December 2024, No 1

GreenDIGIT[1] Project Introduction and Launch

Lowering the environmental impact of digital services and technologies tends to become a significant priority for both the operation of existing digital services and the design of future digital infrastructures. Energy consumption and carbon footprint are the two most talked about environmental impacts, and indeed, digital infrastructures today contribute 3 to 4% of the total greenhouse gas (GHG) emissions in the world, with a growth of 8% per year. In particular, the part of the networking infrastructure alone is responsible for 2 to 14% of digital impacts, according to various sources, mainly due to their electricity consumption.

*To keep Research Infrastructures at the forefront of scientific excellence, new technologies and solutions must be developed to reduce their environmental footprint, a challenge shared across all domains of society. GreenDIGIT brings together 4 major distributed Digital Research Infrastructures (RIs) at different lifecycle stages, **EGI, SLICES, SoBigData, EBRAINS,** to tackle the challenge of environmental impact reduction with the ambition to provide solutions that are reusable across the whole spectrum of digital services on the ESFRI landscape and play a role model.*

1. Objectives

- **Objective 1:** Assess status and trends of low impact computing within 4 DIGIT RIs (EGI, SLICES, SoBigData, EBRAINS) and in the broader digital service provider community of ESFRIs, to produce recommendations and roadmaps for providers for during and beyond the project.
- **Objective 2:** Provide reference architecture and design principles, as well as an actionable model for RIs about environmental impact assessment and monitoring, reflecting on the whole RI lifecycle and including the digital infrastructure components and their interaction with the broader environment.
- **Objective 3:** Develop and validate new and innovative technologies, methods, and tools for digital service providers within European RIs through which they can reduce their energy consumptions and overall environmental impact.
- **Objective 4:** Develop and provide technical tools for researchers, that assist them in the design, execution and sharing of environmental impact aware digital applications with reproducibility, Open Science and FAIR data management considerations.
- **Objective 5:** Educate and support digital service providers in the RI communities about good practices on environmental impact conscious lifecycle management and operation of infrastructures and services.

2. Concept and Approach

GreenDIGIT will address the challenge of achieving sustainability and reduced environmental impact of the future European RIs by addressing a complex of tasks which together will facilitate evolving the European Research Area as an ecosystem with reduced environmental and climate footprint, which impact is measurable and controllable.

Furthermore, GreenDIGIT will capture good practices and existing solutions and will develop new technologies and solutions for all aspects of the digital continuum: from service provisioning to monitoring, job scheduling, resources allocation, architecture, workload and Open Science practices, task execution, storage, and use of green energy, along with covering the entire continuum

[1] GreenDIGIT: [Green]er Future [DIGIT]al Research Infrastructures

of resources (IoT, RAN, Far-Edge, Edge, Cloud) typically/generally present/used in modern RIs. GreenDIGIT will deliver these solutions as building blocks, with a reference architecture and guidelines for RIs to lower their environmental footprint.

User-side tools and Virtual Research Environments will also be expanded with energy usage reporting and reproducibility capabilities to motivate users to apply low-energy practices. The new solutions will be validated through reference scientific use cases from diverse disciplines and will be promoted to providers and users through an active dissemination and training programme, in order to prepare the next generation of digital RIs with a low environmental footprint.

3. Overall Methodology

GreenDIGIT methodology includes several steps and interconnected activities required for the consistent systems design, development, operation, and management, which originated from the modern system engineering practices that include all system life-cycle stages and ensure continuous system evolution and improvement. GreenDIGIT will realise the following development stages important to consistently address the expected outcomes and scope:

- **Landscape analysis:** This stage includes the following two key tasks.
  - **Survey of Practices:** Assessing how digital Research Infrastructures (RIs) manage energy consumption, implement sustainability policies, and comply with local regulations.
  - **Methodology Development:** Creating a framework to evaluate RIs' energy efficiency and sustainability, based on best practices from industry and other sectors. Insights from this analysis will guide all subsequent project activities, including architecture design, technology development, policy formulation, and training.
- **Development of the Reference Architecture for the future sustainable and low impact RIs:** A comprehensive architecture for sustainable RIs in relation to energy efficiency and reduced environmental impact will be designed in this stage. By incorporating cutting-edge technologies and practices energy efficiency and environmental impact minimization are expected to be enhanced.
- **Specification of requirements:** Detailed requirements for sustainable RIs will be defined, covering both technical (e.g., metrics, monitoring) and non-technical (e.g., management, capacity building) aspects.
- **Tools for experimental research automation and research reproducibility:** Tools for automating and optimizing research workflows (e.g., Jupyter Notebooks, CWL) will be developed, integrating energy monitoring to support low-impact, reproducible research aligned with Open Science and FAIR principles. The project intends to put the research reproducibility in the foundation of future sustainable Open Science. The research reproducibility itself is a strong factor for reducing energy and resource consumption for recreating and re-running experiments. GreenDIGIT will extend existing research reproducibility approaches with effective and resource/energy aware data management and sharing, enabling FAIR data principles with energy, durability and sustainability awareness.
- **Development of technology enablers, tools and applications:** Solutions for optimizing energy consumption across the RI continuum, from IoT and 5G networks to datacenters and cloud infrastructures, will be developed. These components will feature standard APIs for seamless integration with existing scientific workflow tools.
- **Prototyping, deployment and evaluation:** The developed solutions will be tested in real RI/datacenter environments, measuring their effectiveness and refining the architecture. Open-source code and implementation guidelines will be provided to promote wide adoption.
- **Policy recommendations and Actions plan** are important factors to drive wide adoption of the proposed solutions and facilitate RIs evolution to environmental sustainability and reduced impact.
- **Self-assessment and Certification.** GreenDIGIT will develop tools for self-assessment and certification in the form of structured questionnaires to facilitate consistent implementation of sustainable, energy efficiency and impact lowering practices and technical solutions. The tools tested and verified in the project will be developed further to create a certification framework that can be used for assessment and potentially audit of the RI environmental sustainability, environmental impact and GHG footprint.

4. Project Structure

GreenDIGIT project includes 11 Work Packages that are organised in two development periods M1-M18 and M19-M36. Figure 1 illustrates the Work Packages interaction that helps to understand the project concept and perspective.

WP3 RI Survey and Landscape Analysis (including existing standards and regulatory documents) provides the basis for both policy and technical development in the project. WP3 outcome feeds into WP4 Architecture, WP5 Tools, WP6 Technology Enablers, and WP8 Policy development.

WP4 Architecture and Requirements for Sustain-

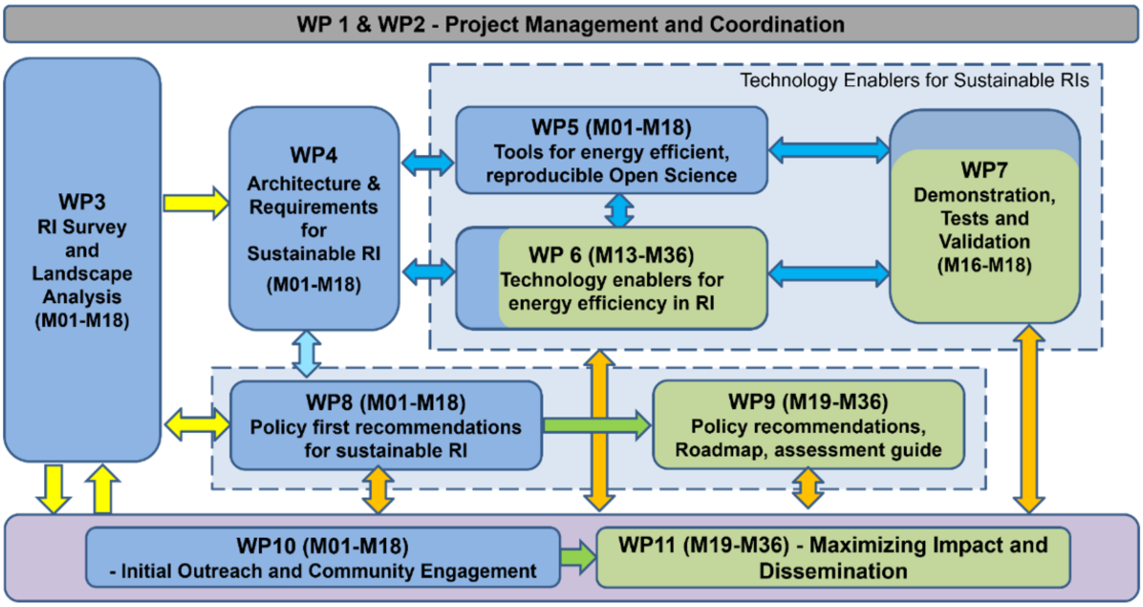

Figure 1: GreenDIGIT project structure showing WPs interaction and timeline: 1st period – blue; 2nd period – green.

able RI performs in tight coordination with WP5 that ensures consistent development of tools for energy efficient, reproducible Open Science and Virtual Research Environment (VRE) which allows sustainability aware scientific workflow management and deployment on the generally distributed RI or data centers.

Technical outcome and sustainability management services and applications are delivered in in WP6, WP7. Policy recommendations and Roadmap for future RI environmental sustainability are developed in WP9.

5. GreenDIGIT Launch and Progress

The GreenDIGIT project at the current stage has established well-coordinated activity of all Work Packages and Tasks to progress with defined Objectives, starting from landscape analysis of the ESFRI Ris, extensive State of the Art (SotA) and current standards and regulation base widely used by data centers, digital and RI operators.

WP3 conducted a survey analysis among digital RIs to review the current landscape of existing practices, metrics, tools, and identified needs and gaps in addressing environmental sustainability in RI operations. The results of this study provide an important landscape context for the environmental impact assessment methodology and guidelines developed in WP3, as well as other Work Packages within GreenDIGIT, helping to identify opportunities for change, potential barriers, and best practices for broader adoption. WP3 is progressing to build upon the survey findings to deliver a methodology to monitor, assess, and quantify the environmental impact of RIs.

WP3 landscape analysis has been supported by necessary desk research to review a range of standards and regulations used to manage and assess the energy efficiency and environmental impact of modern data centres, IoT and edge-to-cloud infrastructure components, and scientific applications. This includes the following groups of International standards: ISO 50001, ISO 50002, ISO 30134, ISO 14001, related ITU-T standards, European standards EN 50600 group, and EU Code of Conduct on Data Centre Energy Efficiency.

WP4 is progressing with the coordinated development of the architecture design patterns for future energy efficient and environmentally sustainable RIs. The development is based on the systematic State-of-the-Art analysis covering best practices for digital infrastructures and software engineering for scientific applications. The proposed Shared Responsibility Model for sustainability is described below. WP4 is in the process of defining design principles for Sustainability by Design that would provide a basis for designing, developing and operating the main functional components of the modern/future digital RI to ensure the reduction of the environmental impact of the whole RI ecosystem serving specific scientific domains.

WP4 works tightly with WP5 in designing the environmentally aware Virtual Research Environment (VRE) and researcher tools that would allow researchers to monitor, control and optimise the environmental impact of RIs and scientific applications. It covers federated data management infrastructures for publishing, storing, searching and indexing reproducible research artifacts. There is ongoing work on improving reproducibility of digital experiments to avoid unnecessary re-iteration which coaligns with the goal of reproducibility, repeatability and replicability. Consumed experiment resources are included in suitable meta data formats. Those are prominently available to researchers making them aware of environmen-

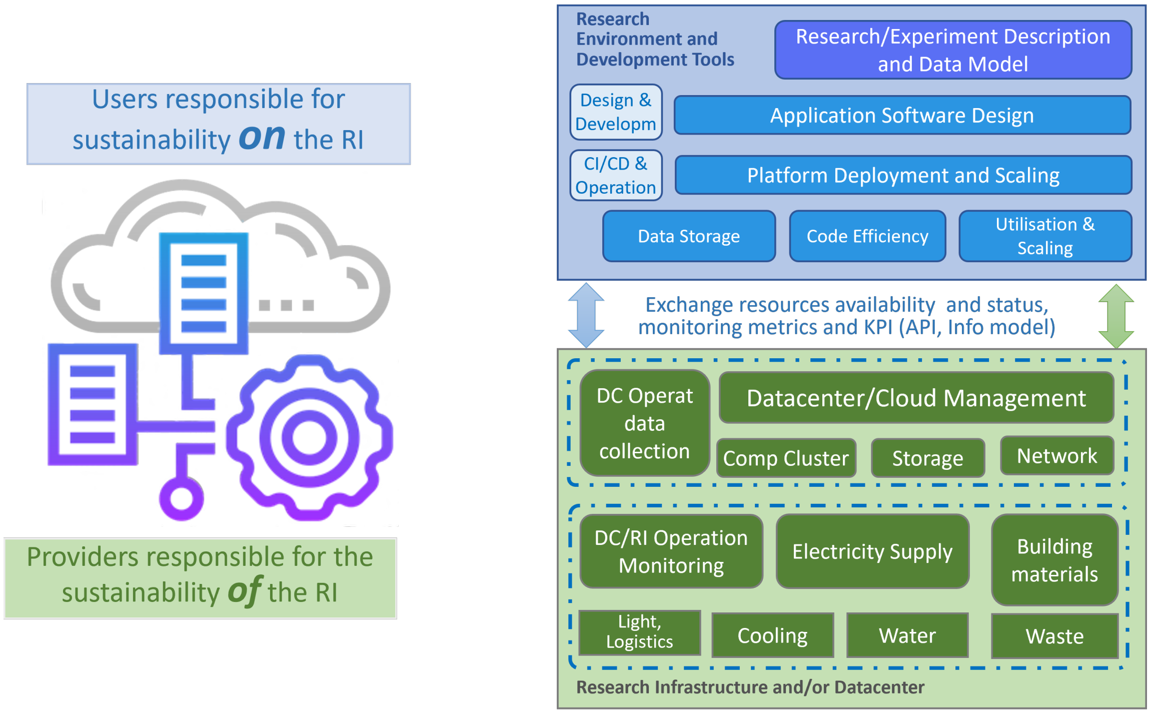

Figure 2: Shared Responsibility Model for RI Sustainability: Sustainability of the RI and Sustainability on the RI (for research applications)

tal footprint. The researcher tooling includes web, command line, API, and Jupyter Notebook interfaces. With the collaboration of European partners from diverse backgrounds, the way researchers experience the work with digital Ris is improved and sustainable, reproducible, and open science practices are introduced.

6. Shared Responsibility Model for Sustainability

Based on the standards and regulations study, supported by the European RI landscape study, the GreenDIGIT project identified a set of requirements and practices related to both RI operators and researchers that need to be brought in compliance and practical implementation. This is defined in the proposed Shared Responsibility Model for achieving sustainable digital RI operation. It clearly defines the roles and responsibilities of both the RI operator and the researcher in reducing the environmental impact and promoting energy-efficient practices. This model is essential for ensuring that both infrastructure management and research practices contribute to environmental sustainability goals. Figure 2 below illustrates the separate responsibilities of the RI provider or operator and the responsibilities of the researcher or research project together with the functional elements that need to support necessary technical solutions for the proposed Shared Responsibility Model.

The following is a suggested breakdown of RI operator and Researcher responsibilities.

RI Operator Responsibilities:

- Energy-efficient infrastructure: The operator is responsible for the underlying infrastructure, including implementing energy-saving technologies such as energy-efficient cooling, virtualization, and workload management systems. They must ensure the infrastructure meets international sustainability standards and provides researchers with tools to monitor their energy use.
- Resource allocation and scheduling: The operator must offer energy-aware resource allocation systems that dynamically adjust resources according to energy availability, and ensure that computational and data storage tasks are allocated in a way that maximizes energy efficiency.
- Monitoring and feedback: RI operators should provide real-time feedback on energy consumption, offer automated suggestions for energy optimization, and make energy-saving features easily accessible to users.
- Use of renewable energy: Operators should maximize the use of renewable energy sources, communicate availability, and ensure that the infrastructure can efficiently switch between energy sources.

Researcher Responsibilities:

- Energy-efficient workflows: Researchers must optimize their workflows by taking advantage of energy-saving features provided by the RI. This includes optimizing data transfers, minimizing resource-intensive operations, and scheduling jobs to align with off-peak energy usage or renewable energy windows.

- Data Management: Researchers should actively manage data storage to minimize energy usage. This includes selecting storage options from different types of data, deleting unnecessary datasets, compressing data, and using energy-efficient storage options.
- Adopt best practices: Researchers need to stay informed about sustainability best practices and incorporate them into their daily workflows, following recommendations and feedback provided by the DRI operator.

Training and Awareness on best practices and technical solutions for environmental sustainability is important for both the operators and researchers. Training and guidelines for energy-efficient practices should facilitate the creation of a culture of sustainability within the research community.

To implement a workable shared responsibility model for sustainable digital RI operation, the architecture design must include solutions that support real-time monitoring, energy-efficient resource scheduling, feedback systems, sustainable data management, and collaboration tools. These components must empower both the RI operator to manage the infrastructure efficiently and the researcher to optimize their workflows with sustainability in mind. The integration of renewable energy, training, and compliance tracking ensures that all parties are aligned in achieving energy and resource efficiency, contributing to a truly sustainable research ecosystem.