IWSG 2022 Workshop

14th International Workshop on Science Gateways

15th-17th June 2022, Trento, Italy

DAY 1 - June 15, 2022		
13:00-14:30	**Registration**	
14:30-17:30	**Session 1 – Science Gateways and Research Cloud Infrastructures**	**Chair: Sandra Gesing**
14.30-15.00	Welcome messages	
15:00-16:00	Keynote 1 - Lessons learnt from building a research cloud infrastructure in Hungary	Peter Kacsuk
16:00-16:30	*Coffee break*	
16:30-17:00	Virtual Research Environments Ethnography: a Preliminary Study	Morteza Arezoumandan, Leonardo Candela, Donatella Castelli, Ali Ghannadrad, Dario Mangione and Pasquale Pagano
17:00-17:15	Everyday orchestration with Docker on Kubernetes	James Deslauriers, Resmi Arjun, Jozsef Kovacs and Tamas Kiss
17:15-17:30	Sharing Data Collections and Models for Ionosphere, Thermosphere and Plasmasphere Research	Tamas Kiss, Gabriele Pierantoni, Dimitris Kagialis, Alexander Bolotov, David Chan You Fee, Ivan Galkin and Anna Belehaki.
18:00-19:00	Evening reception	

DAY 2 - June 16, 2022		
8:30 - 9:00	**Registration**	
9:00-12:00	**Session 2 – Science gateways application to various scientific disciplines and domains**	**Chair: Sandro Fiore**
9:00-10:00	Keynote 2 - A journey towards end-to-end services for instrument based e-Science	David Abramson
10:00-10:30	Developing the German Human Genome-Phenome Archive	Kersten Breuer, Jordy D. Orellana Figueroa, Moritz Hahn, Koray Kırlı, Jens Krüger, Léon Kuchenbecker, Manikandan Ravichandran and Zehra Hazal Sezer
10:30-11:00	*Coffee break*	
11:00-11:30	Science Gateways in EOSC: The NEANIAS Visualisation Gateway	Eva Sciacca, Mario Raciti, Mel Krokos and Benjamin Kyd
11:30-11:45	The Health Data Research Innovation Gateway: A platform to discover, request access to health data and demonstrate trustworthiness in research across the United Kingdom and beyond	Susheel Varma, Paola Quattroni, Chris Milner, Neil Postlethwaite, Ruth Milne and Elaine Brannigan
11:45-12:00	GeoEDF: A Framework for Designing and Executing Reproducible Geospatial Research Workflows in Science Gateways	Rajesh Kalyanam, Jungha Woo, Lan Zhao, Carol X. Song and Jack Smith
12:00-15.00	**Social Lunch - Ristorante Antico Pozzo (Vicolo della Sat, 6)**	

15:00-18:00	**Session 3 – Science Gateways for Geoscience and Climate Change research**	**Chair: Tamas Kiss**
15:00-16:00	Keynote 3 - Large scale data challenges in the Climate domain	Sandro Fiore

16:00-16:30	An EOSC-enabled Data Space Environment for Climate Science	Donatello Elia, Sandro Fiore, Fabrizio Antonio, Guillaume Levavasseur, Paola Nassisi, Alessandro D'Anca, Sylvie Joussaume and Giovanni Aloisio
16:30-17:00	*Coffee break*	
17:00-17:15	A Community Gateway for Data Usage & Data Archive Metrics in the Climate Domain	Alessandra Nuzzo, Fabrizio Antonio, Maria Mirto, Paola Nassisi, Sandro Fiore and Giovanni Aloisio
17:15-17:30	Interactive and Flexible Environment for on-demand Climate Data Analysis	Christian Pagé, Abel Aoun, Alessandro Spinuso, Mats Veldhuizen, Ian van der Neut, Klaus Zimmermann, Lars Bärring and Phillip Kershaw
17:30-17:45	The use of the Unidata Science Gateway as a cyberinfrastructure resource to facilitate education and research during COVID-19	Mohan Ramamurthy and Julien Chastang (Cancelled)
19:00-22:00	**Social Dinner - Forsterbräu Trento (Via Paolo Oss-Mazzurana, 38)**	

DAY 3 - June 17, 2022

8:30-9:00	**Registration**	
9:00-12:30	**Session 4 – Community Building and e-Infrastructure Challenges**	**Chair: Christian Pagé**
9:00-10:00	Keynote 4 - The reasons why the Science Gateways Community needs an institute	Sandra Gesing
10:30-10:30	Distributed but Integrated: Forming a Science Gateway from Multiple Parts	Olaf Brandt, Holger Gauza, Jan Kaltenbach, Steve Kaminski, Jens Krüger, Fabian Paz, Saker Halima, Fabian Wannenmacher, Johannes Werner and Thomas Zajac

10:30-11:00	*Coffee break*	
11:00-11:30	A dynamic and extensible web portal enabling the deployment of scientific virtual computational environments on hybrid e-infrastructures	Marica Antonacci, Daniele Spiga, Davide Salomoni, Diego Ciangottini, Vincenzo Ciaschini, Alessandro Costantini, Giacinto Donvito, Doina Cristina Duma, Federica Fanzago, Emidio Giorgio, Alessandro Italiano, Massimo Sgaravatto, Vincenzo Spinoso, Stefano Stalio, Mirco Tracolli and Marco Verlato
11:30-12:00	Design of a Decentralised Autonomous Scientific Publishing Organisation	Sandi Gec, Petar Kochovski, Jovan Buragev and Vlado Stankovski
12:00-12:15	An Industry Gateway for Rapid Reconfiguration of Manufacturing Processes	Tamas Kiss, Gabor Terstyanszky and Resmi Arjun
12:15-12:45	Panel on “FAIR principles and Science Gateways”	
12:45-13:00	Closing remarks	

IWSG2022 Keynotes

June 15	**Speaker*: Peter Kacsuk*** **Title*: Lessons learnt from building a research cloud infrastructure in Hungary***

Abstract: An ambitious plan to build a state-of-the-art, high quality, reliable and robust cloud infrastructure for the researchers of the Hungarian Academy of Sciences was formulated in 2015. This was the beginning of a long road that led to the creation of a cloud infrastructure that now serves all the Hungarian researchers and even researchers in companies. In fact, since 2015 we can distinguish three phases of building the cloud infrastructure and every phase significantly extended the usability of the cloud and opened it up for wider and wider user communities. This was a real success story and the goal of my talk is to share with the audience the experiences that we collected during this three phases and also to explain the lessons learnt in the different phases and how they helped us to gradually improve our cloud and make it a widely used and very popular research infrastructure in Hungary.

The first phase between 2015-2019 resulted in the creation of a reliable and robust research cloud that was good enough to serve the basic requirements of the researchers of the Hungarian Academy of Sciences. However, by the end of 2019 it became clear that the capacity became insufficient and furthermore the original objectives of the research cloud should be extended concerning the type of services we want to provide and the target communities we would like to serve. As a result, the second phase was started in 2019 with new objectives that reflected the new requirements like supporting big data and AI applications in a much more comprehensive way than it was before. The target user community also should have been significantly extended and the level and quality of user support should have been further improved. One of the major achievements of the second phase was the creation of a large set of reference architectures that can significantly reduce the user effort to create complex software systems in the cloud. The second phase was successfully completed in February 2022 and recently we entered to the third phase where the two major objectives are to provide intensive support of building data repositories in the cloud and integrate our cloud with the three major European cloud infrastructure initiatives: EGI Fedcloud, EOSC and SLICES.

During the talk objectives, implementations, results and lessons learnt for each phase of the Hungarian research cloud program will be explained in order to show the way how to build large-scale and complex research infrastructure and how to avoid pitfalls and other potential problems that are very frequent in running large-scale infrastructure programs.

June 16	**Speaker*: Sandro Fiore*** **Title*: Large scale data challenges in the Climate domain***

Abstract: This talk will highlight key challenges regarding climate data management at scale discussing in particular some case studies of great relevance like the CMIP experiment and the associated Earth System Grid Federation infrastructure.

In such a context, over the last few years, a novel concept of data space has been progressively introduced, revealing to be very attractive and suitable for scientific domains too. The talk will also show potential directions for such a concept in the climate community by focusing on the (i) the central role of data (“data value”, “data strategy”, “data economy”) in public and private sectors, (ii) a more effective and sustainable approach to data sharing and re-use, (iii) trustworthy data repositories, (iv) a pervasive adoption of FAIR data principles, and (v) a

holistic approach where legislation, governance, standards, services and infrastructure will synergistically contribute to enable data handling at all levels.

June 16 **Speaker*: David Abramson***
Title*: A journey towards end-to-end services for instrument based e-Science*

Abstract: As data and computation assume increasing importance in modern science there is a need for integrated infrastructure that supports both basic and sophisticated workflows. To date there has been progress at building platforms that support part of this vision, however, there are often gaps that require manual intervention by trained users. In this seminar I describe the state of infrastructure at the University of Queensland, a work-in-progress journey aimed to realise the vision to support researchers with varying computing skills. We describe some of the components, namely a meta-data management and provisioning system called UQRDM; a sophisticated data storage platform called MeDiCI;, a novel instrument interface called CAMERA that facilitates processing on multiple compute platforms with minimal data duplication; domain specific data management repositories based on the OMERO, XNAT and CLOWDER tools; and multiple high throughput image processing portals called UQIPP and NIMROD. When used together these services allow a researcher to provision storage, capture data on a range of instruments, process it using multiple HPC and Cloud platforms; and publish the results for wider distribution. We will Illustrate the infrastructure by describing its use in multiple bioscience applications.

June 17 **Speaker*: Sandra Gesing***
Title*: The Reasons Why the Science Gateways Community Needs an Institute*

Abstract: The science gateways community is by nature an interdisciplinary community with quite a few different roles: from researchers and educators to creators and providers of science gateways serving research areas in the sciences, humanities and arts. While 80-90% participants at the European IWSG (International Workshop on Science Gateways) and the US Gateways Conference series are predominantly from computer science and engineering, the number of users belonging to the science gateway community is much bigger in research domains beyond these “usual suspects”. Additionally, open science and FAIR (Findable, Accessible, Interoperable, Reusable) initiatives as well as research involving machine learning and artificial intelligence (AI) methods have a lot of momentum in the research community in the last seven to eight years and demand solutions that are tailored to the different communities and allow researchers to focus on their research questions. Science gateways are predestined to serve as a solution with their design for supporting the sharing of simulations, data and workflows while applying research infrastructures including sensors and lab instruments. The open science, FAIR and AI communities grow fast, evident in initiatives such as GO FAIR in Europe and the US, publications on FAIR and AI and the amount of presentations at conferences such as RDA (Research Data Alliance) while the science gateways community seems to be not participating in this fast growth. There are some grassroot efforts that address science gateways and FAIR, for example, but the uptake is not comparable. The talk goes into detail why an institute is necessary to improve the outreach to the community at large to accelerate research and reach researchers and educators who are not aware of science gateways yet.