Contribution ID: **13** Type: **Poster Presentation**

# Towards digital twin oriented Virtual Research Environment

LTER-LIFE is a large-scale research infrastructure in the making; it aims to provide a state-of-the-art e-infrastructure to study and predict how changes in climate and other human-induced pressures affect ecosystems and biodiversity. One of the grand challenges in ecology is to understand and predict how ecosystems are impacted by changes in environmental conditions, and external pressures. Predicting how complex ecosystems will behave under different scenarios requires combining empirical and observational data. To truly predict how ecosystems and landscapes will respond to current and future global change, we need to increase the availability of existing environmental long-term datasets, integrate disparate types of ecological data, and create a user-friendly and secure cloud-based digital modelling and simulation platform that can be used to link data to models and scenarios. This can be done by creating Digital Twins of entire ecosystems. A Digital Twin is "a digital replica of a living or non-living physical entity". Building Digital Twins of ecosystems has only recently become possible as Big Data, artificial intelligence (AI) applications and analytics, advanced computing infrastructure, and the FAIR principles have been developed and made available for ecology, ecosystem restoration, and biodiversity science. This approach brings together data scientists, informaticians, and ecologists in research communities around ecosystems. Detailed information on the LTER-LIFE project can be found at https://lter-life.nl/.

The Virtual Laboratories provide VRE components, data, models, and other Big Data tools (e.g., AI) and allow scientific users to discover, access, and integrate research assets for specific scientific purposes and manage the deployment, execution, and provenance of in silico experiments (analysis and simulations) on local or remote platforms. This includes the assets: 1) FAIR datasets of long-term studies of plants and animal populations as well as of environmental data, 2) FAIR models such as (a) process-based models that formalise ecological knowledge on abiotic and biotic relationships, or (b) data-driven models that explore relationships between data using classical statistics, or more contemporary artificial intelligent tools. These models themselves are made FAIR, so that they can be discovered, accessed, and reused in new scientific workflows. 3) Rules of interactions between assets (i.e., data and models), including solutions for scaling and enhancing interoperability of datasets, which are collected with different methodology. And 4) Tools for scenario studies to explore the effects of in situ local management strategies on ecosystem functioning.
These assets can be coupled in various ways (which is enabled by the LTER-LIFE platform) depending on the objectives of the specific research. Together, they will ultimately compose a Digital Twin of a specific ecosystem, which can then be used for scientific research or scenario studies to address societal issues.

With LTER-LIFE, we will offer a single point of entry for scientists studying entire ecosystems or components. The LTER-LIFE portal gives users tailored access to assets and facilities. LTER-LIFE develops: 1) Data bases and process models that are findable, accessible, interoperable, and reusable (FAIR, by applying (meta)data standards, controlled vocabularies, ontologies, and persistent identifiers). 2) Workflows that allow coupling of these data and models at the desired temporal and spatial scale, and that provide Big Data methodology to analyse these data (AI and machine learning approaches such as deep learning and active learning which enable new theory discovery and novel insights from high-dimensional sparse data spaces) as well as tools for uncertainty analysis, scenario studies and forecasting. 3) Virtual Labs that build upon infrastructure services and components of a VRE, allowing interdisciplinary communities of scientists to collaborate via a digital platform (including Big Data storage, cloud-based modelling, and dynamic simulations). And 4) Basic and advanced training in the use of the infrastructure.

LTER-LIFE will be an open-source infrastructure and users can either use the developed assets within their own local environment or run it in the cloud, facilitated by LTER-LIFE through our partner SURF. Setting up the LSRI LTER-LIFE as an open-source infrastructure aligns well with the current FAIR and Open Science developments. New data or models can be added as assets, which can then be used by other users. Workflows and data as used in projects will receive a digital identifier and will be stored by LTER-LIFE's partners DANS-KNAW and SURF.

**Primary authors:** ZHAO, Zhiming; Dr KISSLING, Daniel; Dr SOUREN, Astrid; Dr ATHANASIADIS, Ioannis; Dr PHILIPPART, Katja; Dr SOETAERT, Karline; Dr VISSER, Marcel; Dr BAKKER, Liesbeth

**Presenter:** ZHAO, Zhiming

**Track Classification:** Track 5: Virtual Research Environment (including tools, services, workflows, portals, ... etc.)