OpenRiskNet

RISK ASSESSMENT E-INFRASTRUCTURE

WP2-Interoperability, Deployment and Security

Ola Spjuth (UU)

General Assembly and Final Consortium meeting
22 October 2019, Amsterdam (The Netherlands)

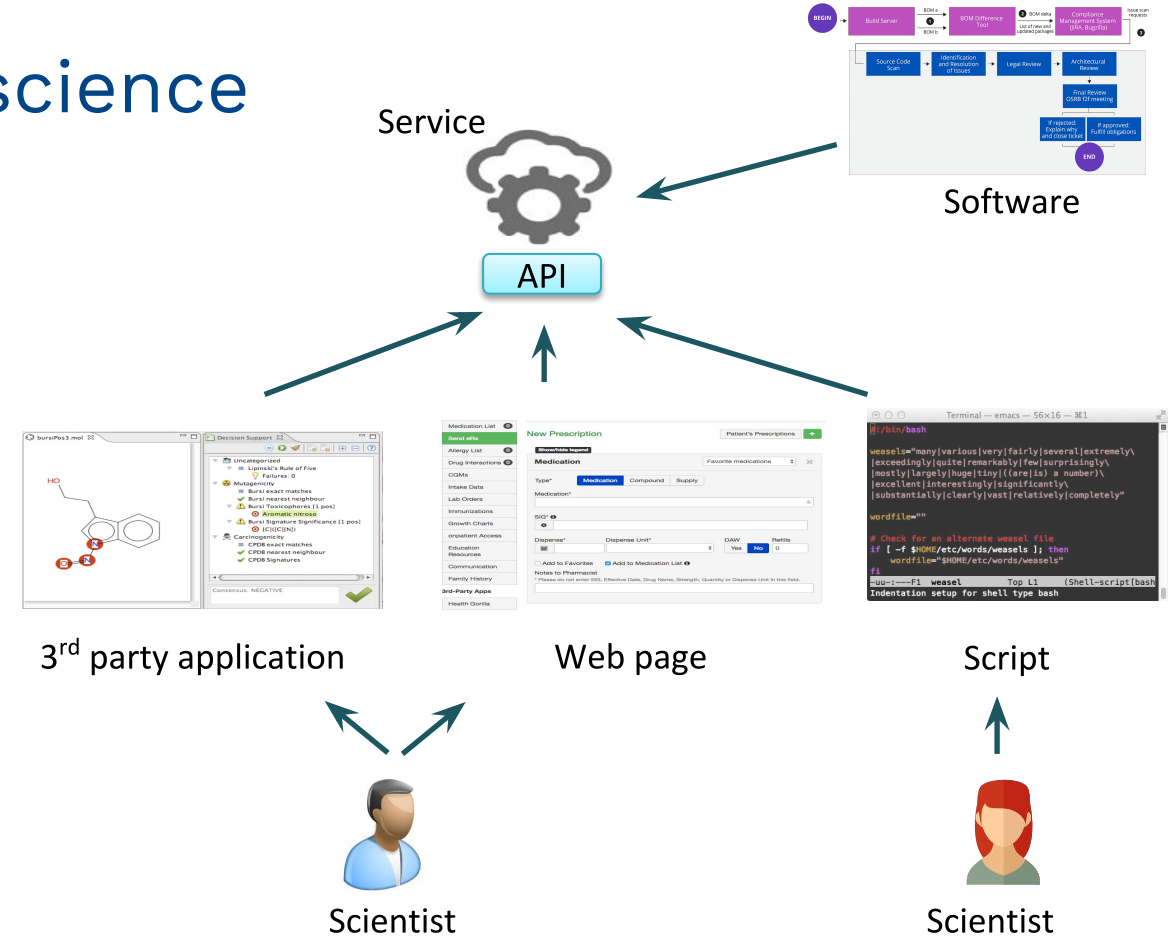
OpenRiskNet: Open e-Infrastructure to Support Data Sharing, Knowledge Integration and *in silico* Analysis and Modelling in Risk Assessment
Project Number 731075

WP2 Objectives

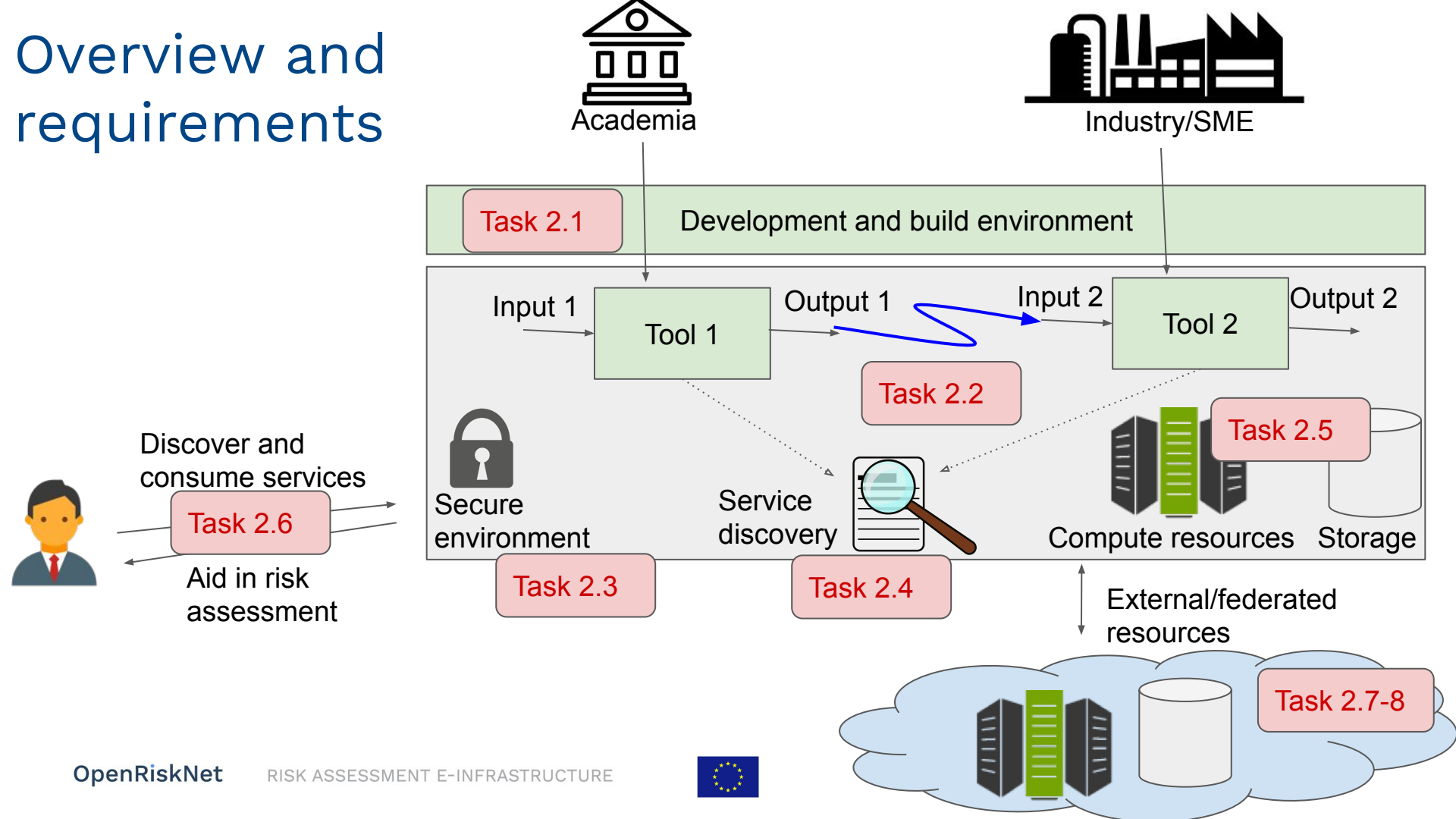
- Allow for interoperability between data and services at the technical level
- Provide mechanisms (APIs) for describing services and data
- Provide mechanisms for discovering relevant services and data
- Define approaches for deploying services as containers
- Allow for services to be accessed securely, as needed by commercial parties
- Allow for new OpenRiskNet virtual environments containing these tools and services to be instantiated by users on cloud or internal architectures
- Provide federation of services and security between different OpenRiskNet virtual environments and with external HPC facilities

Service-oriented science

- Standardize
  - Agree on e.g. interfaces, data formats, protocols etc.
- Decompose and compartmentalize
  - Experts (scientists) provides services
  - Achieve interoperability by exposing data and tools as Web services via APIs



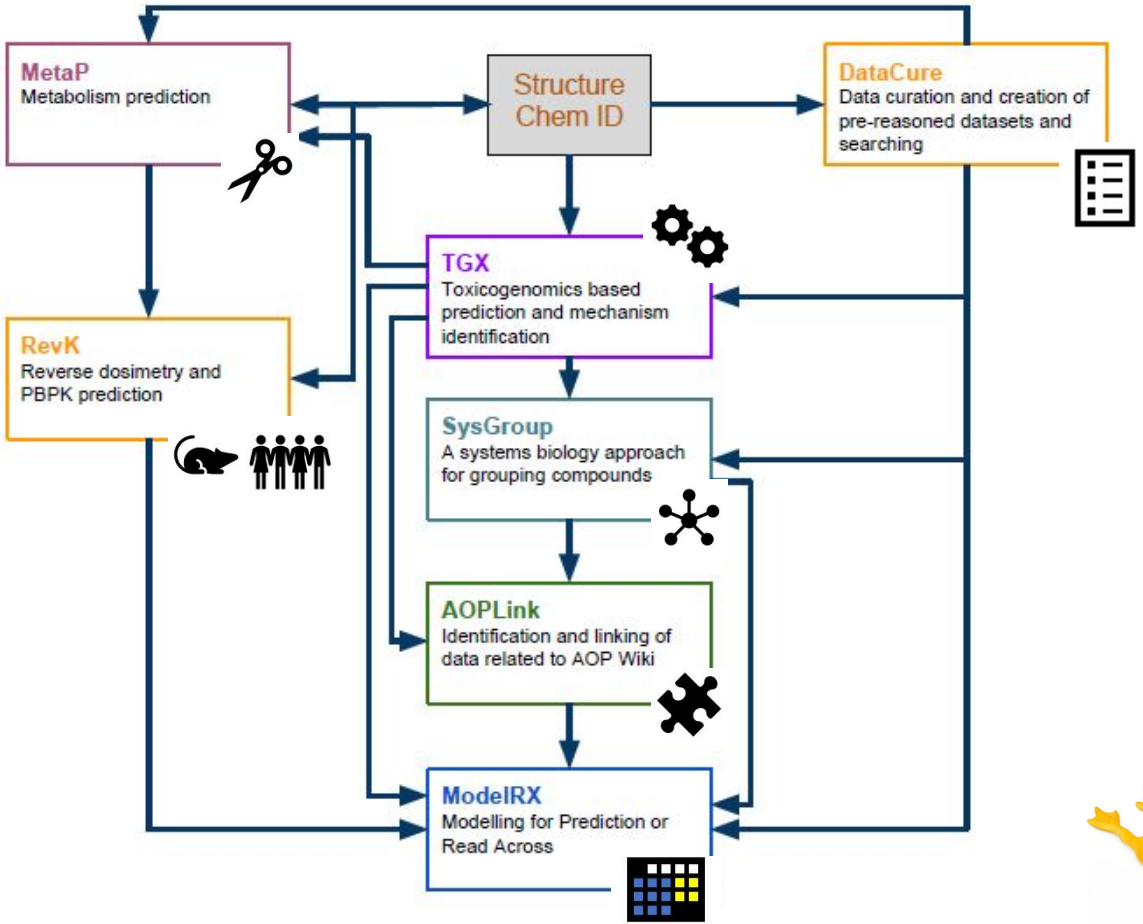
Overview and requirements



Case studies revisited

Key requirements:

- Extensible
- Interoperable
- Scalable



OpenRiskNet e-Infrastructure approach

- *Package tools as Software Containers (WP4) in a development environment (T2.1) and deploy them in Virtual Research Environment (T2.5) with adequate security (T2.3)*
- *Tools can be used/connected via interoperable APIs (T2.2) and discovered in the system (T2.4)*
- *Allow for connecting VRE with external HPC (T2.7) and federate data between VREs (T2.8)*
- *Provide a reference instance for demonstrational purposes (T2.6)*

- Use modern and established tools and frameworks supported by industry
  - Reduce risk and improve sustainability
- Offer an agile and scalable environment to use, and a straightforward platform to extend
- Deploy on physical but primarily on virtual (cloud) resources
- Microservices implemented as Docker containers

Task 2.1: Development environment

- Relying on GitHub for source code, issues, wiki
  - Github Organization: http://github.com/openrisknet
  - Issue tracker: https://github.com/OpenRiskNet/home/issues
  - Wiki: https://github.com/OpenRiskNet/home/wiki
- Red Hat OpenShift: container-based applications
  - Builds on Kubernetes container orchestration
  - Scaling, resilience
  - Hardened security, Single-sign-on (SSO)
  - Best-practices/utilities for Continuous Integration, Continuous Deployment (CICD)
- Status: Operational production and development environment
  - Reported in D2.1

Task 2.2: API specification and semantic interoperability

Aim: Ensure the interoperability of services via common APIs.

Status:

- After evaluations, decided on using OpenAPI (Swagger) description of REST APIs, supplemented with JSON-LD semantic descriptions using OpenAPI extension mechanism
- Semantic descriptions using controlled ontologies
- Example descriptors generated for some of our services
- Query tool created to explore use of these OpenAPI/JSON-LD descriptors
- Series of use cases for queries created
- *Described and reported in D2.2 and D2.4*

Remaining during November:

- Contribute to D2.6
- More???

Task 2.3

Task 2.3: Establish security environment

Aim: Control access to the OpenRiskNet virtual environments and security and privacy of data within it.

Progress:

- Deployment of Red Hat KeyCloak providing distributed, federated Single Sign On solution
- ORN reference site configured to use LinkedIn and GitHub as Identity Brokers
- Other VEs can be configured according to local needs
- Deployment has been streamlined and is part of the core VE creation process
- Keycloak upgraded to version 4.8.3.Final
- Many partner applications have been configured to authenticate against Keycloak
- Provides basic user statistics when user logs in
- *Described and reported in D2.3.*

Remaining during November:

- None
- Contribute to D2.6?

Task 2.4

Task 2.4: Services discovery

Aim: Allow to discover services or datasets running inside the VRE using their semantic annotations defined in Task 2.2.

Progress:

- OpenRiskNet service registry running at http://orn-registry-openrisknet-registry.prod.openrisknet.org/
- Semantic querying of services demonstrated via OpenRiskNet Query Tester (Web-based UI tool): https://orn-query-test.cloud.douglasconnect.com/
- Documentation on how to annotate services to support discovery
- Sample services fully annotated (lazar prediction service (JGU), Jaqpot modelling service (NTUA) and the chemidconvert service (DC))
- *Reported in D2.3*

Remaining during November:

- Contribute to D2.6. More????

Task 2.5: Deployment of virtual infrastructures and container orchestration frameworks

Task 2.5

Aim: Allow users to instantiate a new OpenRiskNet virtual environment on a cloud infrastructure or an in-house server/workstation.

Progress:

- Establish documented means and processes to deploy OpenRiskNet virtual environments to public and private cloud systems, as well as local computer resources.
- Enable scheduling and orchestration of microservices implemented as software containers in an elastic, fault-tolerant multi-node environment, supporting both long-running and short-running operations.
- Facilitate locating and adding additional services and tools to the virtual environment.
- Setup and configuration of security, logging and auditing frameworks capable of monitoring performance and producing sufficiently detailed audit trails
- Provide a mechanism for upgrading services
- Reported in D2.3

Remaining during November:

- Contribute to D2.6.

Task 2.6

Task 2.6: Establishment and maintenance of OpenRiskNet reference instance

The OpenRiskNet reference instance will be a portal that provides project members access to all the approved tools and services and provide an environment in which the case studies can be exemplified.

- Setup of reference site on a public cloud.
- Incorporation of approved OpenRiskNet services and tools
- Provide ability to generate and document workflows demonstrating the case studies
- Provide instructions for how users can spawn their own individual ORN virtual environment

- Operational at https://home.prod.openrisknet.org/
- Service listing: http://registry.prod.openrisknet.org/

Remaining during November:

- Continued maintenance. Contribute to D2.6.

Task 2.7

Task 2.7: Interconnecting virtual environment with external infrastructures

Establish mechanism for how a virtual environment can interact and utilize external infrastructures, such as GRIDs and HPC batch systems. Such infrastructures may be more suitable for particular IO/CPU-intensive tasks such as 'omics analysis, and there are relatively big resources existing that are available to the research community.

Progress:

- Nextflow now supports K8S and OpenShift as batch executors
- Workflows can be launched within VRE and burst into public cloud infrastructure
- Enabled infrastructure switching dependant of workload resource requirements
- Squonk enhanced to support execution of Nextflow pipelines

Remaining during November:

- TBC

Task 2.8: Federation between virtual environments

Enable data and services from one virtual environment to be accessed from another virtual environment.

Progress:

- Investigated data federation platforms iRODS, OneData, and OwnCloud/NextCloud
- Demonstrated Data Federation with NextFlow
- Investigated federation of Kubernetes clusters using Kubernetes Federation
- Reported in D2.5

Remaining during November:

- None

Technology partners

Problems encountered

Fragility in the OpenStack cloud environments on SSC, in OpenShift and GlusterFS storage caused significant delay and extra work. Maintenance effort was greater than expected.

- Worked with infrastructure providers (SSC) to resolve OpenStack problems
- Worked with Red Hat to harden OpenShift installation process
- Created ‘Orchestrator’ to streamline provisioning process
- Created ‘event log’ to capture problems and process
- Invested in automation to fix routine problems

Net impact was that more effort was spent on infrastructure than was originally anticipated.

WP2 Tasks

ID	Title	Partners involved	Start month	End Month	Deliverables
T2.1	Creation of development environment	IM, UU, EwC, CRG, NTUA	1	6	D2.1
T2.2	API specification and semantic interoperability	JGU, EwC, NTUA, IM, UU, UM	1	24	D2.2; D2.4
T2.3	Establish security environment	IM, UU, EwC	1	18	D2.3
T2.4	Services discovery	EwC, UU, IM	1	18	D2.3
T2.5	Deployment of virtual infrastructures and container orchestration frameworks	UU, IM, EwC, CRG, NTUA, UM	1	18	D2.3
T2.6	Establishment and maintenance of OpenRiskNet reference instance	UU, IM, EwC	13	36	D2.6
T2.7	Interconnecting virtual environment with external infrastructures	CRG, UU	13	30	D2.5
T2.8	Federation between virtual environments	UU, EwC	13	30	D2.5

WP2 Deliverables

ID	Title	Due month	Lead partner	Type	Dissemination level
D2.1	**Development infrastructure online**	**6**	**IM**	**Demonstrator**	**Public**
D2.2	**Initial API version provided to providers of services**	**6**	**JGU**	**Demonstrator**	**Public**
D2.3	**Report on deployment of virtual infrastructures with service discovery and container orchestration**	**18**	**UU**	**Demonstrator**	**Public**
D2.4	**Final API available for internal and external service providers**	**24**	**EwC**	**Demonstrator**	**Public**
D2.5	**Compute and data federation**	**30**	**UM**	**Other**	**Public**
D2.6	Reference OpenRiskNet system available online	36	UU	Demonstrator	Public

WP2 Milestones

ID	Milestones	Due month	Lead partner	Means of verification
MS2	Initial API version provided to providers of services	6	CRG	Online publication of API documentation
MS5	Operational deployment of virtual infrastructures with service discovery and container orchestration	18	IM	Reference virtual environment online
MS8	Final API available for internal and external service providers	28	UU	Specifications are reported on project website

WP2 Performance Metrics

Metrics	Status M36
Existence of reference virtual instances of the e-infrastructure	VRE is accessible by end users at: https://home.prod.openrisknet.org/.
Status report from regularly executed automatic testing procedures of core and services	TBC
Generation of list of all available services using the discovery service with all relevant information	This functionality is operational. OpenRiskNet Service Registry available at: http://registry.prod.openrisknet.org/

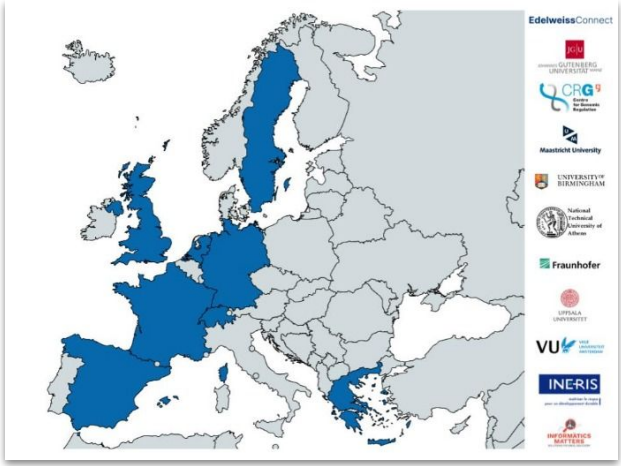
WP2 summary M36

- Only minor deviations from DoW
  - Semantic interoperability and Federation
- All deliverables submitted on time
- Only D2.6 remaining

Acknowledgements

OpenRiskNet (Grant Agreement 731075) is a project funded by the European Commission within Horizon 2020 Programme

Project partners:



P1 Edelweiss Connect GmbH, Switzerland (EwC)
P2 Johannes Gutenberg-Universität Mainz, Germany (JGU)
P3 Fundacio Centre De Regulacio Genomica, Spain (CRG)
P4 Universiteit Maastricht, Netherlands (UM)
P5 The University Of Birmingham, United Kingdom (UoB)
P6 National Technical University Of Athens, Greece (NTUA)
P7 Fraunhofer Gesellschaft Zur Foerderung Der Angewandten Forschung E.V., Germany (Fraunhofer)
P8 Uppsala Universitet, Sweden (UU)
P9 Medizinische Universität Innsbruck, Austria (MUI)
P10 Informatics Matters Limited, United Kingdom (IM)
P11 Institut National De L'environnement Et Des Risques INERIS, France (INERIS)
P12 Vrije Universiteit Amsterdam, Netherlands (VU)