Easy access to HPC workflows and data mangement

Martin Golasowski
IT4Innovations, Czech Republic

EasyBuild User Meeting 2025

Extreme Data Analytics in an Exascale Era with HPC-Cloud-AI Convergence

EXA4MIND

Extreme Data Analytics as Enabler for Science/SMEs/Industry

At Top-Level EU Supercopmuting Centres, Cross-System With Automatised Workflows

Connecting to European Data Spaces, EOSC and EUDAT

How to supercompute ...

- Pick a supercomputer
- Fill out a request for allocation
- Get it approved
- Set up an account
- Set up a SSH key
- Know how to use Linux terminal
- Login in to the Supercomputer
- Learn how to launch jobs in SLURM
- Learn about modules
- Learn about storages and data transfer
- ... finally compute something

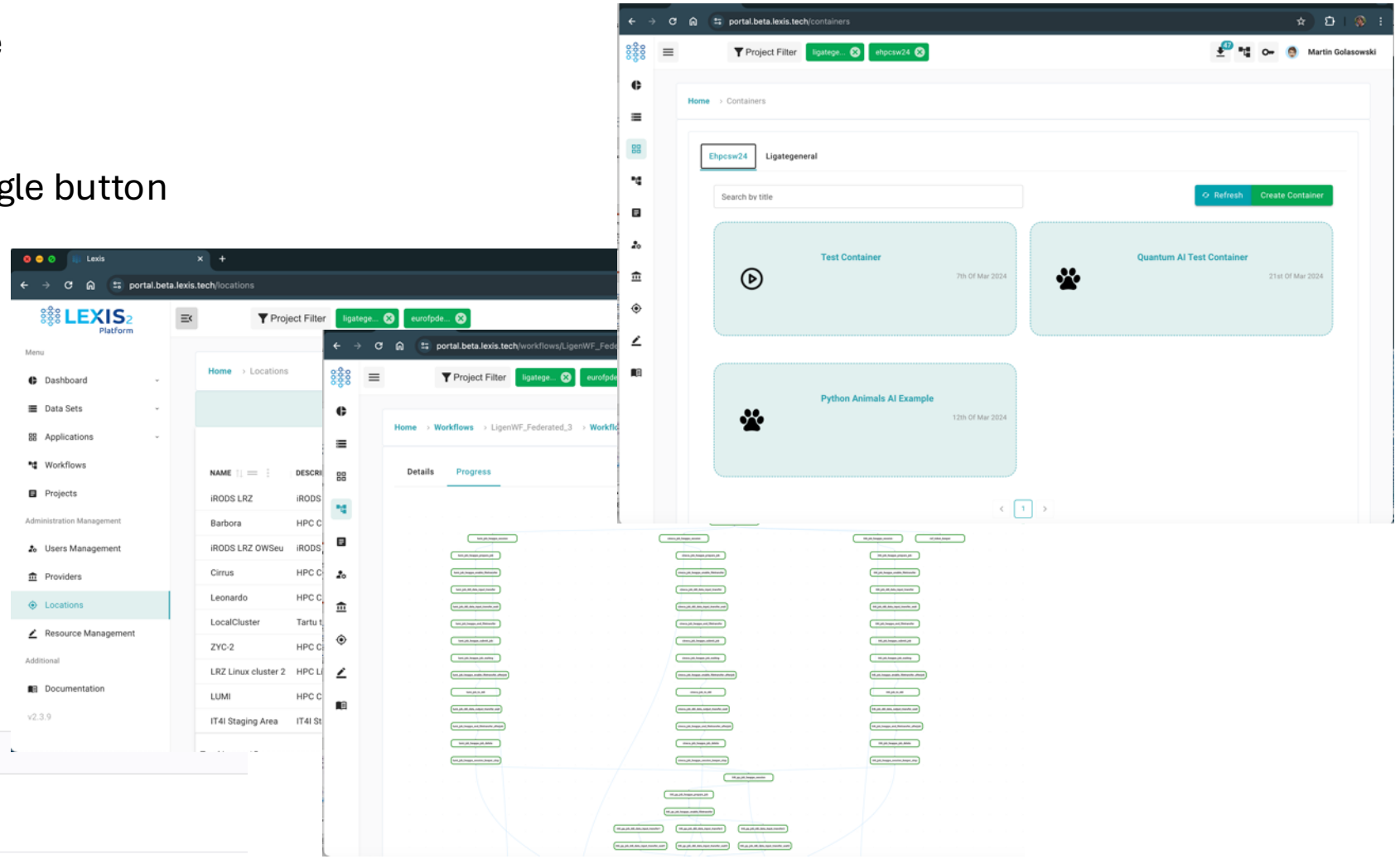
Easy and Safe access

- Let’s have a nice web interface
- Allocations at one place
- Across multiple clusters
- Launch applications with a single button
- Get logs
- Manage data in iRODS
- Use common web login
- … and many more

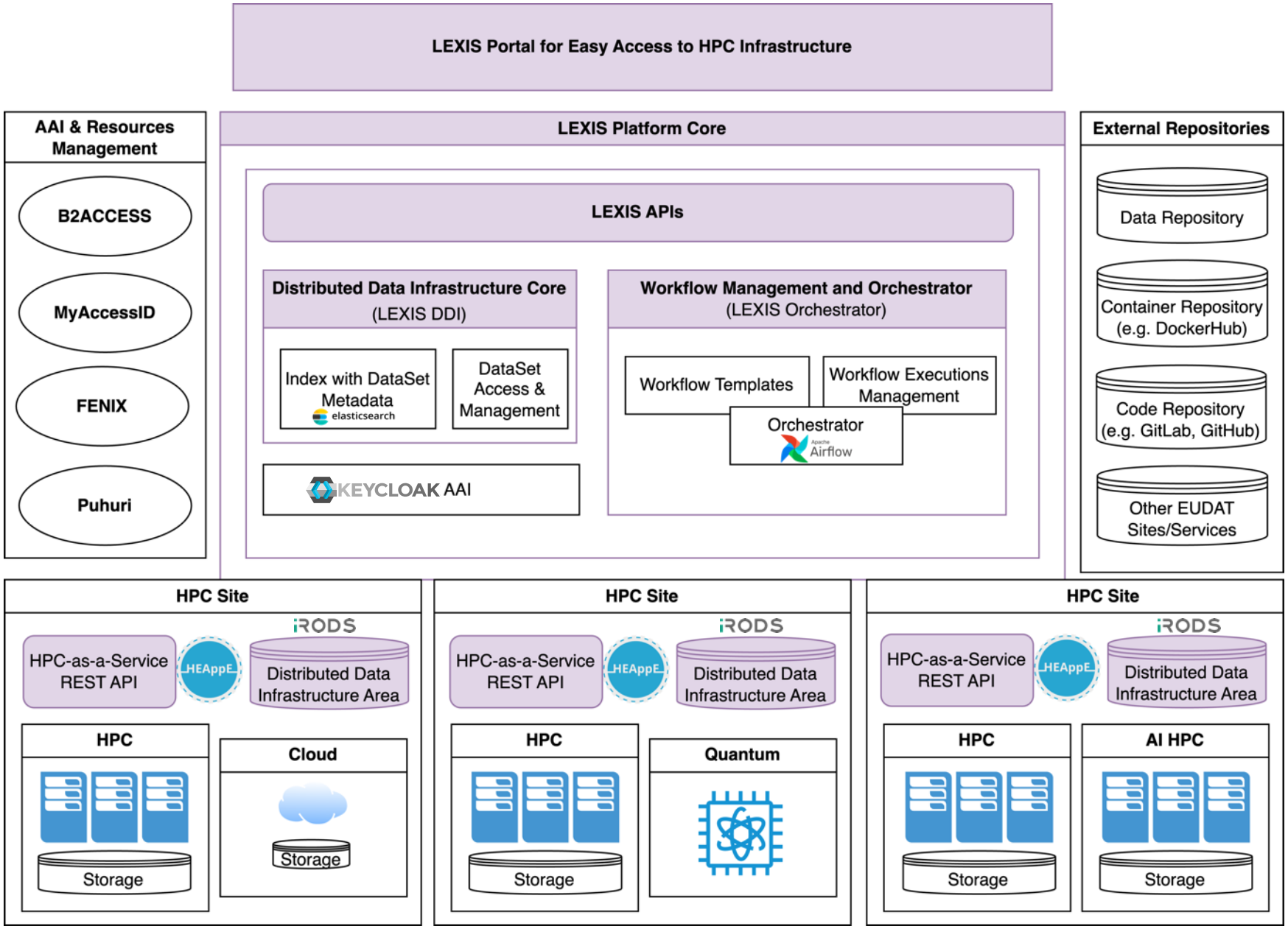
Visit for more:

docs.lexis.tech

opencode.it4i.eu/lexis-platform



LEXIS Platform

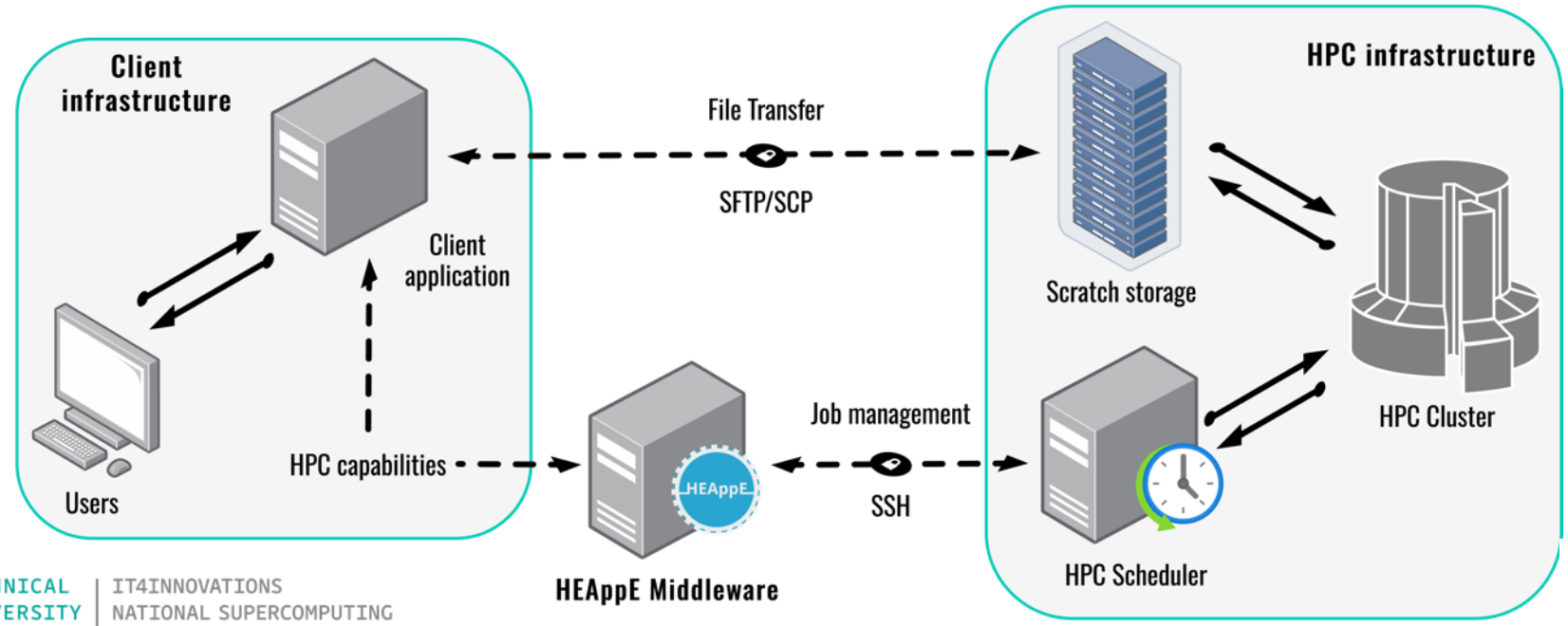


- Federation of **European computing centres**
- **Hiding** of technical and **operational differences** across organizations
- **HPC & Cloud** service providers, Data providers
- Unified & distributed **data management**
- Orchestration
- **Federated Authentication & Authorization** Infrastructure (AAI)

HPC-as-a-Service

High-End Application Execution Middleware

- Providing HPC capabilities as a service to client applications and their users
- Secured and restricted access to HPC infrastructure
- REST API for easy access and integration
- Authentication and authorization to provided functions
- Monitoring and reporting of executed jobs and their progress



HEAppE

HashiCorp Vault

.NET Core

REST API

jupyter

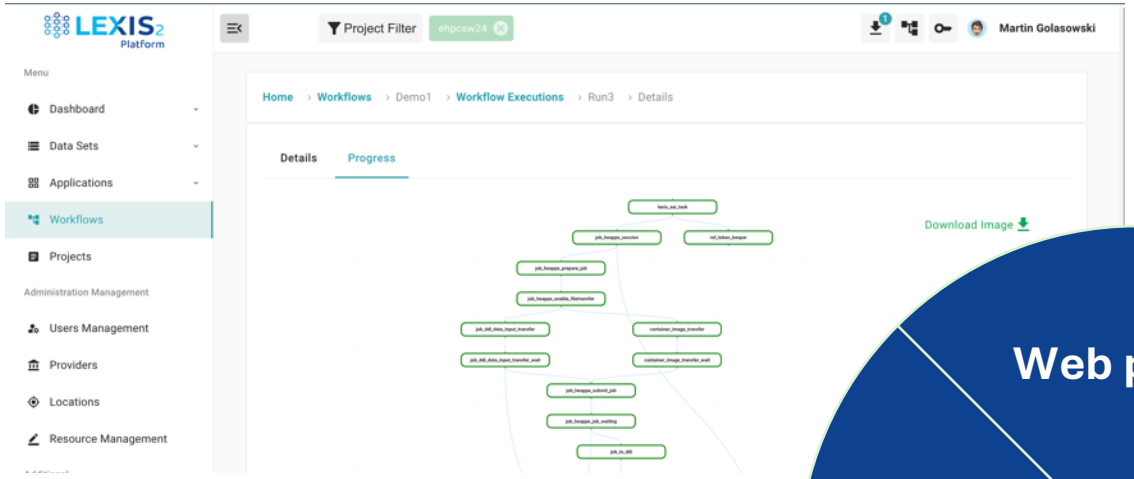
Swagger

docker

slurm workload manager

Altair | PBS Works™

LEXIS Platform



Web portal

Public REST API

Python library SDK

Terminal based interface

workflowManagement Actions relating to management of Workflows and Workflow Executions

GET	/workflow	Return list of available LEXIS Workflows	
POST	/workflow	Create a new LEXIS Workflow on the system	
GET	/workflow/{workflowId}	Return detailed info on LEXIS Workflow for given Workflow ID	
DELETE	/workflow/{workflowId}	Delete LEXIS Workflow on the system	
GET	/workflow/{workflowId}/execution	List the current available LEXIS Workflow Executions.	listWorkflow
POST	/workflow/{workflowId}/execution	TODO: Needs implemented with TOSCA 1.3 Capabilitise. Create a new LEXIS Workflow Execution by providing remaining inputs	createWorkflowExecution
GET	/workflow/{workflowId}/execution/{workflowExecutionId}	Returns LEXIS Workflow Execution detail.	getWorkflowExecutionDetail
DELETE	/workflow/{workflowId}/execution/{workflowExecutionId}	Cancel a LEXIS Workflow Execution.	cancelWorkflowExecution

README.md

py4lexis

Package py4lexis provides functions to manage datasets within the LEXIS platform. Uploading datasets' files is performed by tuspy Python package, i.e. by TUS Client.

Install

We recommend to use virtual environment, e.g.:

```
python -m venv ./venv
```

vated virtual environment, type:

```
-m pip install py4lexis
```

```
preprocessing_dataset_path_geographical_data_path": "project/proj934a1763631a81e

flowExecutionName": "Run3",
flowExecutionConsent": True

are_data():
epared = WRF_EXEC_PARAMS
epared["inputParameters"]["preprocessing_start_date"] = datetime.datetime.now().str
prepared["inputParameters"]["preprocessing_start_date"] = '2022120400'
return prepared

def submit_wf_execution(token):
    res = requests.post(LEXIS_API + WORKFLOW_ID + '/execution',
    json = prepare_data(),
    headers = {'Authorization': 'Bearer ' + token})

    if res.status_code != 201:
        logger.error(res.content)
        return

    return res.json()
```

```
(venv) mgolas@Martins-MacBook-Pro ~/I/L/L/Py4Lexis (develop)> python examples/upload.py
Progress:  |                    ----------------------------------| 30.0% Uploaded
```

LEXIS Platform

Selected use-cases

LIGATE Project

Application for molecular docking simulation – private IP by DOMPÉ
LEXIS provides access to workflows with this application running on HPC
Without direct access to the binary or source code

OpenWebSearch.eu

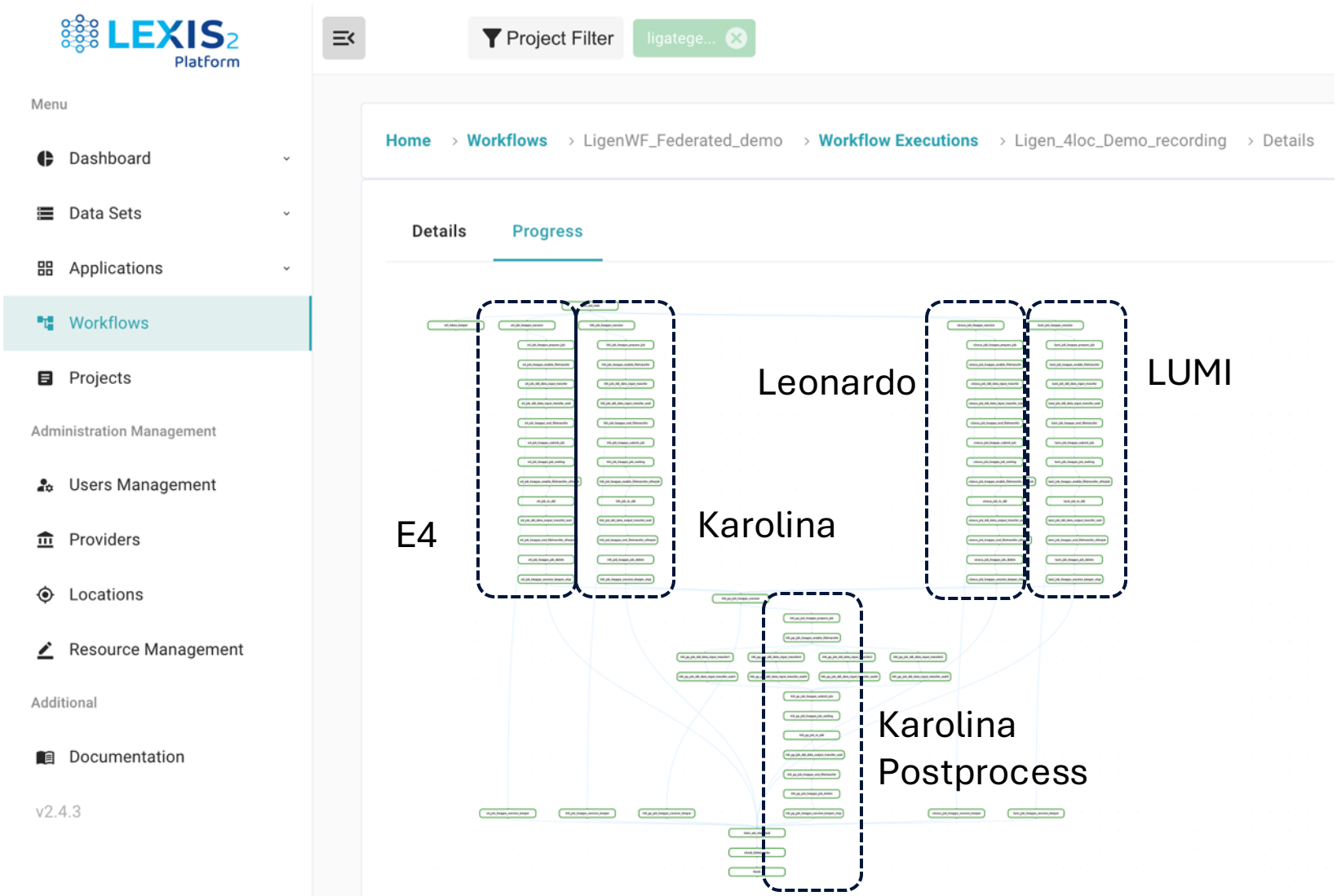
European open web index processed through LEXIS on several HPC locations (LRZ, IT4I, CSC, DLR)
Public indices made available trhough the LEXIS Portal

LEXIS Platform

User story - DOMPÉ Company & LiGen Software

- DOMPÉ has its own code co-developed with POLIMI & CINECA
- DOMPÉ LiGen code IP does not allow to share LiGen SW (even binary)
- DOMPÉ would like to extend its customer base thanks to LEXIS Platform and EuroHPC infrastructure without losing IP
- DOMPÉ would like to make available LiGen to non-profit academic research and public institutions as a Platform-as-a-Service solution for drug discovery

Federated execution on 4 HPC clusters at once

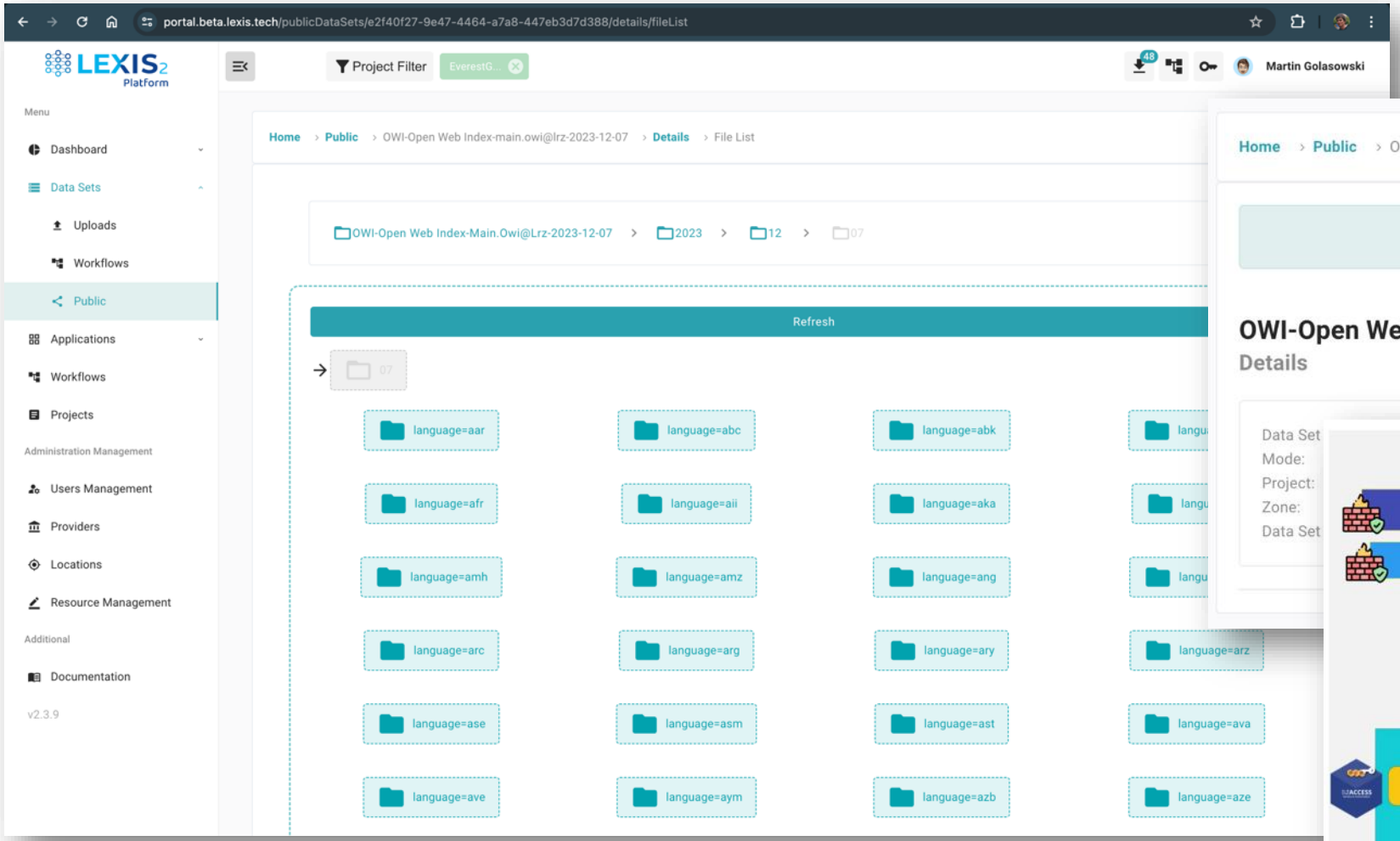


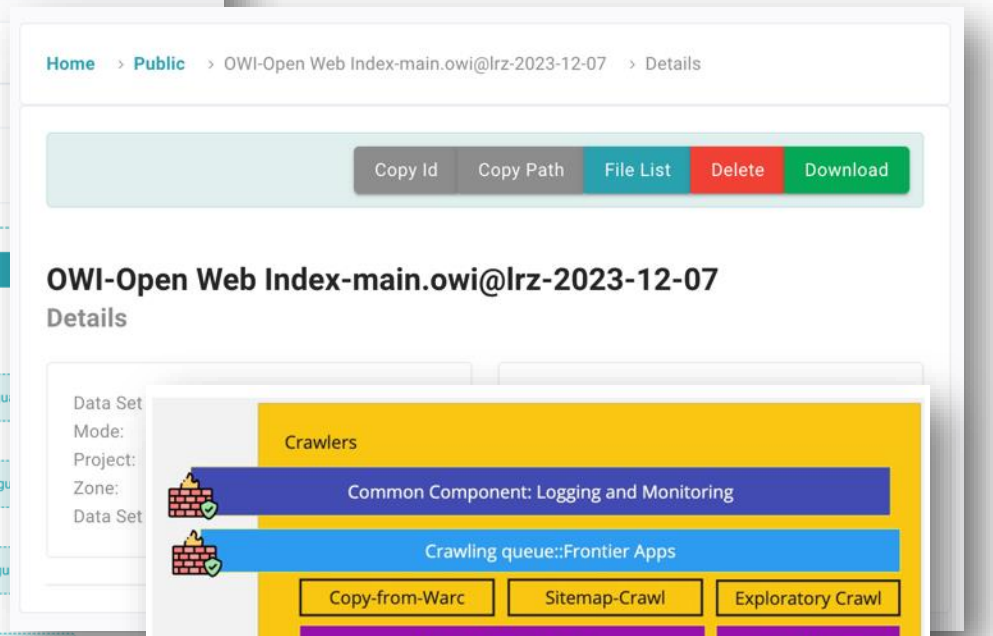
Executed on 4 locations

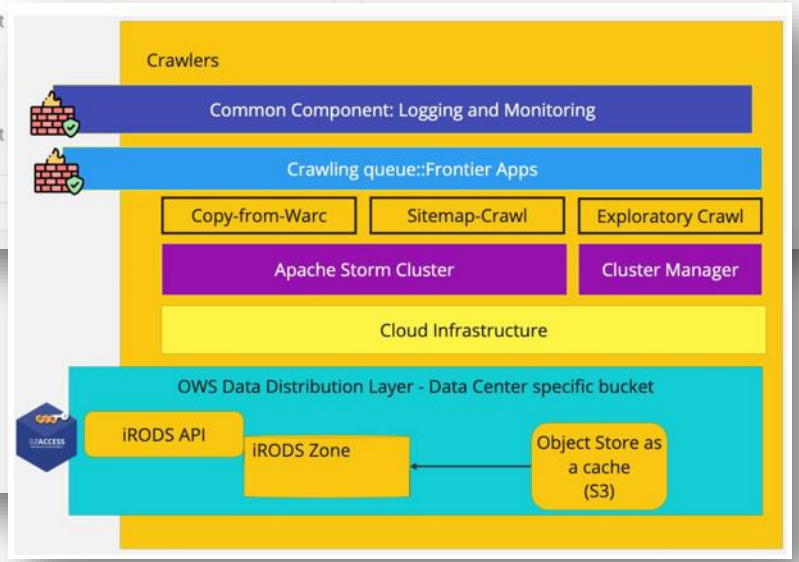
- LUMI
- Leonardo (CINECA)
- Karolina (IT4I)
- E4 AMD Cluster

OpenWebSearch.eu - Public web indexes

LEXIS2 Platform







VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

FLOREON+

- Uses LEXIS workflows for simulations and data processing
- Developed since 2006
- Decision support system for crisis management domain

Hydrology

- Monitoring of hydrological situation and flood forecasting

Traffic

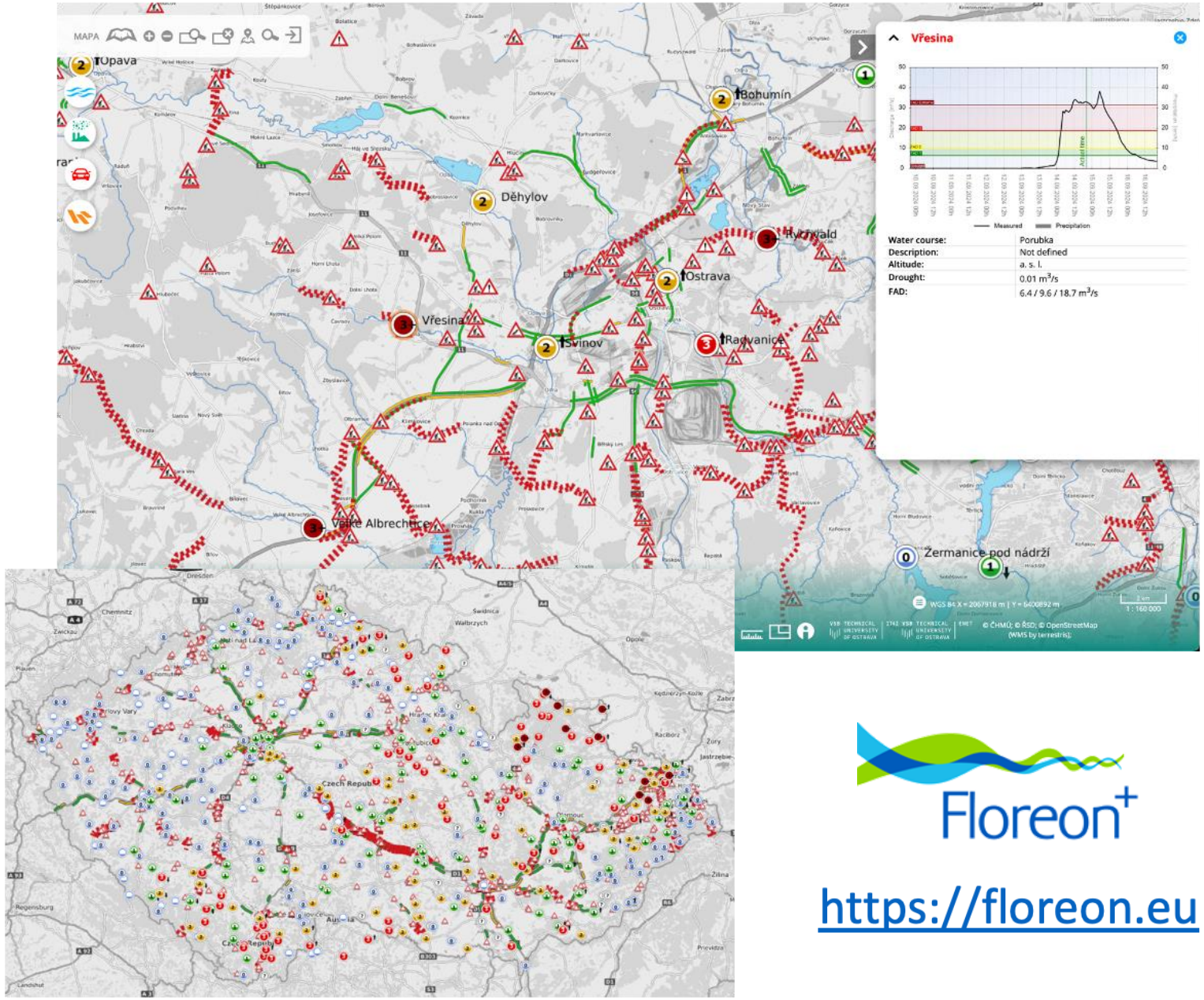
- Monitoring and modelling of traffic situation

Air quality

- Monitoring of air quality and simulation of the spread of hazardous substances

Landslides and subsidience

- Tracking of elevation point movements

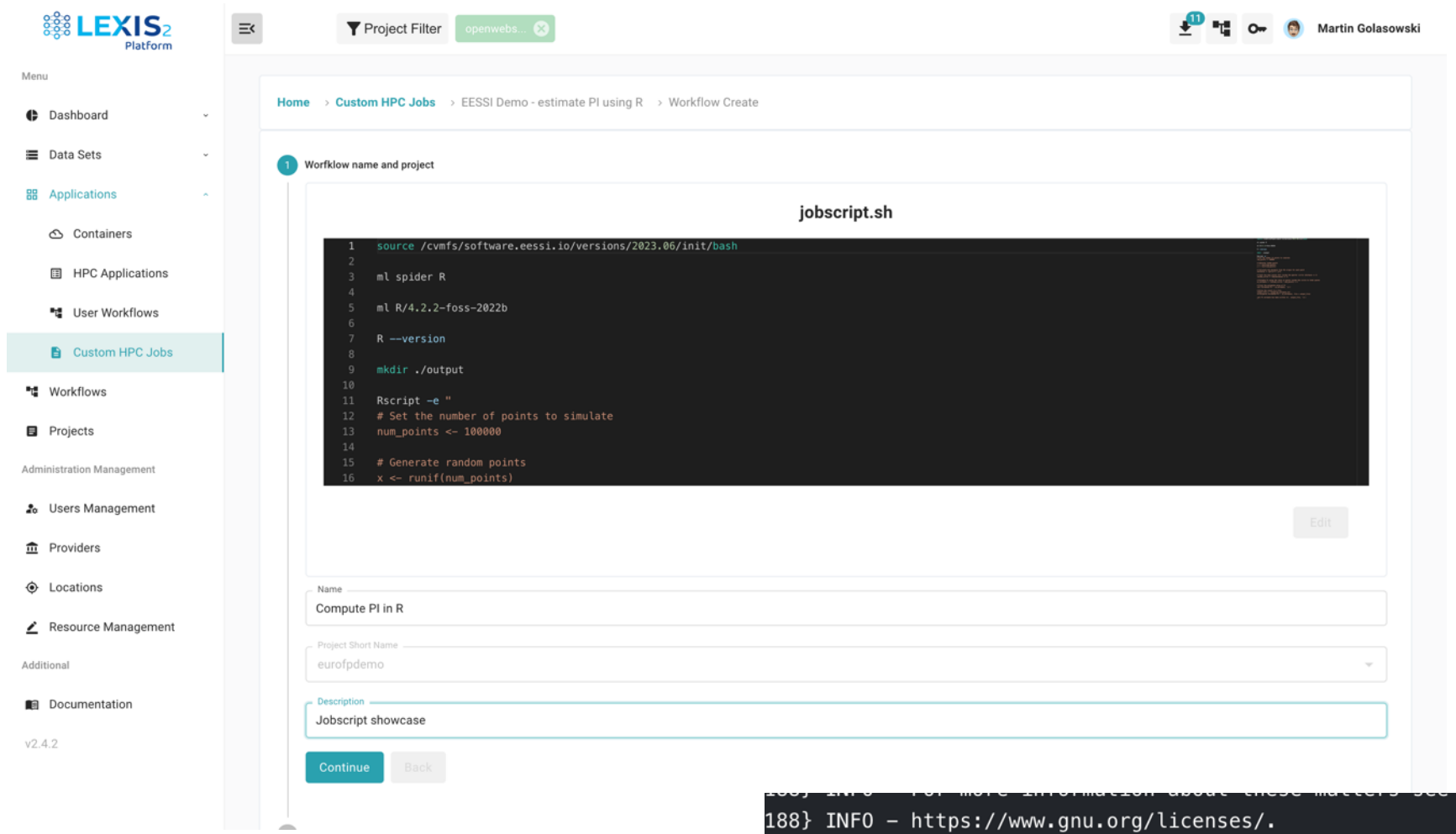


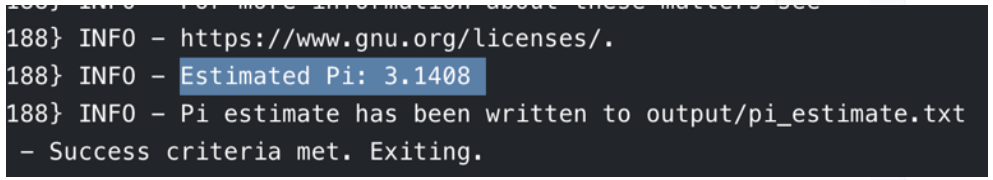
Floreon+

https://floreon.eu

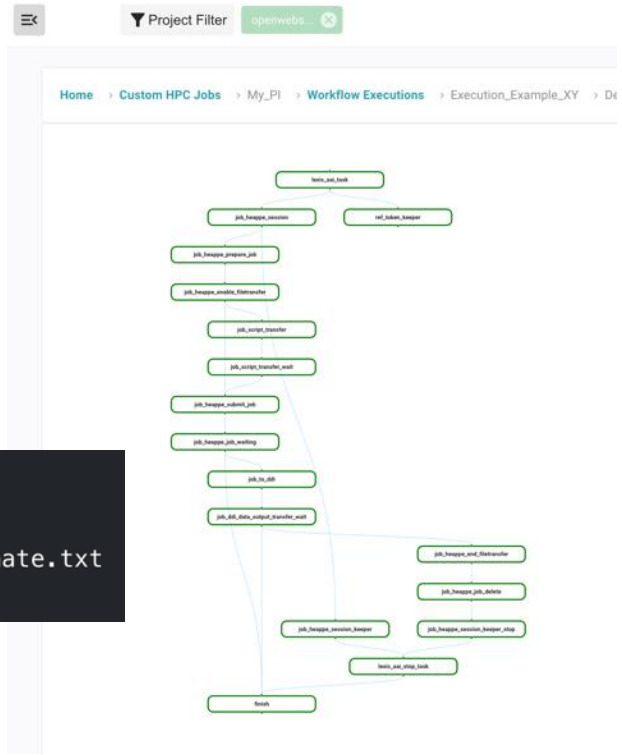
VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

Custom jobscripts in a workflow





- Run custom R script
- Use common EESSI HPC module
- Workflow & output in GUI



VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

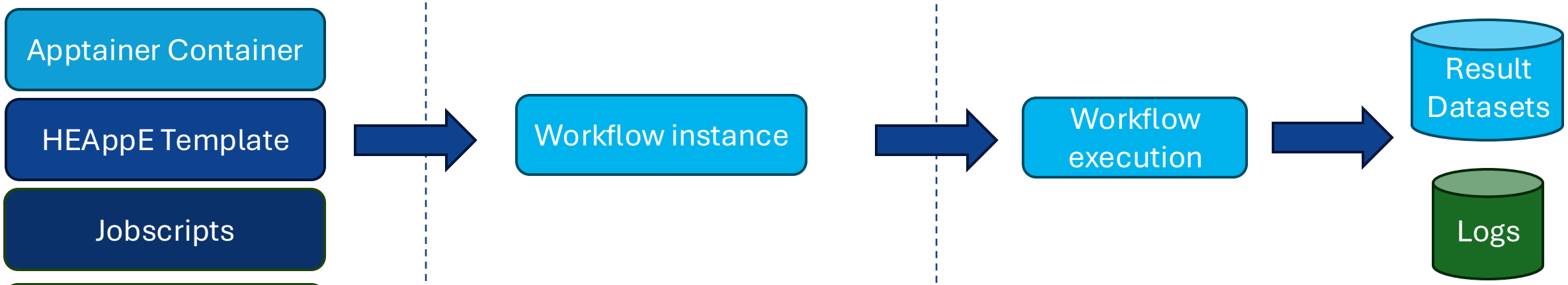
Workflow execution model

LEXIS2 Platform

Templates

Instance

Execution



- HPC application deployed (compiled) by specialist
- Custom workflows with K8S, schedules, etc.

- Target allocation / resource
- Data handling
- Default execution-time-parameters

- Input / output datasets
- Application params / environment vars
- Scheduling policy
- Static location

Airflow integration in LEXIS

LEXIS Platform

Apptainer Container

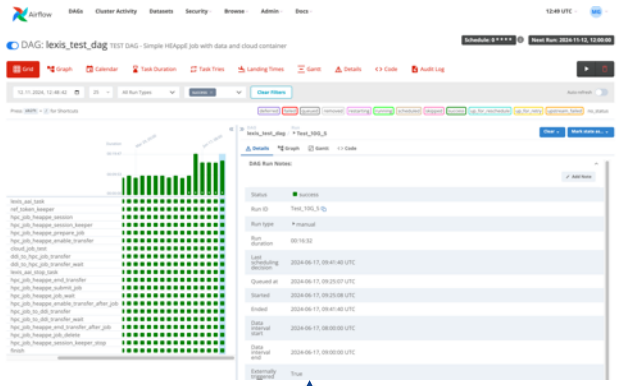
HEAppE Template

Jobscripts

CWL

Custom Workflow DAG

Apache Airflow



Workflow specification

Integration with LEXIS APIs

LEXIS Provider

Operators
DDI, HEAppE,AAI

LEXIS API Plugin

AAI, Container/HEAppE
DAG Templates
YAML Parser

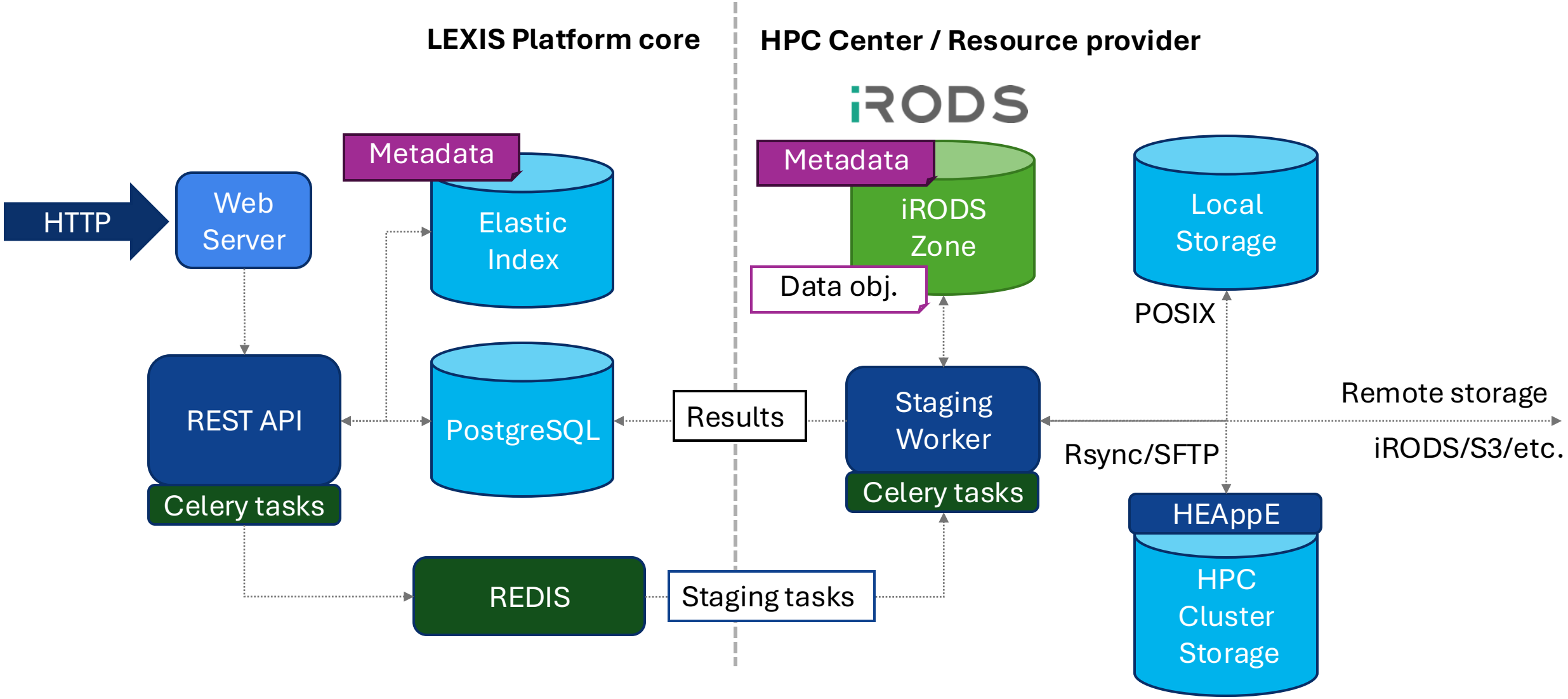
DDI

HEAppE

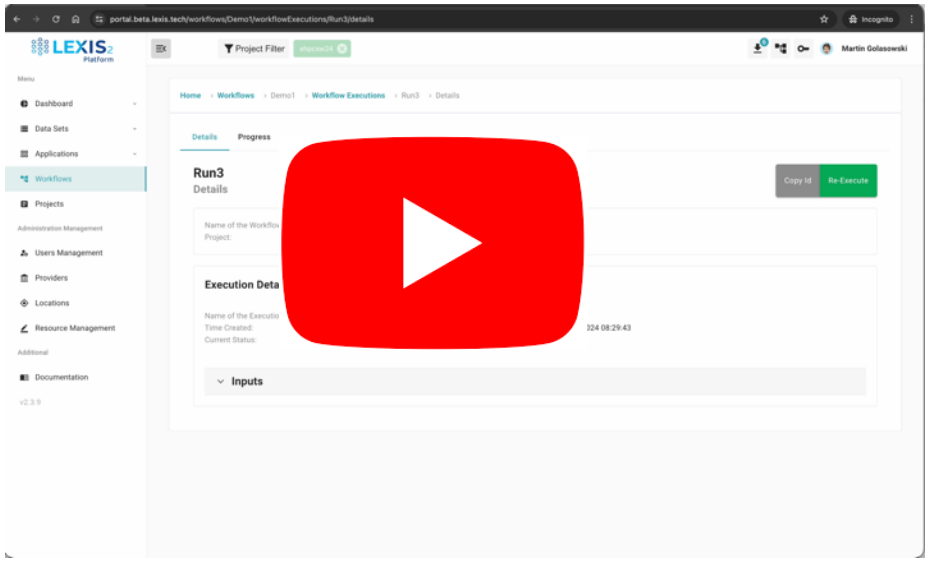
AAI

UserOrg

Distributed Data Interface – data staging



AI workflow demo with Python in Apptainer

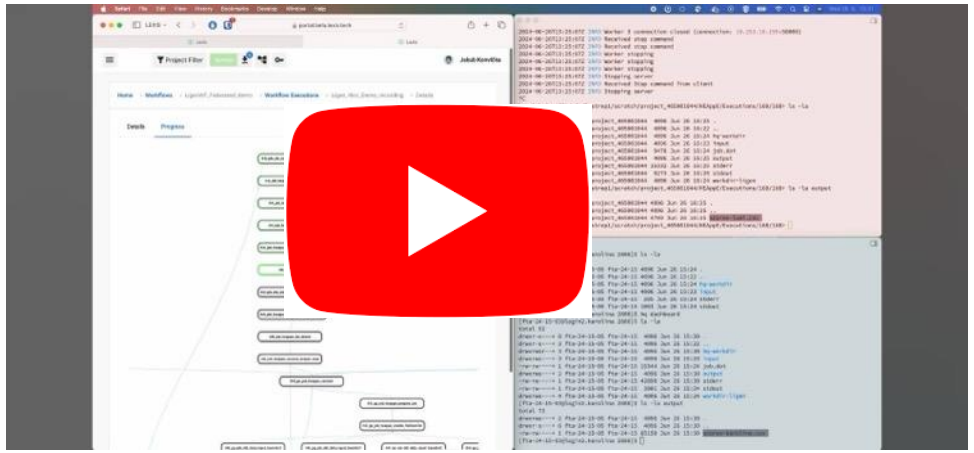


Check it out!

Thank you!

LEXIS2 Platform

Urgent computing federated workflow execution



Check it out!

www.lexis.tech

VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER