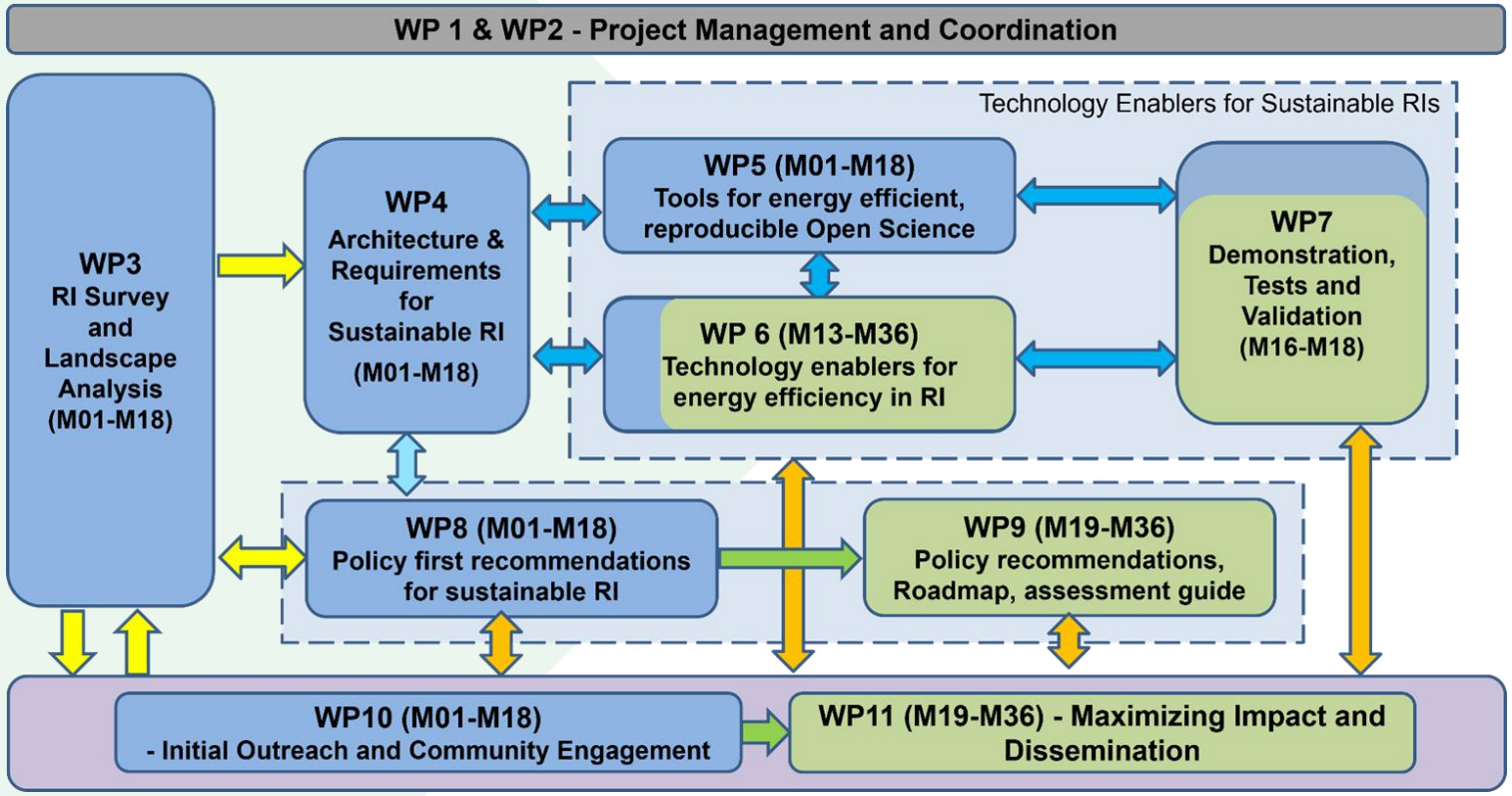
GreenDIGIT
Recent Architecture Developments
For discussion

Yuri Demchenko, UvA

WP4-WP5-WP6 Design Workshop

23-24 January 2925

GreenDIGIT Workpackages Interaction



General structure

- WP3 – Survey and Landscape analysis (among ESFRI RIs)
- WP4, WP5 – Technical development and foundation
- WP6, WP7 - Implementation and demo
- WP8, WP9 – Policy recommendations, (self-)assessment guide, roadmap
- WP10, WP11 – Dissemination, outreach, branding & Training

Goals of this Workshop

(4) General/Conceptual

- Verify, revise, extend Architecture: Shared Responsibility Model (T4.2/D4.2), VRE&FDMI (T5.1, T5.3) and Researcher tools (T5.2)
- RI Lifecycle (RILC) and Lifecycle Analysis (LCA) (T4.3 ⇔ T4.2, T3.2/T3.3)

(1) Datacenter/Operator Group

- Reach consensus on infrastructure metrics and implementation (T4.4, T6.1)
- Clarity on brokering algorithms (T6.3, T6.4), frameworks (T7.1), power grid support (T6.5), validations (T7.2, T7.4)
- Establish connection with IoT optimization algorithms (T6.2)

Researcher Development Environment (RDE)

- (2) Researcher interactive Energy Efficiency assessment tools (T5.1, T5.2, T4.2+, T4.4?)
  - Information model and API to: RI/datacenter & external data (geo/energy, Vendor/benchmarks)
  - Plugins/Libraries for RDE and Jupyter Notebooks (also Workflow Management Systems)
- (3) Understand initial design for federated infrastructure and reproducible platform design (T5.3, T7.3)

Joint WP4-WP5-WP6 work planning

- DevOps environment and work style - scrum/Sprints/Jira?, Kanban/Trello?

Recommended Outcome/Recommendations (1)

- Metrics to be defined to include: Measurement, Usage, CO2 footprint
  - Decide on/specify information granularity (e.g. for power for what region, unit?)
- What sources for available information to use? What to measure on our own facilities?
- Metrics&Monit/Reporting (MMR): Consider sharing experience and code by IFCA/CSIC on tools and implementation, Home Assistant (?)
- (MMR): Provide information, experience and recommendations with/on using Scaphandre
- (MMR): Decide on the first/initial stage of implementation, planning for the next stages
- Consider/discuss architecture templates/patterns for cloud by SZTAKI (OpenStack) – e.g. Dashboard for sustainability monitoring and deployment control
- Implement elements of the interoperability of items in monitoring, reporting, control via common info model, schema, semantics, registry
  - Possibly start with GOCDB, tools used for monitoring -> schema, API, JSON
  - Factor of reproducibility
- Information/data model in EGI accounting system – ready or to be implemented?

Recommended Outcome/Recommendations (2)

- Architecture: Consider development/vision on network infra continuity and optimisation of energy efficiency – check with 6Green project
  - Stakeholder shared responsibility and cooperation/awareness based on Service Agreement/User awareness on migration/optimisation – Extend biz type of stakeholder (binding) agreement/relation to user/tenant to be concise
  - Backpressure in sustainability: stakeholders and owners - Exposure level to tenants

T4.2 Architecture Development

- Architecture related discussions

GreenDIGIT Architecture Definition Methodology

Why we need Architecture?

- Architecture is a way to coordinate and establish a common language between: Operators, Developers, Researchers/Users, Policymakers
- A basis for linking standards and regulations to functional architecture components

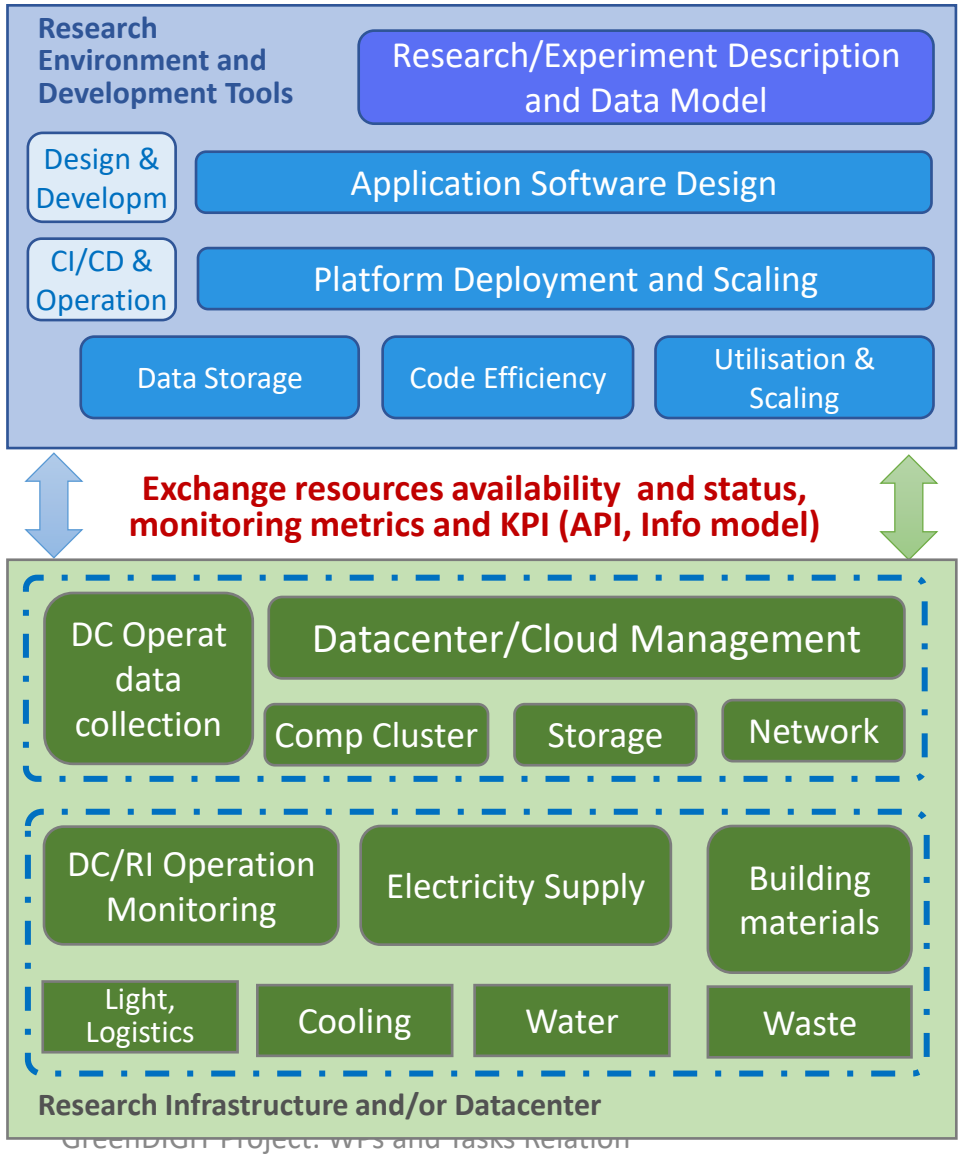
- General view on the RI Ecosystem Optimisation and Green IT
  - Horizontal, vertical, lifecycle
  - RI Operators and Researchers
  - RI continuum: (Research Object) – Sensor - (RAN) – Edge – Cloud – Workflow - Researcher
- Sustainable architecture design principles
  - As a basis for modelling and metrics for RI infrastructure operation and optimisation
  - **Shared Responsibility Model**: RI provider/operator and Researchers/Projects
  - **Sustainability by design** – A novel concept to be introduced addressing different aspects and stages
- Linkage with existing Standards and Regulations to ensure Sustainable Architecture Design principles support compliance with the standards, regulations and audit
  - To provide the opportunity for RI/datacenter operation (and design) optimisation (through the whole lifecycle)
- System Engineering and Design (thinking) approach in Green research and technologies

Shared Responsibility in Sustainability – Reflecting Operational and Management Aspects and Roles

GreenDigit

Users responsible for sustainability ***on*** the RI

Providers responsible for the sustainability ***of*** the RI



Standards and regulations
Software Development Quality and Design Patterns

Project/Researcher Responsibility:
Applications Development, Deployment, Operation, Energy usage and KPI monitoring

Sustainability by Design

Provider/Operator Responsibility:
Research Infrastructure or Datacenter, Monitoring Energy and environmental impact metrics and KPI

Standards and regulations
Datacenter and RI Building and Operation

GreenDIGIT Project: Novel and Consolidating Approach
Shared Responsibility in Sustainability and Sustainability by Design

GreenDigit

Users responsible for sustainability ***on*** the RI

Providers responsible for the sustainability ***of*** the RI

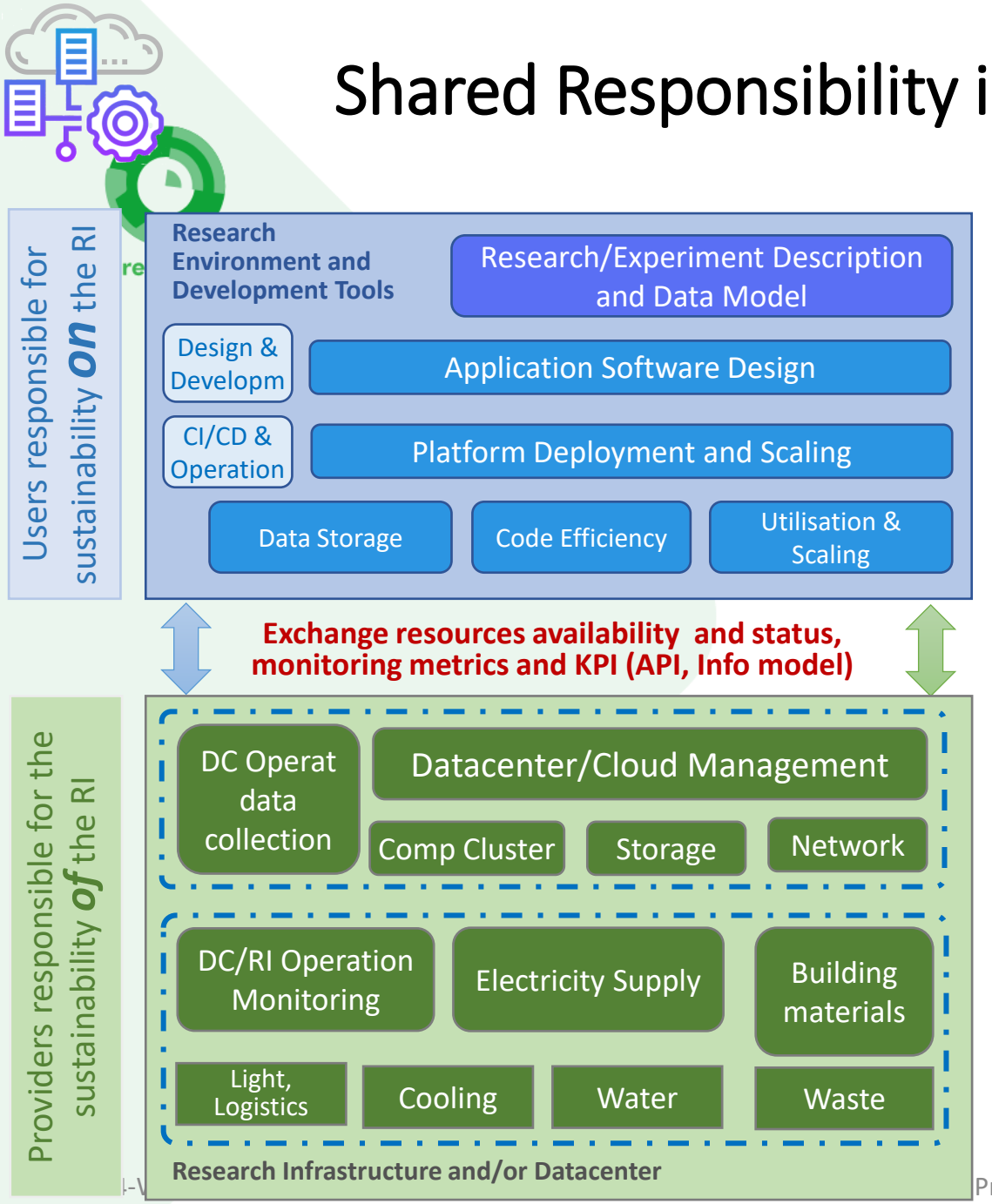
Researcher/ Project Responsibility:
Applications Development, Energy usage and KPI monitoring

Sustainability by Design Challenge

RI/Datacenter Provider/Operator Responsibility:
Monitoring Energy and environmental impact, metrics and KPI

- **Architecture for Sustainability by Design**
  - Functional components, layers, API, Requirements
- **Software and application components that can be optimised during design and controlled during operation**
  - Green aware API including necessary energy, performance, environment information
- **(!) Common information/data model and metadata (naming)**
  - Including Requirements, KPI, Metrics
  - Create a basis for reproducibility
  - + FAIR for Sustainability
- **RI and applications lifecycle**
  - RI lifecycle stages (concept, design, development, deployment, operation, decommissioning) and scientific workflow and research data

Shared Responsibility in Sustainability – Roles and Actors



Project/Researcher Responsibility:
- Applications Development,
- Deployment, Jobs submission & Energy Efficiency optimisation
- Operation, Workflow execution
- Energy usage and KPI monitoring

Experimenter/Researcher Demand:
- Access to datacenter (internal) monitoring data and custom configuration
- Energy usage and KPI & metrics monitoring

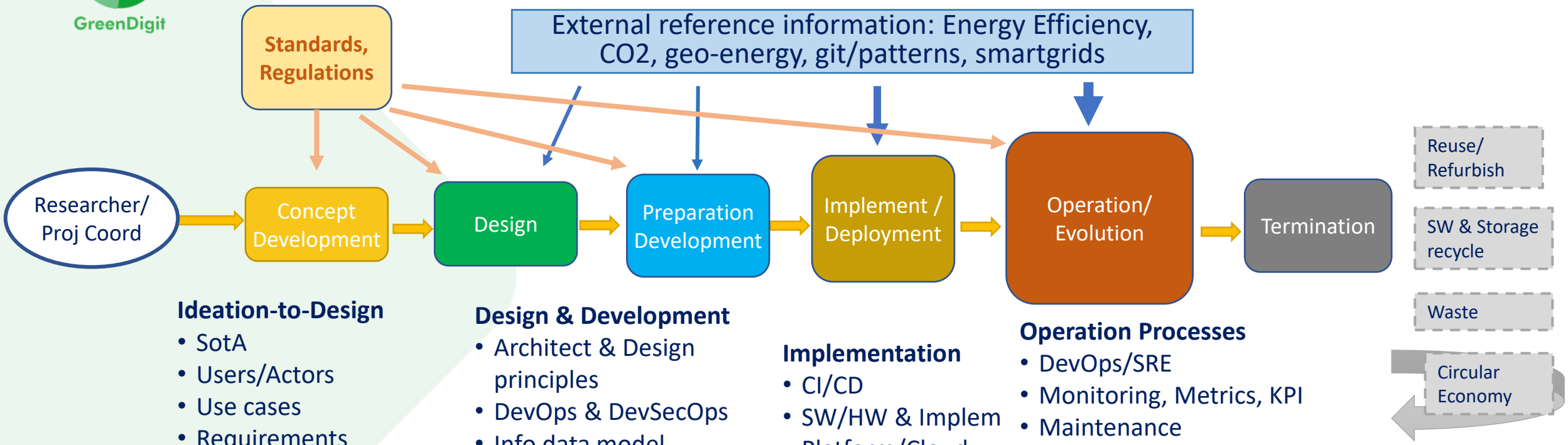
Provider/Operator Responsibility:
- Operation of Research Infrastructure
- or Datacenter,
- Monitoring Energy and environmental impact metrics and KPI
- Waste
- Lifecycle and evolution

Actors:

Operator group
- Facility owner
- Datacenter operator
- RI (overlay) operator

User group
- Project/Group user
- Researcher/ Scientific User
- Researcher/ Experimenter on RI or datacenter workflow

RI Lifecycle Stages, Activities and Factors



Ideation-to-Design
- SotA
- Users/Actors
- Use cases
- Requirements Engineering
- Architecture
- Standards/Regulations

Design & Development
- Architect & Design principles
- DevOps & DevSecOps
- Info data model
- Git, patterns, metrics, schemas, Repository
- API/WebAPI
- Testing, Logbook
- Project Management
- Tech Documentation
- Report, Compliance

Implementation
- CI/CD
- SW/HW & Implem Platform/Cloud
- Feedback

Operation Processes
- DevOps/SRE
- Monitoring, Metrics, KPI
- Maintenance
- Audit, Certification
- Electricity, Water, Buildings

Research Infrastructure Lifecycle
- All stages and processes to be In compliance with standards and regulations
- Continuous monitoring
- Shared Responsibility model and User /operstor tools

Environmental Sustainability for RI

- Standards and best practices
  - ISO, ITU-T, EN, EU JRC Datacenter Code of Conduct
  - Recent EC Delegated Regulation (EU) 2024/1364 of 14 March 2024 – Mandatory reporting since 15 Sept 2024
- KPI and metrics for energy efficiency and environmental impact/ sustainability

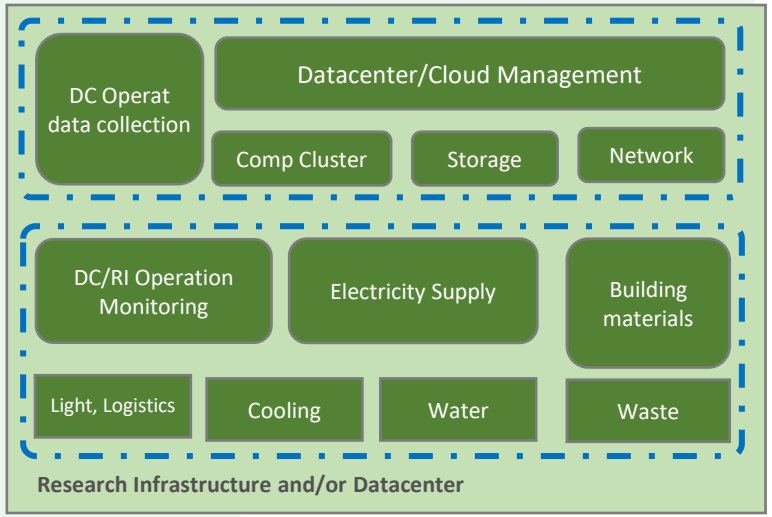
Linking KPI, Metrics and Design Patterns – To be clarified and extended

KPI (Key Performance Indicators)

- Energy consumption (kWh).
- PUE (Power Usage Effectiveness): Ratio total energy used to energy consumed by IT
- Carbon footprint: Amount of CO_2 emissions associated with energy usage.
- Uptime or reliability: How well the infrastructure maintains consistent service
- Resource utilization efficiency: Resources (servers, storage, etc.) usage w/o wasting energy.

Metrics

- CPU usage and load: The utilization levels of computing resources.
- Energy draw per server.
- IT equipment energy consumption
- Temperature: The internal temperature of servers or cooling systems.
- Cooling system efficiency



Design Solutions/Patterns

Under RI provider control/management

- Energy-efficient hardware
- Green energy sources.
- Monitoring and energy efficiency analytics and optimisation at the level of RI provider
- Advanced cooling systems
- Smart grid integration

Under user/developer control

- Energy and environment aware applications development
- Modelling, simulation and testing as part of the development process
- Green software practices and templates
- Virtualization and containerization

Linking KPI, Metrics and Design Patterns – To be clarified and extended

KPI (Key Performance Indicators)

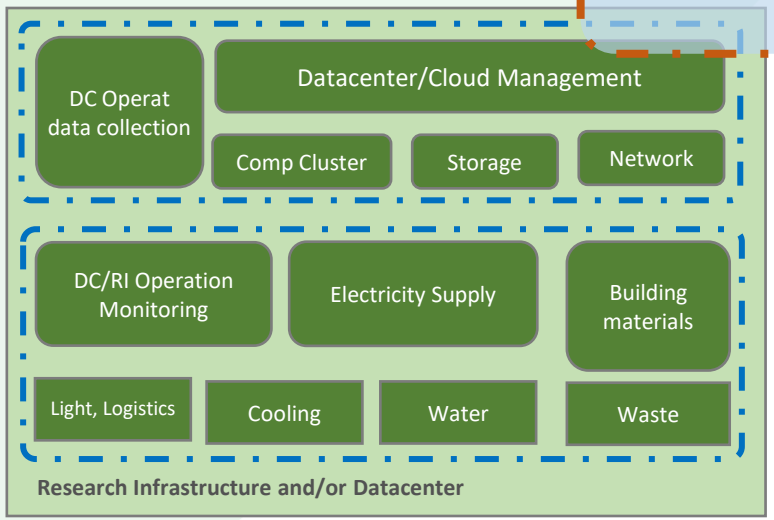
- Energy consumption (kWh).
- PUE (Power Usage Effectiveness): Ratio total energy used to energy consumed by IT
- Carbon footprint: Amount of CO_2 emissions associated with energy usage
- Uptime or reliability: How well the infrastructure maintains consistent service
- Resource utilization efficiency: Resources (servers, storage, etc.) usage w/o wasting energy.

Metrics

- CPU usage and load: The utilization levels of computing resources
- Energy draw per server.
- Network energy consumption
- Temperature: The internal temperature of servers or cooling systems.
- Cooling system efficiency

Based on the previous EGI/partners work to develop monitoring and metrics for EGI infrastructure of datacenters

- **ACE project and EGI survey**
- **RI landscape survey**
- **Standardisation and audit requirements**



Design Solutions/Patterns

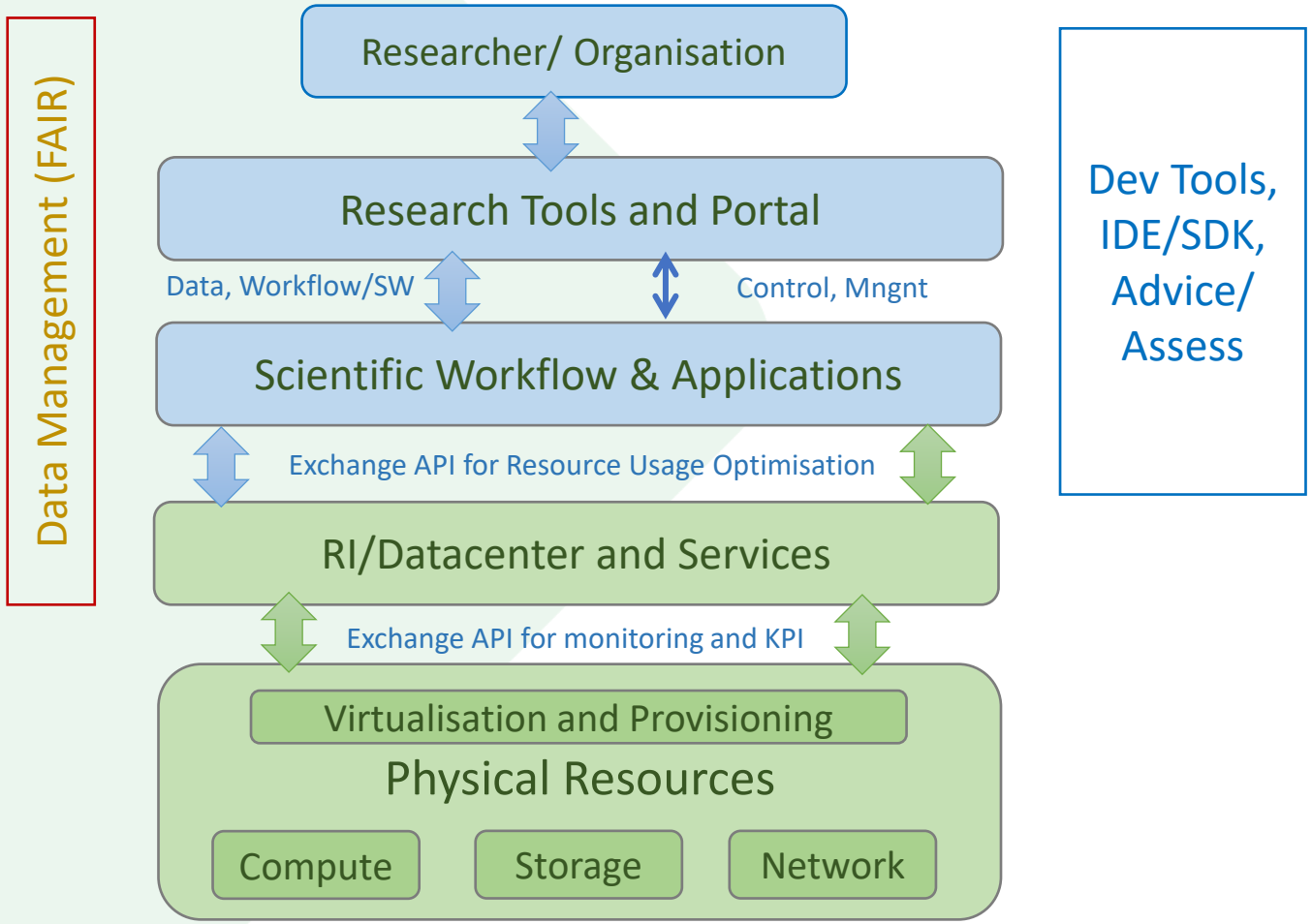
Under RI provider control/management

- Energy-efficient hardware
- Green energy sources.
- Monitoring and energy efficiency analytics and optimisation at the level of RI provider
- Advanced cooling systems
- Smart grid integration

Under user/developer control

- Energy and environment aware applications development
- Modelling, simulation and testing as part of the development process
- Green software practices and templates
- Virtualization and containerization

RI Sustainability by Design Components/Aspects: Motivated by the Shared Responsibility Model



- **Architecture for Sustainability by Design**
  - Functional components, layers, API, Requirements
- **Software and application components that can be optimised during design and controlled during operation**
  - Green aware API including necessary energy, performance, environment information
- **(!) Common information/data model and metadata (naming)**
  - Including Requirements, KPI, Metrics
  - + FAIR for Sustainability
- **RI and applications lifecycle**
  - RI lifecycle stages (concept, design, development, deployment, operation, decommissioning) and scientific workflow and research data

GreenDIGIT RI Continuum: Stages, Zones, Tiers, Layers

- Workflow
- Compute: IoT – Edge (Far - Near) – Cloud – Storage/Visualise
- CPU/GPU – VM – Container - Serverless
- Storage: RAM – Fast - Object – Archive – Gleisure/Gold
- Network: RAN – WiFi – LAN – Fast Network - Backplane

Multi-factor model for Energy Efficiency and Performance Optimisation