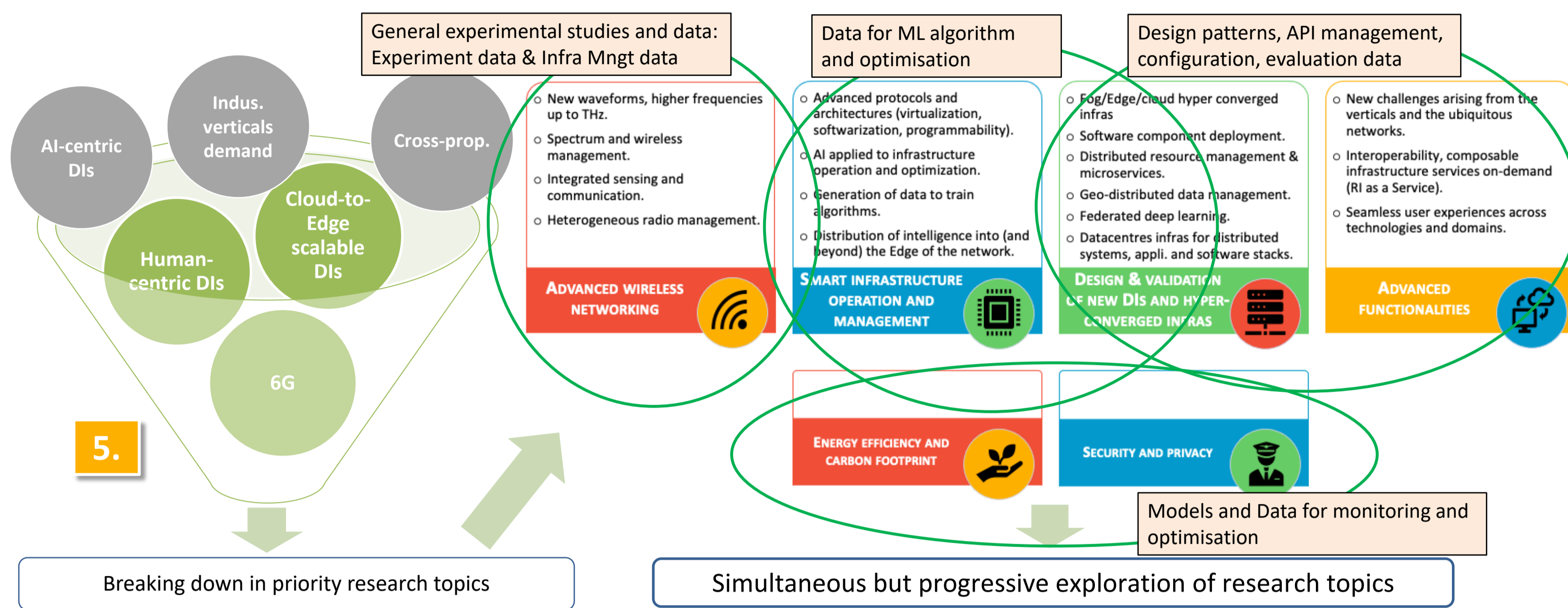


SLICES Data Management Infrastructure for Reproducible Experimental Research on Digital Technologies

Yuri Demchenko, Paola Grosso, Shashank Shrestha (University of Amsterdam)

Different Types of Data for Different Experimental Studies



Data Types in SLICES-RI Experimental Studies

General experimental studies and data documentation and publication

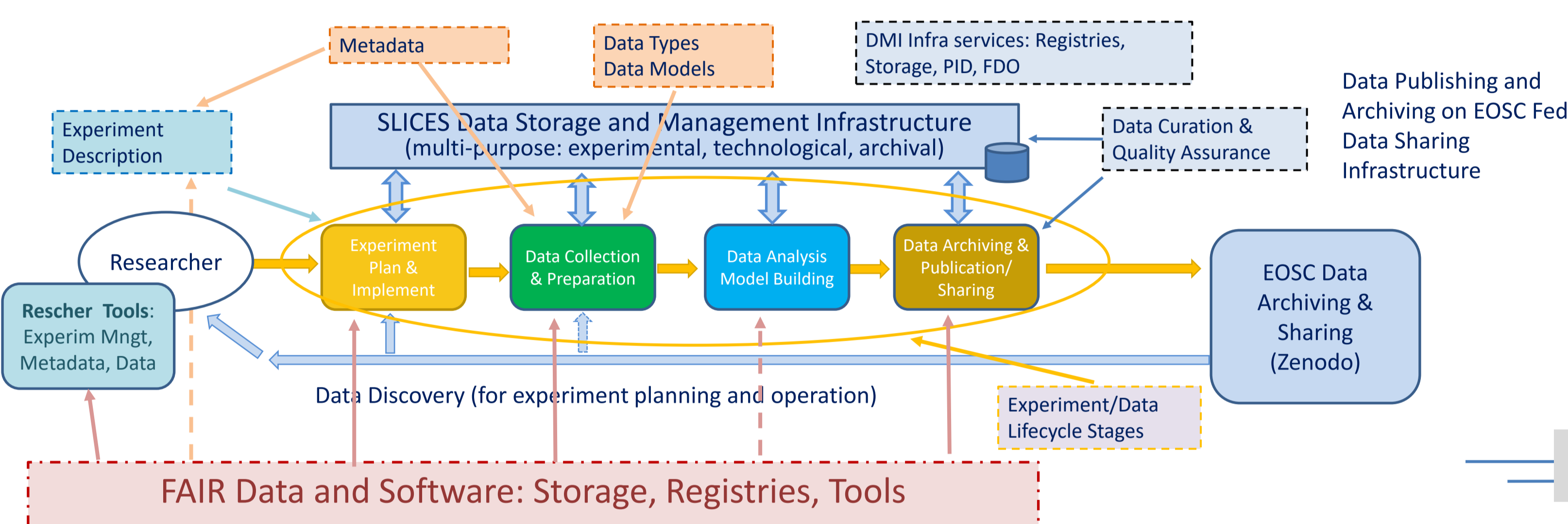
- FAIR (Findable, Accessible, Interoperable, Reusable) data principles are key for experimental data sharing
 - Metadata profiles to be defined for major types of experiments and supported by data and metadata management tools
 - Infrastructure management information to be recorded as experiments environment
 - Research Object (RO) and RO-Crate archiving template
 - FAIR Digital Object (FDO) and PID infrastructure
 - SLICES FDO profile (SFDO) and Metadata profile
- Data produced for AI/ML algorithms training for smart infrastructure optimisation and management (including energy efficiency, performance, resilience, sustainability)
- Data modelling and data lineage (staging documenting)
 - AI/ML models serialization and portability

New Digital Infrastructure architecture elements and design patterns

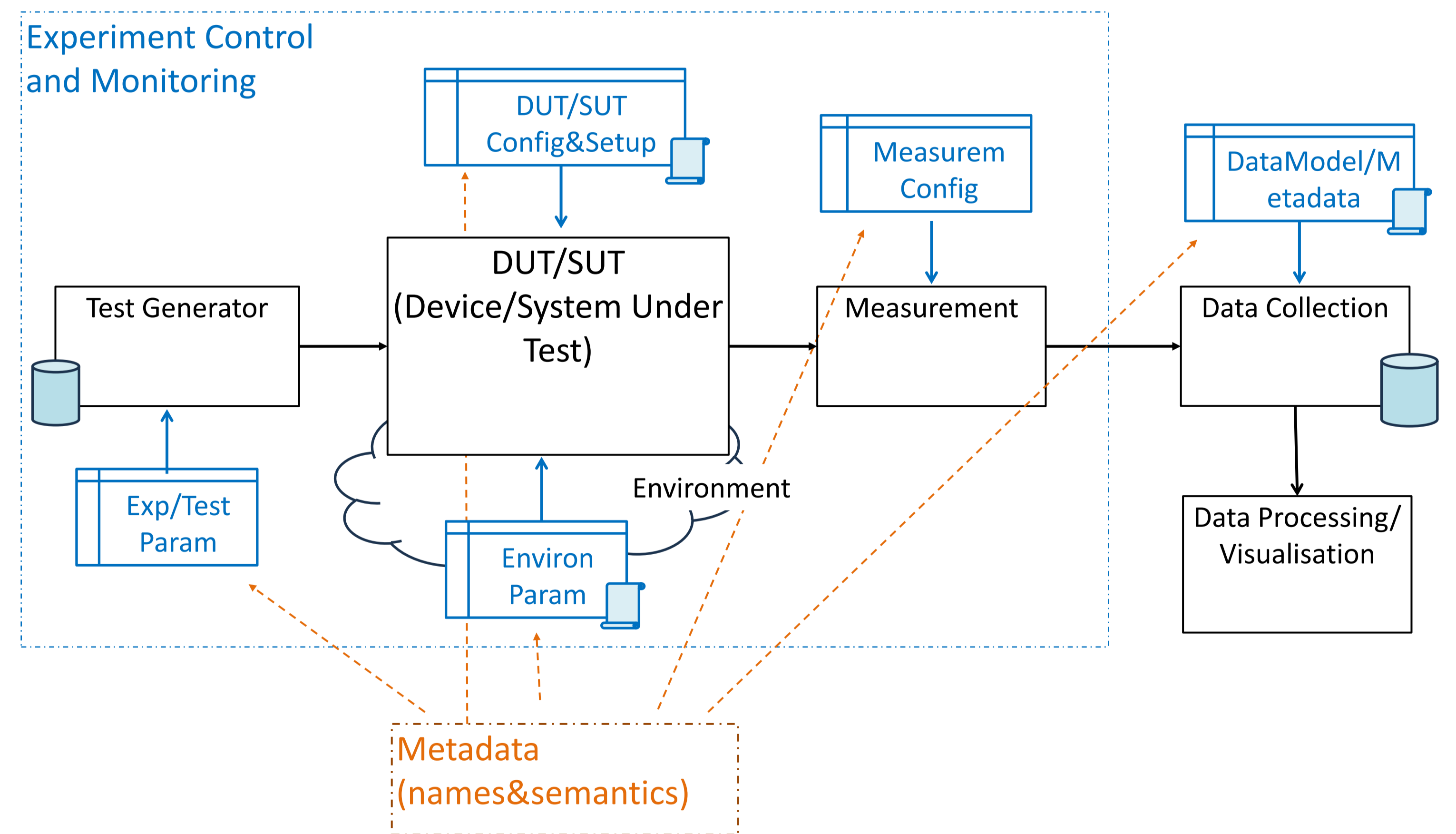
- Infrastructure and design patterns
- Metadata for API description, identification, composability

Focus on **repeatability** and **reproducibility** with the future support of **replicability**

SLICES Experimental Data Lifecycle and Research Workflow



Generic Experiment Model for Reproducibility



Experiment Description and Data Modeling: Metadata Requirements

SLICES Experimental Data/Metadata groups

- Architecture and services
- General Metadata definition and management
- Experiment description and metadata
- Domain specific (e.g. SLICES Blueprint Architecture)
- Metadata Management tools

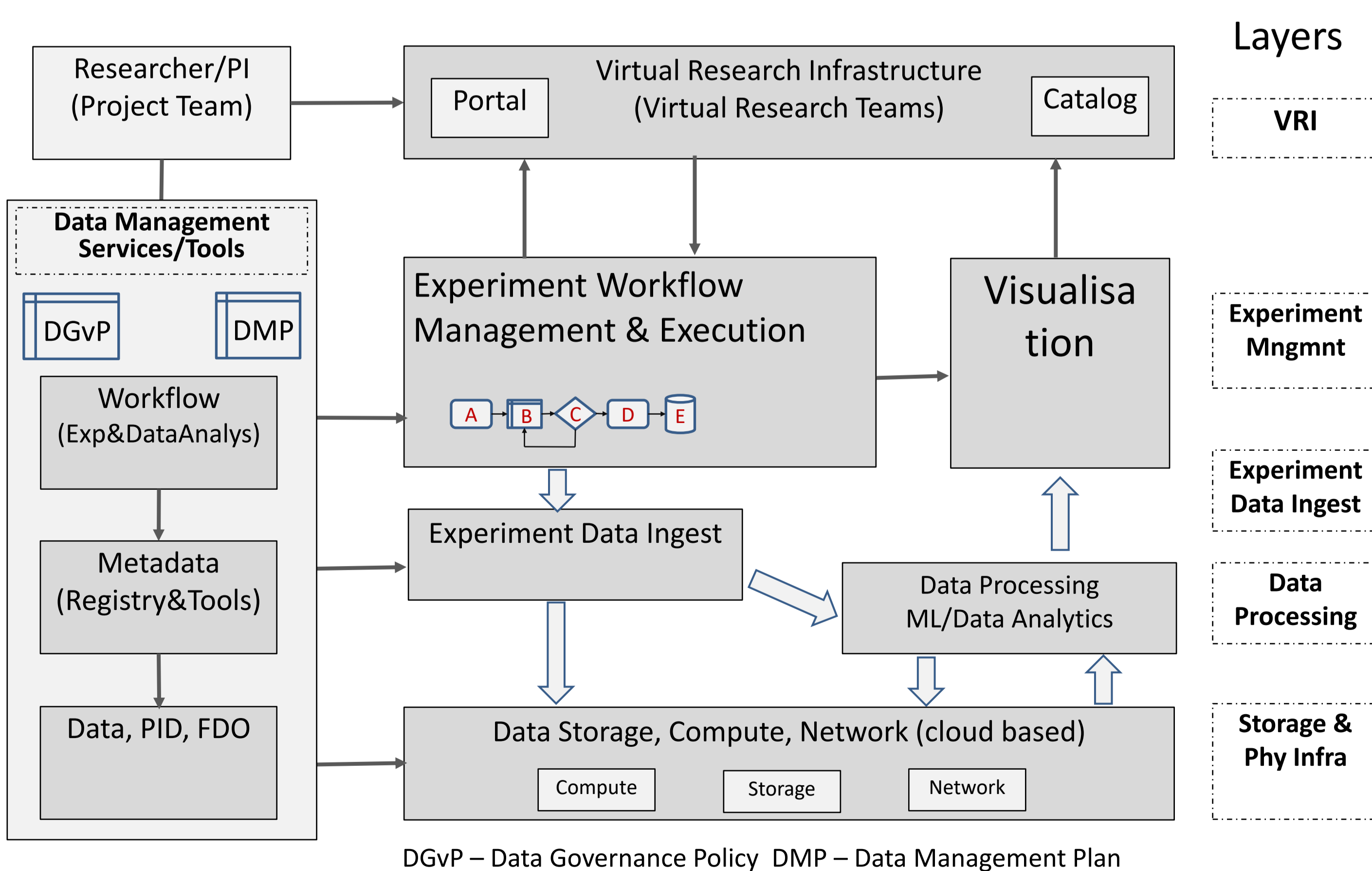
Data Types in Experiment operation

- Experiment Variables
- Configuration (equipment, DUT, Test Generator)
- Environment
- Test data
- Measurement data (data model, metadata)
- Orchestration and workflow
- Storage API/schema

What metadata should describe

- Data models: storage, databases, metadata
- Experiment
 - Orchestration; configuration; equipment: DUT, test generators, measurement; data storage; data models/metadata
- Dataflow: Stages, transformations, lineage/provenance, data models
- Workflow: Stages, Operations/conditions, workstations

Experimental Data Management Infrastructure



Data Management Infrastructure Layers

Data Management Infrastructure Layers to separate data management and governance concerns and actors/roles

- Layer 4 - Experiment Infrastructure configuration and management
- Layer 3 - Experimental data collection/recording

- Layer 2 - Data processing
- Layer 1 - Data Storage, Archiving, Exchange

- Data Management Services and Tools (Data Management Plane)
- Data Management Plan and Data Quality Assurance, FAIR compliance
- Metadata registries and tools
- Data Security and Data protection, GDPR

References

SLICES-RI - <https://slices-ri.eu/> SLICES-PP - <https://www.fairsfair.eu/> GreenDIGIT <http://www.greendigit-project.eu/>
 Yuri Demchenko, et al, SLICES Data Management Infrastructure for Reproducible Experimental Research on Digital Technologies, Proceedings Workshop on FutureG Experimental Test Platforms for Advanced Systems Implementation and Research, IEEE Global Communications Conference, 4–8 December 2023, Kuala Lumpur, Malaysia
 Fdida, Serge, et Al, SLICES, a scientific instrument for the networking community, Computer Communications, 2022, ISSN 0140-3664, <https://doi.org/10.1016/j.comcom.2022.07.019>

