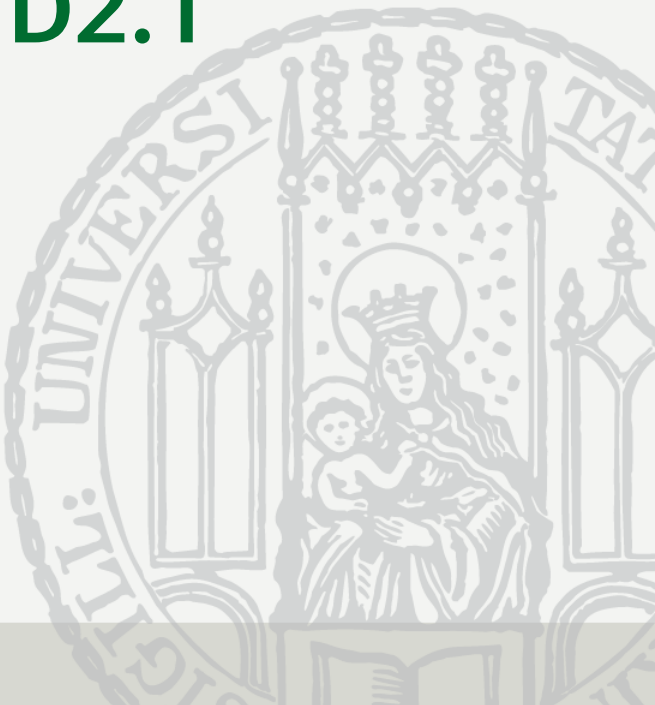
LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

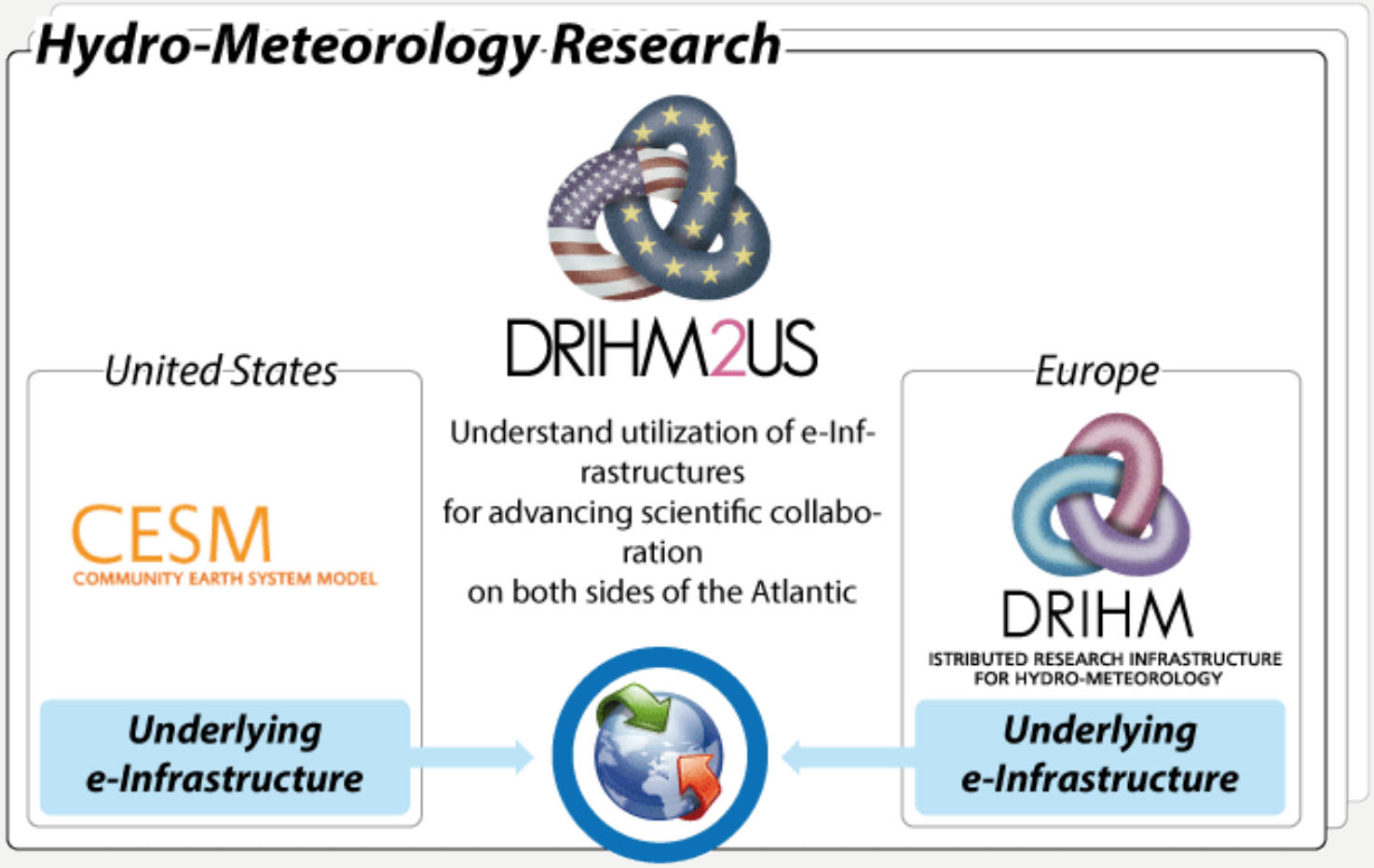
Christian Straube

Review of DRIHM2US D2.1

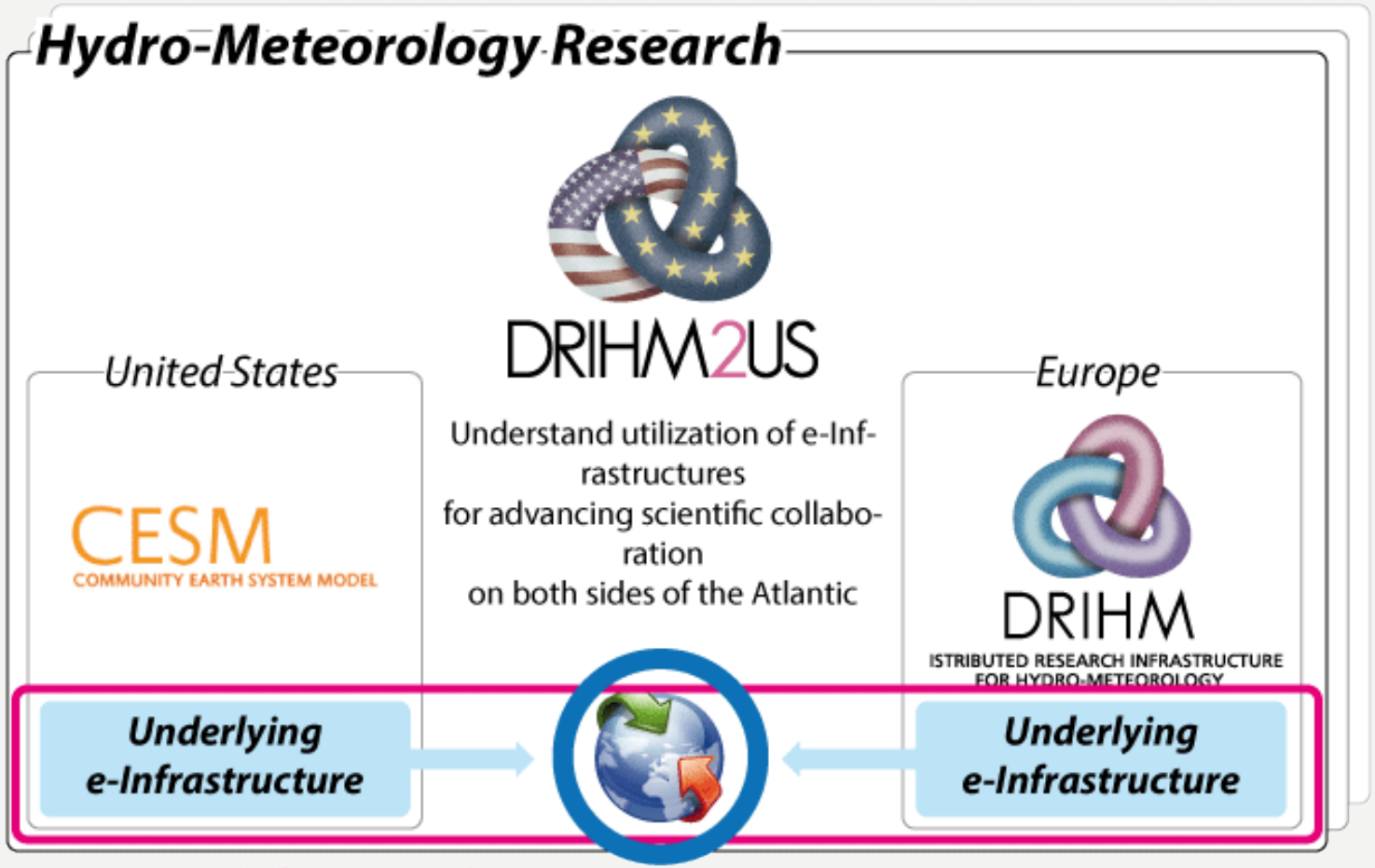
Report on current approaches



Context of WP2

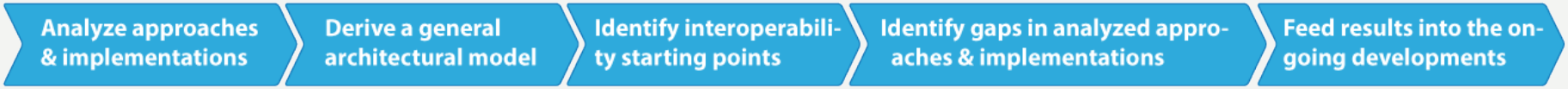


Purpose of WP2



Analyse different e-infrastructure approaches in the U.S. and Europe (as far as they relate to hydro-meteorological research)

Methodology of WP2



Tasks of WP2

Analyze approaches & implementations → **Derive a general architectural model** → **Identify interoperability starting points** → **Identify gaps in analyzed approaches & implementations** → **Feed results into the ongoing developments**

Task 2.1 Analyze current approaches (LMU)

- Comparative analysis of different approaches underlying the organization of e-infrastructure for HMR in Europe & U.S.
- Analysis core aspects: Geospatial cataloguing interoperability; Data & Model integration; Web processing Service Compatibility

Task 2.2 Derive a generic model (LMU)

- Derive the generic model
- Takes results from Task 2.1 and abstracts them
- Model will serve as input for gap analysis

Task 2.3 Opportunity & gap analysis (IMATI)

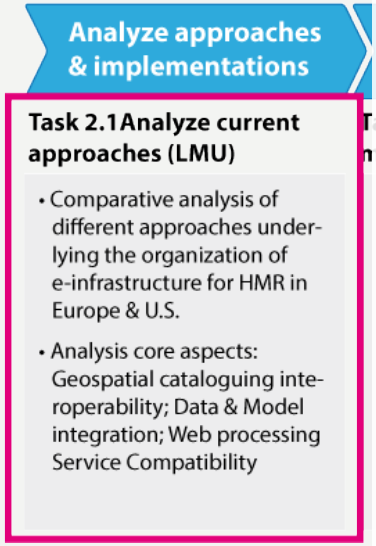
- Understand how gaps identified in the results of Task 2.1. and 2.2. may be closed or managed

Use general architectural model of Task 2.2 for gap identification

- Ensure that gaps do not become a major obstacle to interoperability
- Main expected opportunities: convergence on standards at different levels, such as those developed under the OCG umbrella or the use of OpenMI

Task 2.4 Future planning (LMU)

- Based on WP2 results (together with oters), describe the main technical elements that should be taken into account in future designs of interoperable and integrated trans-continental infrastructures



Summary

Report on the comparative analysis of the different approaches underlying the organization of e-infrastructure for HMR in Europe and in the US.

Three major sources

1. Literature screening
2. Analysis of respective project deliverables
3. Output presented at the DRIHM2US expert networking sessions of work package 3 (as reported in D3.3[1])

Analysed element – e-Infrastructure

- Technology and organisational model that support research.
- Embraces networks, Grids, data centres, collaborative environments, service registries, single-sign on mechanisms, certificate authorities, training and help-desk services.
- Seamless integration of these concepts.
- Community-specific e-infrastructures (like the ones relevant for HMR) also feature community-specific mechanisms like model coupling or data format conversions.

[1]DRIHM2US Project: Domain Expert Networking Report. Deliverable D3.3, July 2013

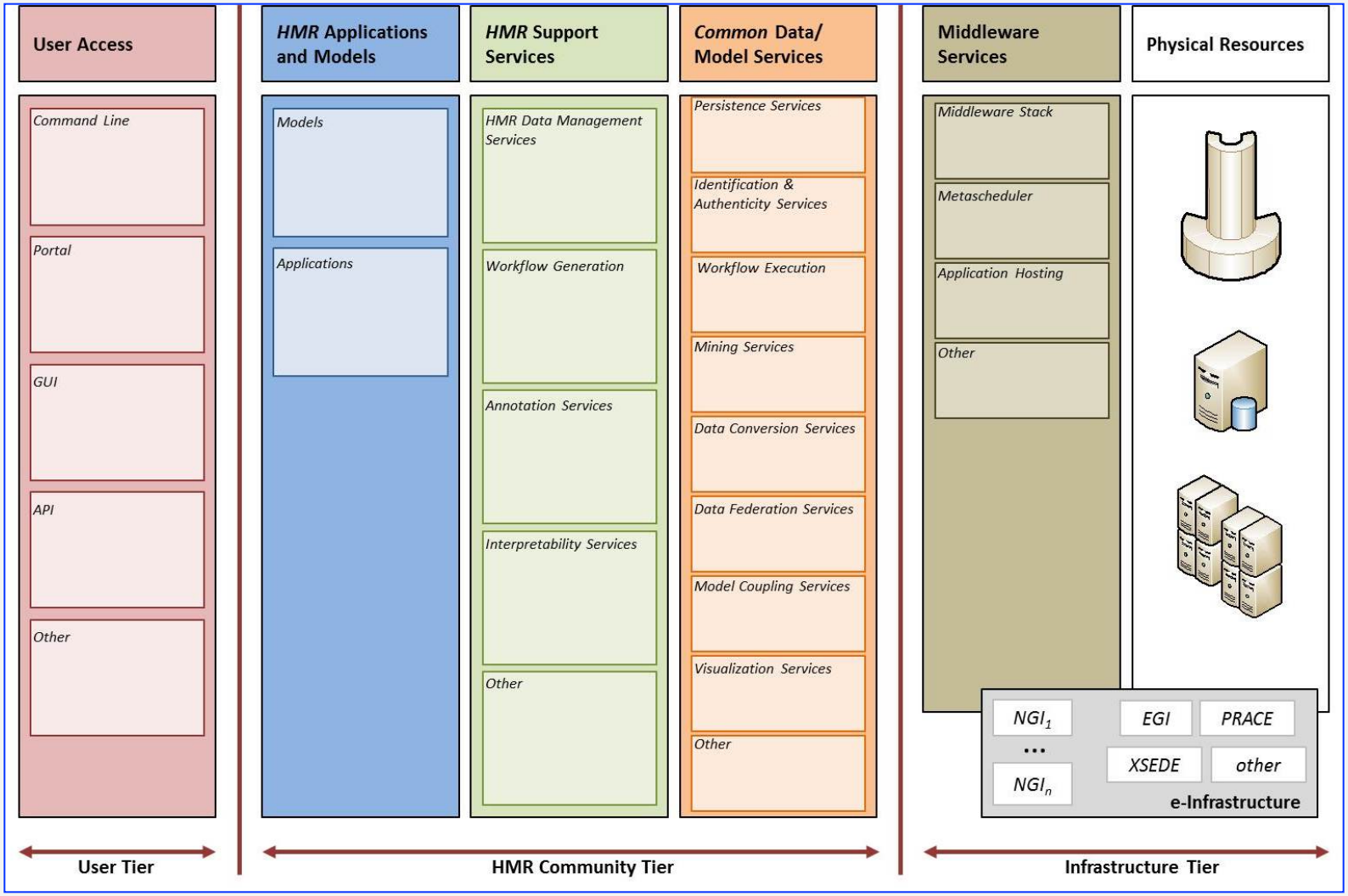
Basic observations

- e-infrastructures deal with different stakeholders, data types and services that somehow interrelate to facilitate global science processes.
- Data generators (like sensors or instruments) and users (like scientists or citizen scientists) gather, capture, transfer and process data - often across the globe in virtual research communities.
- They draw upon support services in their specific scientific communities (e.g., HMR), which typically comprise tools to locate data, process it, annotate it or interpret it.
- The support services themselves make use of a broad set of common data services including data storage, data identification, data authentication, data mining, and workflow/ task execution.

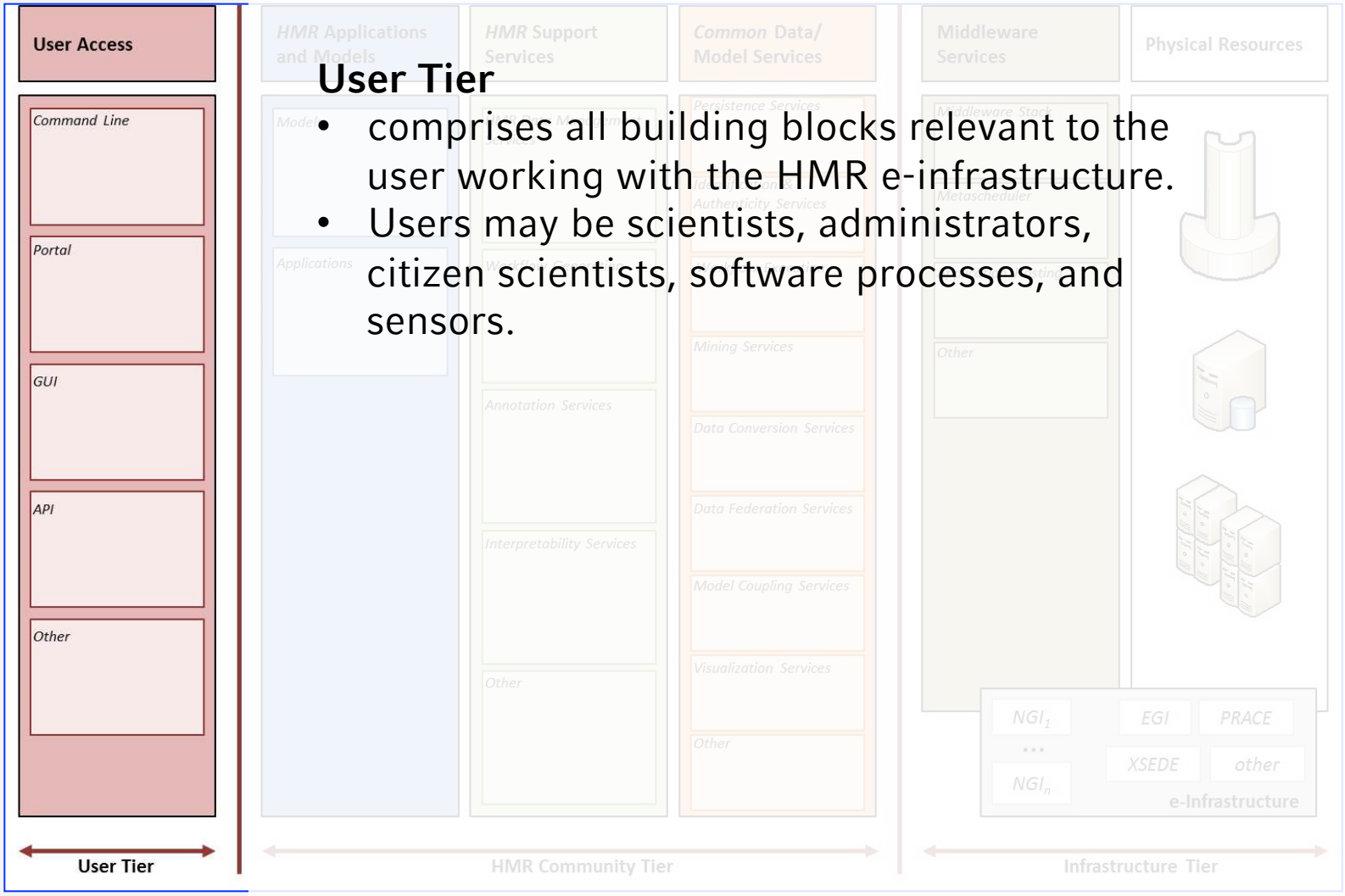
Analyzed reports

- The European Commission High Level Expert Group on Scientific Data report "Riding the Wave" from October 2010
- The European Commission report "Open Infrastructures for Open Science - Horizon 2020 consultation report" by Richard L. Hudson and Carlos Morais Pires
- The 2009 White Paper "Strategy for a European Data Infrastructure" of the European data initiative PARADE (the Partnership for Accessing Data in Europe)
- Wilkins-Diehr, Nancy, Dennis Gannon, Gerhard Klimeck, Scott Oster and Sudhakar Pamidighantam: TeraGrid science gateways and their impact on science, published in Computer, 41(11):32{41, 2008} [11]
- The XSEDE architecture ratio as described in [16]
- The Second European Union-Australia Workshop on Research Infrastructure [8]

Generic Reference Framework



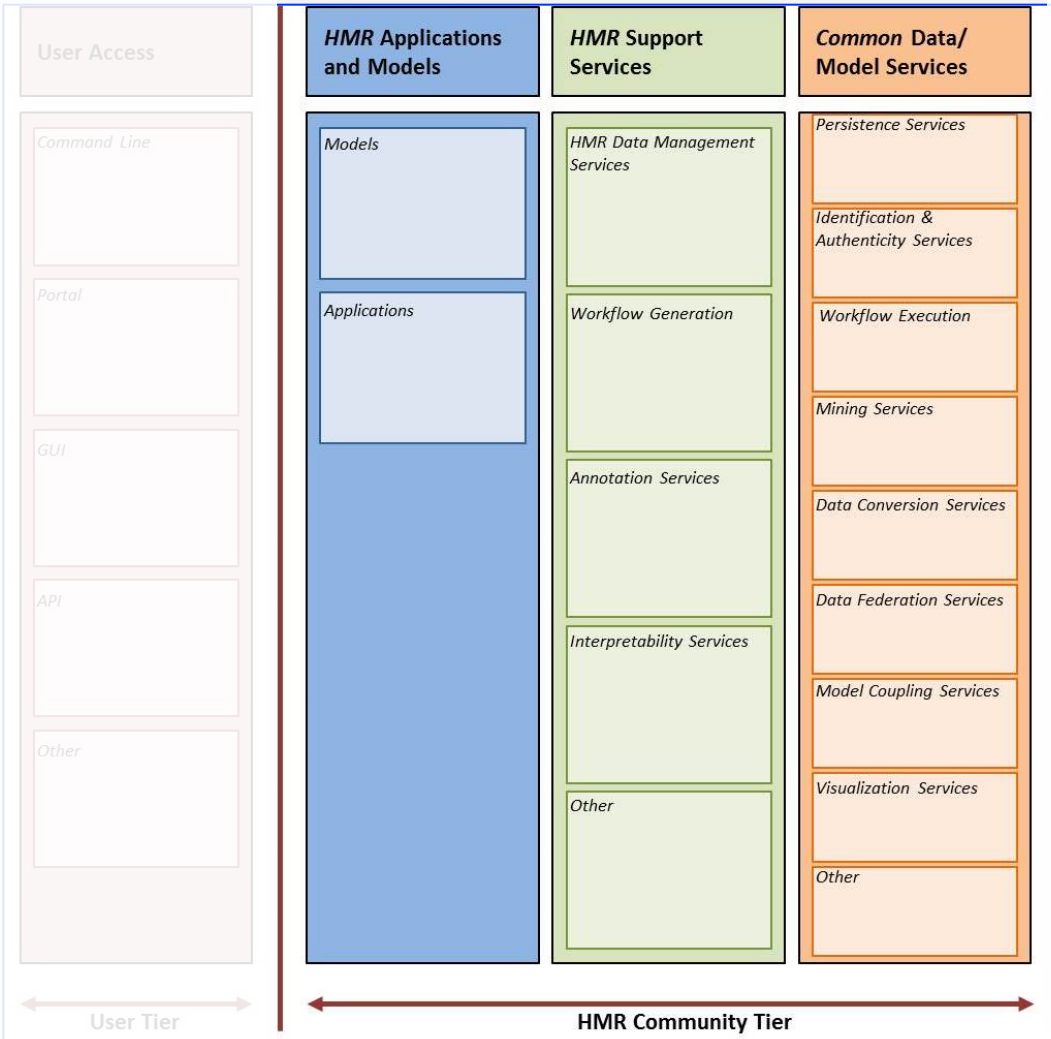
Generic Reference Framework



User Tier

- comprises all building blocks relevant to the user working with the HMR e-infrastructure.
- Users may be scientists, administrators, citizen scientists, software processes, and sensors.

Generic Reference Framework



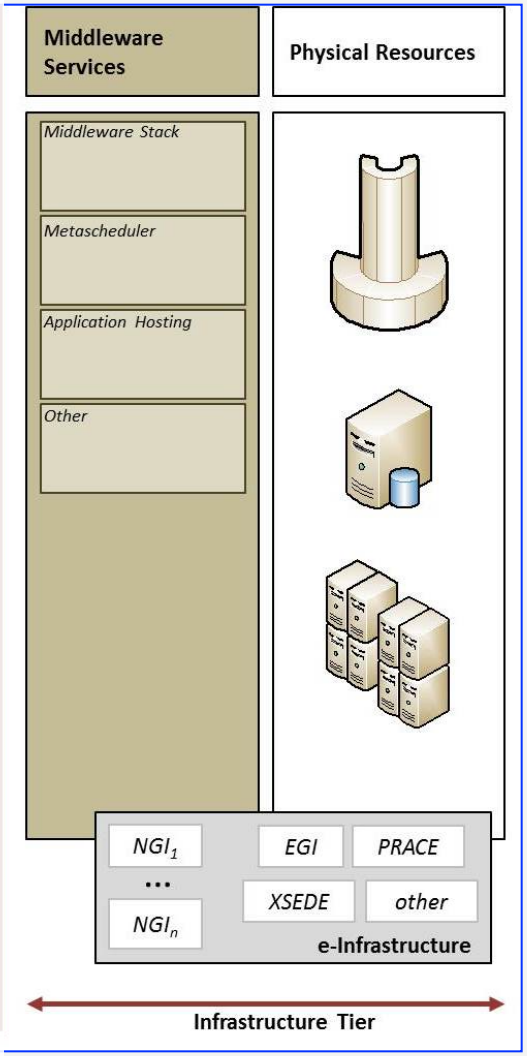
HMR Community Tier

- Reflects model and data usage patterns of HMR applications
- HMR applications draw upon HMR specific support services.
- Support services make use of common (i.e., generic) data/model services (storage, data identification, archiving, authentication, task execution, data conversion etc.)

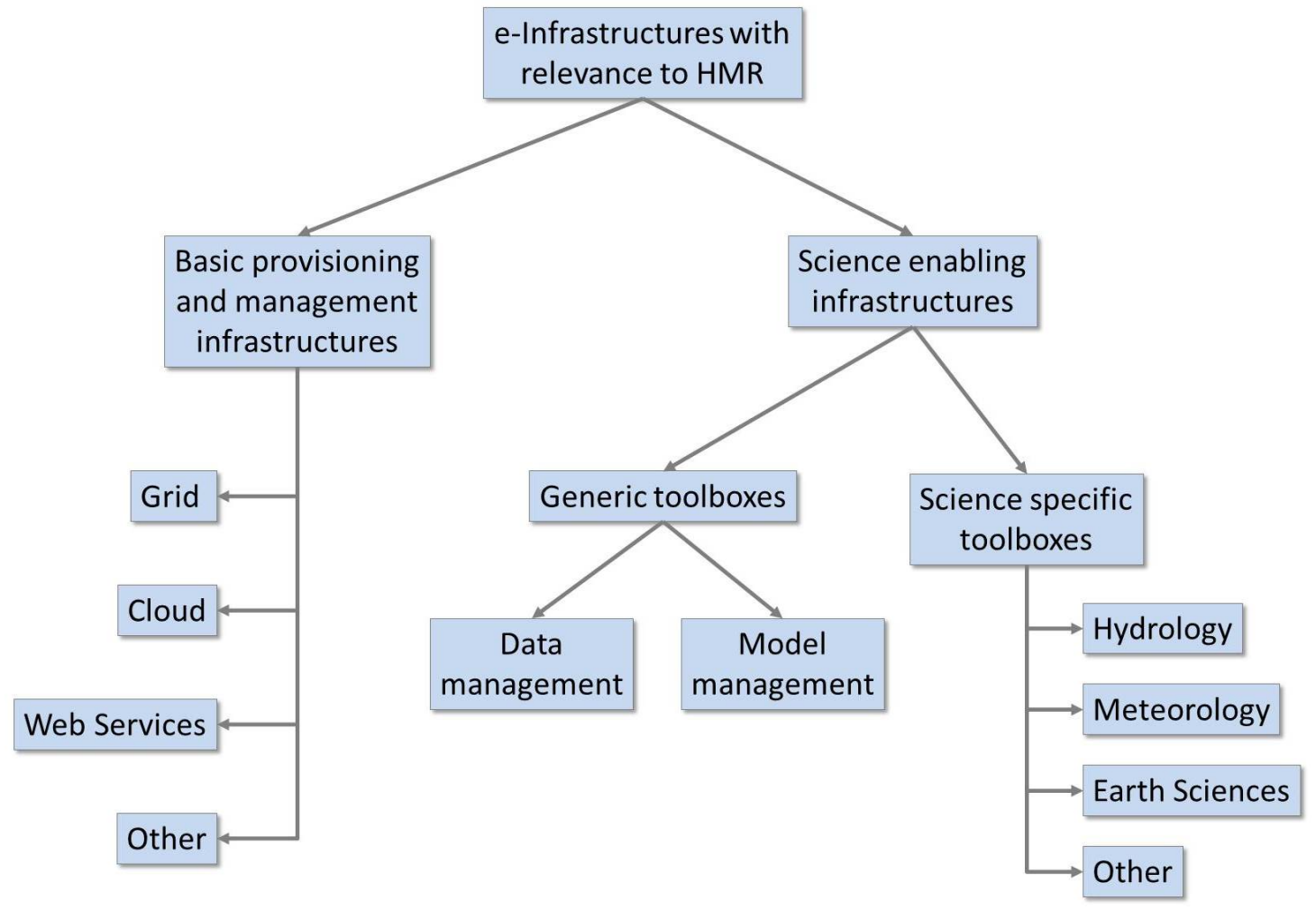
Generic Reference Framework

Infrastructure Tier

- Provides physical resources and access & management mechanisms
- Resources may in total or in parts be "contributed to" e-Infrastructures
  - the National Grid Infrastructures (NGI), to the European Grid Infrastructure (EGI), to the Partnership for Advanced Computing in Europe (PRACE), to the Extreme Science and Engineering Discovery Environment (XSEDE) or others.
- Access is typically enabled by middleware technology, meta schedulers, application hosting environments etc.



Classification to select analyzed e-Inrastructures and projects



Multiscale Applications on European e-Infrastructures (MAPPER), EU
Analysed: Multi scale service stack

Distributed Research Infrastructure for Hydro-Meteorology (DRIHM), EU
Analysed: Research infrastructure

Community Earth System Model (CESM), US
Analysed: Earth system model

Earth System Modeling Framework (ESMF), US
Analysed: Earth system model

Hydrologic Information System (HIS), US
Analysed: Hydrologic information system

Water Information Research and Development Alliance (WIRADA), Australia
Analysed: Framework

Analysis summary

		Project/Initiative					
Tier	**Category**	**DRIHM**	**CESM**	**ESMF**	**HIS**	**MAPPER**	**WIRADA**
User Access	CLI	0	0	++	0	0	0
	Portal	++	+	+	0	0	0
	GUI	+	-	0	+	+	++
	API	0	+	+	+	+	0
HMR Applications and Models	Models	*	*	*	0	n/a	*
	Applications	*	*	*	0	n/a	*
HMR Support Services	Data Mgmt.	0	+	-	+	0	+
	Services						
	Workflow						
	Generation	+	0	+	--	0	+
	Services						
	Annotation						
	Services	0	-	-	--	--	-
	Interpreta-						
	bility	-	--	--	--	--	--
	Services						

On the left framework tiers and categories

--	not available
-	very rudimentary
0	available as part of the tools or services
+	available in large parts
++	full availability
*	several
n/a	not applicable
X	main focus

Analysis summary

		Project/Initiative					
Tier	**Category**	**DRIHM**	**CESM**	**ESMF**	**HIS**	**MAPPER**	**WIRADA**
User Access	CLI	0	0	++	0	0	0
	Portal	++	+	+	0	0	0
	GUI	+	-	0	+	+	++
	API	0	+	+	+	+	0
HMR Applications and Models	Models	*	*	*	0	n/a	*
	Applications	*	*	*	0	n/a	*
HMR Support Services	Data Mgmt. Services	0	+	-	+	0	+
	Workflow Generation Services	+	0	+	--	0	+
	Annotation Services	0	-	-	--	--	-
	Interpreta-bility Services	-	--	--	--	--	--

User Tier

- Typically project/ initiative related, since it facilitates access to the HMR Community Tier.

Lacking

- a generic HMR user API
- a common HMR portal system for all HMR scientific activities over e-infrastructures.

Analysis summary

		Project/Initiative					
Tier	**Category**	**DRIHM**	**CESM**	**ESMF**	**HIS**	**MAPPER**	**WIRADA**
User Access	CLI	0	0	++	0	0	0
	Portal	++	+	+	0	0	0
	GUI	+	-	0	+	+	++
	API	0	+	+	+	+	0
HMR Applications and Models	Models	*	*	*	0	n/a	*
	Applications	*	*	*	0	n/a	*
HMR Support Services	Data Mgmt. Services	0	+	-	+	0	+
	Workflow Generation Services	+	0	+	--	0	+
	Annotation Services	0	-	-	--	--	-
	Interpreta-bility Services	-	--	--	--	--	--

HMR Community Tier

- several models & applications available.
- Large variety of HMR Support Services **available**, mostly "case-specific" but probably expandable
- HMR-specific annotation services and interpretability services are **not available**, only rudimentary.

Analysis summary

		Project/Initiative					
Tier	**Category**	**DRIHM**	**CESM**	**ESMF**	**HIS**	**MAPPER**	**WIRADA**
Common Data/Model Services	Persistence Services	0	+	-	0	--	-
	Identification and Authenticity	-	0	0	0	0	--
	Workflow Execution	+	--	+	--	0	+
	Mining Services	--	+	--	0	--	--
	Data Conversion Services	+	0	+	0	0	-
	Data Federation Services	0	0	0	0	0	-
	Model Coupling Services	0	0	0	0	++	+
	Visualization Services	0	0	+	--	0	--
	Standards Based	0	-	++	+	0	++

HMR Community Tier

- Common data/model services are **available** in various constellations, but **not generic and interoperable enough**

Analysis summary

		Project/Initiative					
Tier	**Category**	**DRIHM**	**CESM**	**ESMF**	**HIS**	**MAPPER**	**WIRADA**
Middleware Services	Grid Middleware	*	n/a	n/a	n/a	*	n/a
	Metascheduler	*	n/a	n/a	n/a	*	n/a
	Application Hosting	*	n/a	n/a	n/a	*	n/a
Physical Resources		*	*	*	*	*	*
e-Infrastructure		*	*	*	n/a	*	n/a
Community Focus	HMR	X					
	Hydrology				X		X
	Geophysics						
	General Earth Sciences		X	X			
	Generic					X	

Infrastructure Tier

- basic infrastructures (NGI, EGI, PRACE, XSEDE) are assumed to be operational or there is no dependency on such infrastructures at all.

Outcome

1. There is a growing landscape of services and tools, which individually are valuable for HMR scientists.
2. No "best HMR e-infrastructure", but best practices
3. A "Best Practices HMR e-infrastructure" can be pieced together even now.
4. HMR related e-infrastructure services and tools are a priori not interoperable.
5. There is a growing need for standardization but the standardization process is not fast enough.

Usage of Results

Analyze approaches & implementations → **Derive a general architectural model** → **Identify interoperability starting points** → **Identify gaps in analyzed approaches & implementations** → **Feed results into the ongoing developments**

Task 2.1 Analyze current approaches (LMU)

- Comparative analysis of different approaches underlying the organization of e-infrastructure for HMR in Europe & U.S.
- Analysis core aspects: Geospatial cataloguing interoperability; Data & Model integration; Web processing Service Compatibility

Task 2.2 Derive a generic model (LMU)

- Derive the generic model
- Takes results from Task 2.1 and abstracts them
- Model will serve as input for gap analysis

Task 2.3 Opportunity & gap analysis (IMATI)

- Understand how gaps identified in the results of Task 2.1. and 2.2. may be closed or managed

Use general architectural model of Task 2.2 for gap identification

- Ensure that gaps do not become a major obstacle to interoperability
- Main expected opportunities: convergence on standards at different levels, such as those developed under the OCG umbrella or the use of OpenMI

Task 2.4 Future planning (LMU)

- Based on WP2 results (together with oters), describe the main technical elements that should be taken into account in future designs of interoperable and integrated trans-continental infrastructures

Summary of D2.1

- Reports on an assessment of HMR related e-infrastructures
- Assessment is based on
  - definition of a reference framework
  - selection of reference candidates according to a coarse-grained taxonomy
- Reference framework itself has been derived from related efforts

  European Commission High Level Expert Group, the European Commission Horizon 2020 efforts, the European data initiative PARADE, the TeraGrid science gateway investigations, the XSEDE architecture ratio description, and the European Union-Australia Workshop on Research Infrastructures.

There are HMR related e-infrastructure building blocks available, but they must be combined successfully for future HMR services and applications.