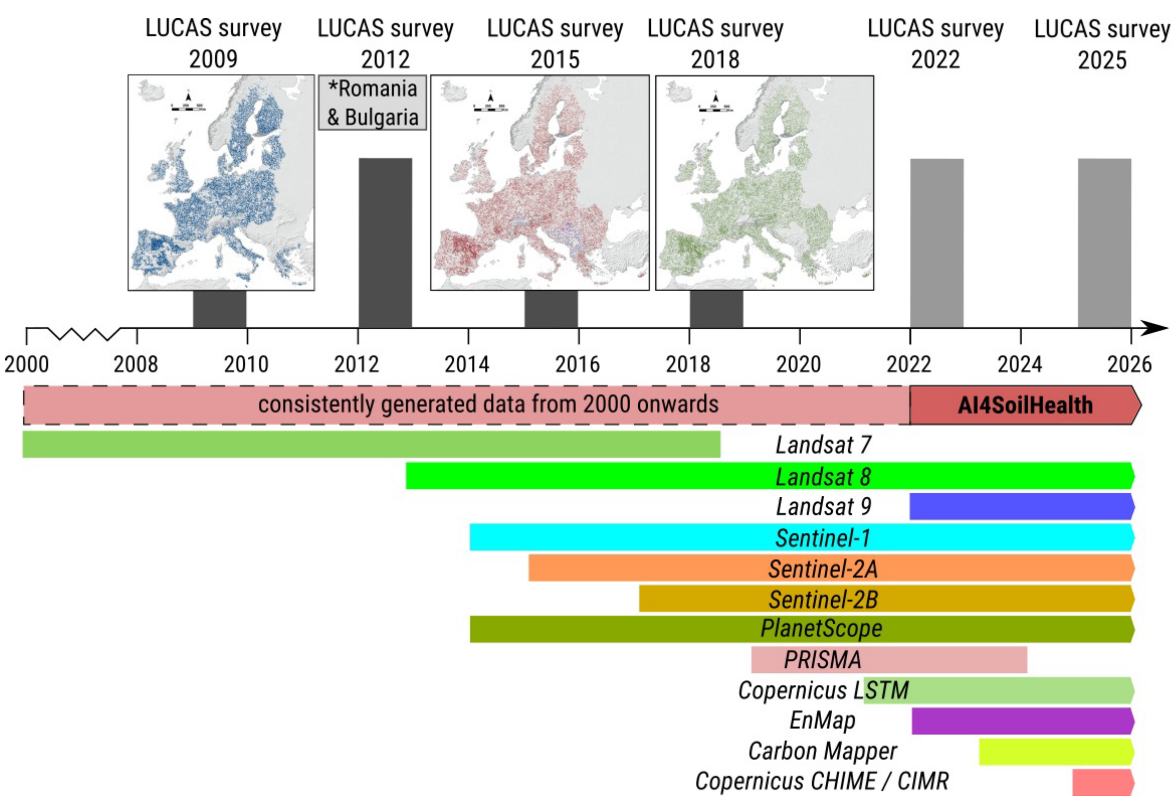

Fig. 2: LUCAS soil samples (Orgiazzi et al., 2018) and connected existing and upcoming EO missions.

This will leverage the latest advances in AI including deep learning to identify and map indicators of soil health at a pan-European scale. Moreover, our ambition is to interface this technology with stakeholder co-design and interoperability that informs decision making and empowers real soil managers at all levels. Consequently, AI4SoilHealth will be acting in a ***"lab-to-customer approach"***, where project outputs are readily applicable for external users.

What can you expect from this project?

To detect change over time, AI4SoilHealth will develop an **AI-based framework for a spatiotemporal EO system to model soil health indicators**. The model will be driven by satellite-derived inputs, and readily available environmental predictors (Fig. 3). Advances in EO will enable the development of satellite products that represent key inputs into soil models, examples being agro-climatic and land use/land management inputs (WP3). Combining AI and mechanistic modeling into hybrid modeling, a robust alternative to pure data-driven machine learning will be demonstrated. Mechanistic AI models incorporate the scientific knowledge that is encapsulated in mechanistic models into AI. This hybrid modeling approach has been introduced by a consortium partner (Reichstein *et al.*, 2019) and will be deployed for a **vertical SOC model with microbial and mineral interactions** (Ahrens *et al.*, 2020) (WP3).

The consortium will ensure that soil health indicators relying on remote sensing data, such as vegetation cover, NDVI / FAPAR, landscape heterogeneity and soil sealing, can be evaluated for different regions of interest and using different types of input data without loss of consistency. To this end, calculation methods will be developed (WP3) that are independent of scale and grid resolution.

	AI4SoilHealth	*LUCAS soil*
Data services	✔ Hundreds of layers at 30 m analysis-ready, complete & consistent	only 250 m resolution maps available
Data storage solution	✔ Cloud-solutions for data storage and sharing	On-demand from ESDAC
Real-time in-situ measurements	✔ Hand-held soil spectroscopy, based on FTIR / VisNIR instruments	none
Soil DNA analysis (biological indicators)	✔ Complete eDNA analysis	limited
Soil enzyme analysis	✔ Complete soil enzyme analysis	limited
Monitoring & Forecasting of ecosystem services	✔ Ecosystem services quantified & mapped	none
Data entry and retrieval on mobile devices	✔ Mobile phone App	none
Certificates and reports	✔ Automated generation of certificates / reports based on chatbots & AI solutions	none
Soil depth relationship	✔ Soils mapped in 3D (0 to 200 cm) with clear representation of soil-depth relationships	top-soil only (0 to 20 cm)

Fig 3: Target functionality in AI4SoilHealth vs LUCAS.

For landscape heterogeneity this poses a particular challenge, because higher resolution input data could easily suggest more landscape heterogeneity. Such effects will be mitigated by scalable calculation methods (WP5) that allow for a seamless transition between pan-European, regional and local soil health assessment.

Several components of the **AI4SoilHealth** methodology and technology have already been developed and tested by participants of the consortium in previous projects. Examples include indicator selection for national soil monitoring schemes (**EJP SOIL**), harmonization of LUCAS (Fig. 2) and national datasets, spatial modeling (Witjes *et al.*, 2022) and spatial decision support management practices such as the impact of bio-based fertilizers (**LEX4BIO**), and prediction of soil properties from soil spectroscopy (**Open**

Soil Spectral Library[5]). Furthermore, **AI4SoilHealth** will integrate the most up-to-date knowledge generated by other international and national projects.

Direct links to the **European Soil Partnership** and **Global Soil Partnership** will secure production of information applicable on multiple scales and purposes. Close collaboration with the **EJP Soil consortia** and **EU Soil Observatory** will be secured to enable alignment with the next generation of soil health monitoring (*e.g.* LUCAS 2025–26).

Project structure

Figs. 4 and 5 provide the general workflow for realizing our vision of AI4SoilHealth, describing the work packages (WP1–7) and some of the key links between them. **WP2** will develop the co-design of AI4SoilHealth with policymakers and stakeholders to determine their needs for the project output. These may include policy needs (planning, regulation, etc.), the appropriate spatial and temporal scales to inform policy, or available soil management options. **WP3** will use the input from WP2 to set the scope for soil health indicator selection and challenge criteria. Moreover, WP3 will assess and develop data, process and hybrid model products for input into WP5. **WP4** sits alongside WP3 focusing on development of new methods that can inform the derivation of new and future proxy indicators in WP3; it includes analyzing data for genomics and spectral measurements. The outputs of WP3, informed by WP4, provide the basis for the development of AI4SoilHealth in **WP5**. This WP will use the latest AI, including deep learning methods, to fuse and synthesize data and model outputs. Moreover, we will use advanced statistical methods, such as change point detection, to reveal where data and mechanistic model mismatches occur, uncovering new or unknown processes or sources of uncertainty; this will feed back into process model development in WP3.

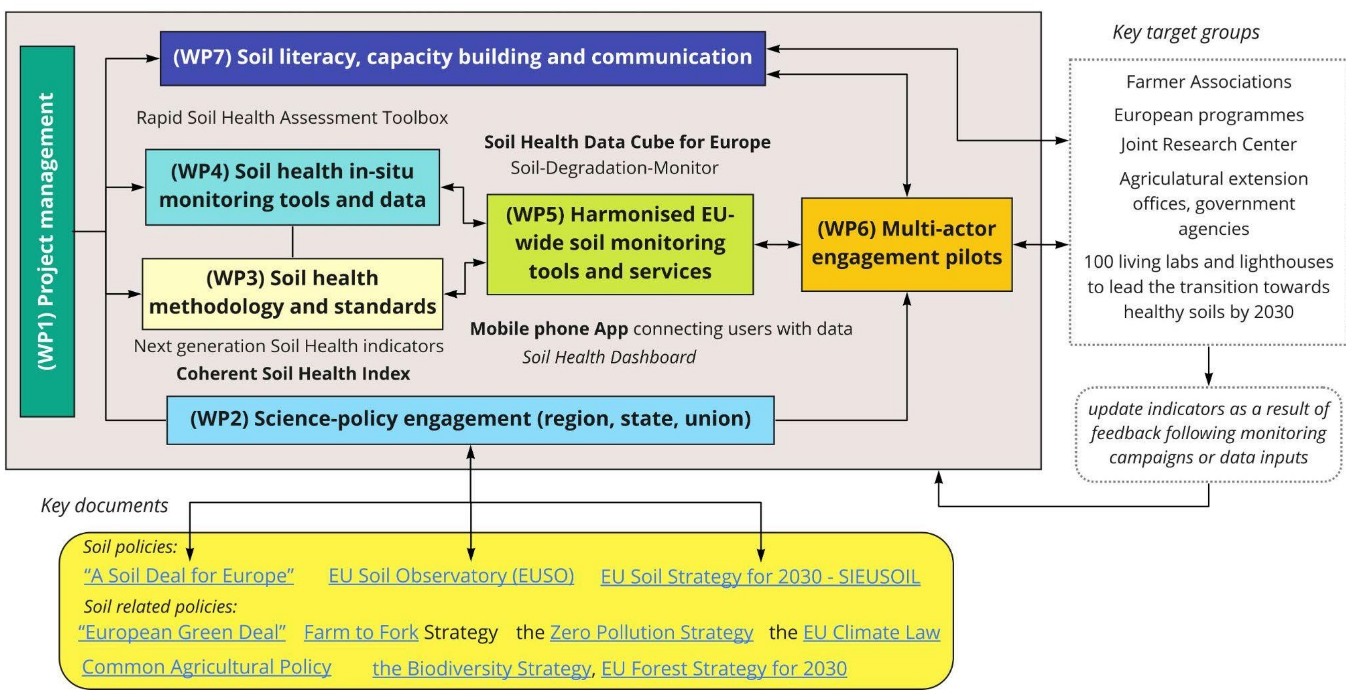

Fig 4: The project overview and linkages between work packages (WP) and key target groups.

5 https://SoilSpectroscopy.org

	Work Package Title	Lead #	Lead Name	PMs	Start month	End month	Duration
WP1	Project Management	1	Aarhus University	66	1	48	48
WP2	Science-policy engagement	22	ISINNOVA	55	1	48	48
WP3	Soil health methodology and standards	1	Aarhus University	316	1	48	48
WP4	Develop in-situ tools for observation and monitoring of soil health	10	Stockholm University	330	1	48	48
WP5	Harmonized EU-wide soil monitoring tools and services	13	OpenGeoHub	394	1	48	48
WP6	Multi-actor engagement pilots	15	Soil Association	340	1	48	48
WP7	Soil literacy, capacity building and communication	15	Soil Association	190	1	48	48
			Total	1691			

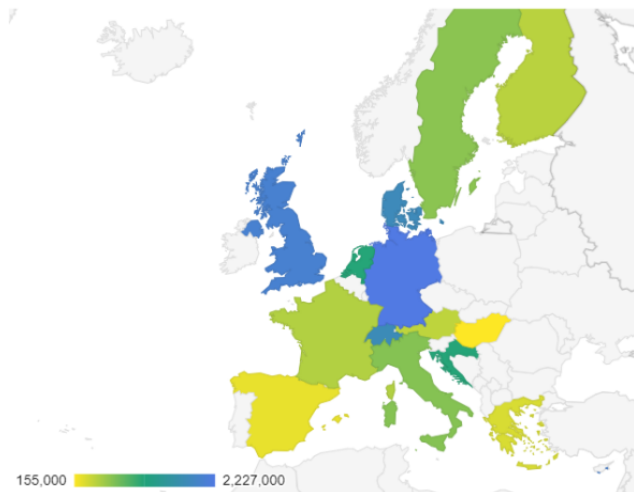

Fig 5: Work Packages with person-months (PM) and distribution of the project budget per country.

WP5 will also provide analysis-ready AI4SoilHealth predictions through a Soil Health Data Cube that will be served to users in **WP6**. These predictions will be compared to the unstructured data from the national pilot areas and living lab test sites to understand the impact of bias or methodological differences. This will help target data collection at living lab test sites, which will inform development for WP5 and WP3, and help them iteratively improve predictions and potentially **increase accuracy** and **usability of the produced soil health indicators** informing the design of soil health data collection strategies and methods. The outputs of WP4, WP5 and WP6 will aid in the co-design of dashboards and tools for policy-makers and soil managers who will be engaged in knowledge exchange and dissemination in **WP7**.

Consortium

Project lead: **Mogens Humlekrog** (greve@agro.au.dk)

Project partners: *Department of Agroecology, Aarhus University, OpenGeoHub foundation, UK Center for Ecology and Hydrology, Soil Association, Planet Labs, IIASA, Institute of Geo-Hydroinformatics at TUHH, University of Zagreb Faculty of Agriculture and Electrical Engineering and Computing, Stockholm University, Aalborg university, Thünen Institute of Climate-Smart Agriculture, Università degli studi Roma Tre, NEIKER, University of Aberdeen, Max Planck Institute for Biogeochemistry, Basel Universität, ETH Zürich, Bern University for Applied Sciences, Aristotle University of Thessaloniki, Natural Resources Institute Finland (Luke), INRAE, Sorbonne University, ISINNOVA, Institute for Soil Sciences at ATK, MultiOne, Digit Soil and Sustinn.*