

e-IRG – Magazine

Magazine

Fall 2018

Highlights from the delegates meetings

Since the publishing of the previous edition of this Magazine, there were two delegates meetings.

Page 4

Serving the user base

Digital Infrastructures for Research – DI4R2018 report

Page 3



e-IRG Knowledge Base

Provides answers to all your e-Infrastructures questions

Page 2

European Open Science Cloud

will give researchers access to e-Infrastructures and services

Page 8



Interview Ivan Dimov

Deputy Minister of Education and Science, Bulgaria

Page 20

ESFRI's 5th Strategy Report & Roadmap 2018

Report of the launch event in Austria.

Page 9



e-IRG Workshop Sofia

Nearly 100 participants attended the e-IRG workshop organised in the framework of the Bulgarian EU Presidency on 14-15 May, 2018 in Sofia, Bulgaria. The workshop had three main topics: challenges resulting from Big Data for HPC; progress on the development of the European Open Science Cloud; and use of European Structural Investment Funds for e-Infrastructures in the Southeast European and Mediterranean region.

Big Data for HPC, EuroHPC, European Open Science Cloud (EOSC) progress, and European investment for e-Infrastructures in the Southeast and Mediterranean region were the main topics addressed at the e-IRG Workshop.

Read further at Page 15 >>



Vi-SEEM virtual research environment emerges as a mini-EOSC for Southeast European and Eastern Mediterranean scientists

Read further at Page 14 >>

National Nodes as the foundation for the European Open Science Cloud

Read further at Page 21 >>

Special Edition of Open e-IRG Workshop

under the Austrian EU Presidency will focus on relationship between Open Science, FAIR data and EOSC

The public Open e-IRG Special Edition Workshop under the Austrian EU Presidency will take place on 20-21 November 2018 in Vienna, Austria, focusing on e-Infrastructures in the broader context of the development of Open Science policies and the aspect of national nodes as foundation for the European Open Science Cloud (EOSC).

The e-IRG Special Edition Workshop will be co-organised by the Austrian Federal Ministry

of Education, Science and Research and the Library and Archive Services of the University of Vienna in the framework of the Austrian EU Presidency, the Second EOSC Stakeholders Forum on 21-22 November 2018 at the Austria Center Vienna and the Launch of the EOSC Governance Structure on 23 November.

The first session in the afternoon of 20 November will address the e-Infrastructures in the broader context of the development of Open Science policies. The invited speakers are:

Professor Barend Mons from the Human Genetics Department at the Leiden University Medical Center, The Netherlands, and Director of GO FAIR International Support and Coordination. His presentation is titled "A Global Challenge: The GO-FAIR Initiative - Beyond Regional Open Science Initiatives".

Professor Roberto Barbera, Department of Physics and Astronomy of the University of



Catania, Italy. His talk will be about FAIR in Ethiopia: "Towards a common Vision on Open Science: FAIR-Data Culture and National Data Infrastructure in Ethiopia".

Ms. Rahel Fidel, Ethiopian Ministry of Science and Technology will speak about the African Open Science Cloud.

The second session in the morning of 21 November will bring the aspect of national nodes as foundation of the European Open Science Cloud on the table and present an analysis of the current status in the different EU countries, along with the development of recommendations and good practices towards national e-Infrastructure Commons in order to ease integration at EU level. ❖

<http://e-irg.eu/e-irg-special-edition-nov-2018>

e-IRG Knowledge Base

Provides answers to all your e-Infrastructures questions

How much is invested in European e-Infrastructure projects? Which European country participates in most international organisations? The e-IRG Knowledge Base provides the answers.

e-Infrastructures provide basic ICT infrastructure to many European science initiatives. e-Infrastructures are complex, as are the funding and the political decision processes. To facilitate these and to provide insight into the complex e-Infrastructure field, the e-IRG Knowledge Base collects lots

of data about e-Infrastructures and makes it available to the community. The Knowledge Base is about facts and figures. The e-IRG Knowledge Base is also about collecting data that is openly available. It combines data from many sources, links data from many sources and presents it in an easily accessible way.

The e-IRG Knowledge Base exists now for over a decade and has seen several upgrades and major changes. The most recent one just got online.

What can you find in the Knowledge Base? The browsing interface at <http://knowledgebase.e-irg.eu> provides major entries into:

- Funded projects under Horizon 2020
- Information organised per country

Part of the core data for the H2020 projects is taken from the European Open Data portal. It is enriched with information from other sources.

The country information related to e-Infrastructures collects data on participation of the country in international organisations, information about the national network infrastructure (NREN) and supercomputer centres.

Up to now more than 600 million Euro funding from H2020 has been invested in

e-Infrastructure related projects. We provide an overview of the top organisations that are most involved (measured by the number of projects) in e-Infrastructures. We do the same for HPC. You may notice some overlap: some HPC projects are also e-Infrastructure projects. In HPC about 500 million Euro has been allocated to active H2020 HPC related projects. ❖

<http://knowledgebase.e-irg.eu>

Rank	Organisation	Number of H2020 e-Infrastructure related projects
1	CNRS (Centre National de la Recherche Scientifique)	20
2	CNRS (Centre National de la Recherche Scientifique)	20
3	European Molecular Biology Laboratory	18
4	European Organization for Nuclear Research	15
5	GRNET S.A. (GRNET)	15
6	Barcelona Supercomputing Center	15
7	Forschungszentrum Jülich (FZJ)	15
8	Science and Technology Facilities Council (Daresbury Laboratory)	14
9	PSNC (Instytut Chemii Bioorganicznej) Polak	14
10	UEDN	11
11	Agencia Estatal Consejo Superior de Investigaciones Científicas (CSIC)	11
12	Stichting Egis (EGIS)	11
13	Commissariat à l'Énergie Atomique (CEA)	11

Serving the user base

Digital Infrastructures for Research – DI4R2018

Conference report

The third edition of the annual Digital Infrastructures for Research (DI4R) conference was held this year at the University of Lisbon campus in Lisbon, Portugal from 9 October to 11 October 2018. The DI4R 2018 conference was jointly organised by EOSC-hub, GEANT, OpenAIRE, and PRACE under the theme "Challenges for Research Communities around Open Science".

This event provided a packed programme of sessions, posters and networking opportunities - leading to brainstorming and discussions on how best to serve the user base. DI4R 2018 has showcased the policies, processes, best practices, data and services necessary for the support of the research process to ultimately advance scientific knowledge through open science.

More than 370 researchers, innovators, data producers, scientific domain experts, librarians, data science practitioners, service providers, project leaders, policy makers and funders from 38 countries, participated and contributed with their voice to DI4R 2018. Following the previous two DI4R events – DI4R2016 in Krakow and DI4R2017 in Brussels, the conference was an opportunity to report and discuss requirements, existing solutions and success stories that highlight the relevance of open science and its enabling building blocks. By bringing together many different stakeholders, the conference has fostered cooperation between infrastructure projects, communities and initiatives.

In the 2018 edition, the Programme Committee was chaired by Sinéad Ryan from Trinity College and Volker Guelzow from the Deutsches Elektronen-Synchrotron (DESY).

The main topics of interest were taken from the following areas:

Area 1: Cross-Domain challenges / Data exchange across domains: researchers, technologist

Area 2: Open Science and skills

Area 3: Computing and Virtual Research Environments

Area 4: Security, trust and identity



Area 5: Digital Infrastructures for the European Open Science Cloud (EOSC) and/or the Electronic Data Interchange (EDI)

Area 6: Business models, sustainability and policies

Area 7: Innovation in Open Science with Small and Medium-sized Enterprises (SMEs) and Industry

The work of the e-IRG delegates and the e-IRGSP5 team were presented in two talks. On 10 October, in the session "EOSC from Theory to Practice", Professor Sverker Holmgren presented "Analysis of National Nodes as foundation for the European Open Sciences Cloud". The scope of the

presentation was to present the preliminary analysis of the requested information concerning issues such as the coordination of national horizontal e-Infrastructures, their governance model, their funding, and the access policies currently in place along with a first set of recommendations for the different stakeholders, namely e-Infrastructure providers, funders, policy makers and users and get some initial feedback.

On 11 October, in the session "Innovation for Open Science with SMEs", Fotis Karayannis presented "Towards a common approach on Key Performance Indicators (KPIs) from e-Infrastructure" in which he emphasized that Key Performance Indicators will play an important role in monitoring the development of projects and services of research infrastructures and e-Infrastructures, and their commitment to the principles of Open Science.

Strong discussions arose in the following parallel sessions: "EuroHPC a new player in the game - PRACE/GEANT", "Thematic Services", "EOSC Service Architecture", "Open Science Services", "Data Management Services", and "OpenAIRE services for Research Infrastructures" which were of great interest for all participants at DI4R. ❖

More information and the full programme of DI4R2018 is available online at:

<https://www.digitalinfrastructures.eu/>

Todor Gurov, e-IRGSP5 project.

	Project acronym	Project Name	URL (likely website too)	Start Date	End Date	Budget in Euro
1	AARC	Authentication and Authorisation for Research and Collaboration (AARC)	http://aarc-project.eu/	2015-05-01	2017-04-30	2.950.060,00
2	AARC2	Authentication For Research and Collaboration		2017-05-01	2019-04-30	2.999.892,89

Highlights from the delegates meetings

Four times a year, the e-IRG delegates, representing their countries gather for a working meeting. These meetings are organised in the country that at that moment holds the Presidency of the European Union. Since the publishing of the previous edition of this Magazine, there were two meetings - in Sofia - from which the minutes are already available. Although the meetings are only open for delegates to attend, there are public minutes published of each meeting.

These are available on the e-IRG web site. Because these delegates meetings are typical working meetings, a lot consists of reporting and discussing ongoing work. Hence when one meeting is finished, the results of the previous meetings are not that interesting anymore, as work has progressed. Here, we will summarise the highlights of the Bulgarian delegates meetings.

52th Delegates Meeting

The-IRG policy document on e-Infrastructure Commons & national nodes and European Open Science Cloud Roadmap in the picture.

The closed 52nd e-IRG Delegates Meeting took place on March 20, 2018 at the Best Western Expo Hotel in Sofia, Bulgaria, during the Bulgarian EU Presidency. A proposal with the main directions for a new focus of the e-IRG was presented and discussed.

Part of the meeting was spent on the next main e-IRG policy document on e-Infrastructure Commons National Nodes.

The two new permanent members in the e-IRG Executive Board were announced, and the proposed changes to the bylaws were discussed.

The current status of the three Working Groups, namely the Overarching Working Group, the Long-Term Preservation of Research Data (LT PRD) Working Group and the Conflict of Interest Working Group - was reported.

There was a discussion about the updates on the European Open Science Cloud (EOSC) Roadmap and FP9. The countries reported important national news and the next

activities under Bulgarian and Austrian EU Presidency were presented.

e-IRG Chair Gabriele von Voigt opened the meeting. She thanked Aneta Karaivanova and the rest of the Bulgarian team for arranging the meeting in Sofia.

Ivan Dimov, Deputy Minister for Education in the Ministry of Education and Science in Bulgaria, on behalf of the Bulgarian EU presidency and the Chair, welcomed the delegates. Ivan Dimov stated that the research funding in Bulgaria is 4,6% of the GDP.

The government published the strategy for research with 4 priority areas, one of which is ICT. They also give appropriate attention to Research Infrastructures, including e-Infrastructures, based on their national roadmap. They have the largest supercomputer power in the South-Eastern European (SEE) region with the recent HPC centres, since 2015, participating in e-Infrastructure, HPC, Grid, Cloud and Big Data initiatives. Ivan Dimov mentioned that the Bulgarian Commissioner Mariya Gabriel is responsible for the Digital Agenda. Bulgaria is a member of EuroHPC. The declaration was signed in October 2017. One of the Bulgarian EU presidency priorities was digital connectivity of the Western Balkans.

The key related ICT events of the presidency were the flagship conference on Research Infrastructures, the HPC for Extreme Scale Scientific and Industrial Applications on April 19, the e-IRG open workshop on May 14-15 and finally the regional event on e-Infrastructures for excellence science on May 15-16, 2018.

The Chair introduced the new nominations and re-nominations: Danko Lupi and Branislav Zimonjic for Bosnia and Herzegovina; Ivan Maric for Croatia; Veronika Macková and Jan Gruntorad for the Czech Republic; Irina Kupiainen and Sami Niinimäki for Finland; Panos Argyrakis and Vasileios Gongolidis, alternating with Maria Gkizeli for Greece; Norbert Meyer and Mariusz Sterzel for Poland; João Ferreira and João Pagaime for Portugal; and Alison Kennedy and Simon MacIntosh-Smith for the UK. The Chair also asked the new members who were present to introduce themselves. Annika Philipps, German Federal Ministry for Education and Research is to be nominated as German delegate for the next meeting due to the German elections. She is currently observer. Mariusz Sterzel is involved in Academic Computing in Poland and has a background in chemistry.

New focus of e-IRG

Some important points come up with the new focus and position of e-IRG towards the main e-Infrastructure developments. e-IRG should be dealing with both short- and long-term issues of e-Infrastructure policies. With

regard to the European Open Science Cloud, e-IRG should be synthesising and analysing national views, providing recommendations to the European Commission, Member States and Associated Countries with a neutral focus in governance. In Data, the Research Data Alliance and GoFAIR, e-IRG should be more active and liaise with relevant groups. e-IRG, the Member States and Associated Countries should work closely together and exchange best practices & experiences regarding user needs, EOSC national nodes, and national roadmaps. e-IRG should continue to play a role for HPC. e-IRG should continue to work together with ESFRI and Research Infrastructure projects as well. As for FP9 (Horizon Europe), e-IRG should strive to have influence in the development of the Framework Programme and work programmes.

Next main e-IRG Policy Document that has to be worked on is the *e-Infrastructure Commons National Nodes*.

Sverker Holmgren gave an introduction, presenting the three main steps of the e-IRG e-Infrastructure Commons and the decision of e-IRG. Step 1 consists of the e-IRG White Paper 2013. This White Paper described the urgency to build a European e-Infrastructure Commons that is open and accessible; continuously adapting to the changing needs; open to new technological opportunities; and able to meet the challenges of implementing the EU's 2020 Strategy. The three distinct core functions of the e-Infrastructure Commons are community building including high level strategy and coordination in one single organisation with a central role for user communities; service provisioning with a flexible, open, and competitive approach to national, European, and global service provision with advanced collaboration among the interested public and commercial service providers; and innovation by implementation of major innovation projects through the best consortia including e-Infrastructure suppliers, industry, users and academia.

Step 2 consists of the e-IRG Roadmap 2016. The goal described in this document is to turn the vision of the e-Infrastructure Commons into reality by 2020. The Research Infrastructures and research communities should be encouraged to elaborate on and drive their e-infrastructure needs. This was started by intensified collaboration with ESFRI. Governments and funding agencies should be encouraged to embrace e-Infrastructure coordination at the national level and construct strong national e-Infrastructure building blocks.

Step 3 consists of the e-IRG Roadmap 2018. This document has to address the national nodes. It was decided at the 51st Delegates Meeting that e-IRG's highest priority should be given to a document dealing with the views from Member States and Associated

Countries on national infrastructures. The following points should be dealt with as well: Go-INFRA, Research Infrastructures and e-Infrastructures, Member State and Associated Countries answer towards EOSC, and update of experiences from the e-IRG-ESFRI collaboration.

Two parallel sessions were organised. The main focus of the document is to come up with a landscape analysis on the status of the e-Infrastructure Commons in the different countries, understanding the dynamics and evolution in the countries, and also do synthesis and analysis work, and come up with clusters - groups of countries with similar approaches, best practices and recommendations for optimal cooperation and coordination. A Working Group was formed to act as the editorial board with the support of the e-IRGSP5 and co-chaired by Sverker Holmgren and Arjen van Rijn, and staffed by e-IRG Chair Gabriele von Voigt, Ulrike Jaekel, Josva Kleist, and Laurent Crouzet. Each e-IRG country delegate is responsible to get the state of his country to be documented in an annex.

The document should be ready in around a year's time.

Extended Executive Board

During the 51st Delegates Meeting it was decided to extend the e-IRG Executive Board by two permanent members. As three candidates have been nominated, an election process took place. The two members elected in the e-IRG Executive Board for two years, until Spring 2020, are Sverker Holmgren and Arjen van Rijn.

Overarching Working Group

Arjen van Rijn reported about the current work on the Overarching Working Group (OA WG) composed of all e-IRG representatives in the ESFRI Strategy Working Groups and the Implementation Group. The hearing of all ESFRI Roadmap 2018 proposals has been finished, along with the harmonization meeting, and the e-IRG OA WG was invited by ESFRI to give input to the evaluation of this process. This has to be discussed at the ESFRI meeting. The deadlines for contributions were end of March for running Roadmap projects, and end of April for new Roadmap proposals, so the OA WG had to come up with comments. The OA WG was expected to contribute to the Landscape Analysis. Notably, not all SWGs have asked for e-Infrastructure landscape analysis inside their group.

Long-term Preservation of Research Data

Rosette Vandembroucke reported on the status of the Long-term preservation of Research Data (LT PRD) document. She reminded that the document was initiated at

an ESFRI Forum in Brussels where she presented e-IRG, and that it was conceived as a Guidelines document for Research Infrastructures. She presented the current status of the document. The document is still a draft, missing the legal aspects section. Preservation is also linked with openness and "FAIR"ness.

Update on the European Open Science Cloud and FP 9

A so-called EC Staff Working Document (SWD) entitled "Implementation Roadmap for the European Open Science Cloud" has been published. The EC SWD proposes a process to get to EOSC. It executes what the EC Communication 2016 promises, reflects on the EC Work Programme 2018-2020 Horizon 2020 programming actions, based on the EOSC Declaration, that was published in June 2017. It contains the Implementation



Roadmap for EOSC along six action lines. Some other and related next steps in 2018, were the following:

April 25th: Revised recommendation on Access to and Preservation of Scientific Information and the revised PSI Directive.

May 28th-29th: Competitiveness Council meeting and Conclusions

June 11th: EOSC 'coalition of doers' Summit 2018 to provide a platform to voice the Council Conclusions on EOSC and to consult with the 'coalition of doers'.

September: Publication of the final draft FAIR Data Action Plan by FAIR data Expert Group

November-December: Launch of the EOSC Governance Structure under the Austrian Presidency. Publication of the recommendations by FAIR data Expert Group. Publication of the recommendations by EOSC High Level Expert Group

A set of action lines for the implementation of the EOSC Roadmap were also prepared, namely architecture, data, services, access & interface, rules and governance. The possible design of the EOSC Governance framework

as resulted from the consultations is also included in the document.

A discussion followed on the positioning of e-IRG in the Governance, taking into account the recent decision that e-IRG should not be directly involved. Based on this, the most appropriate block for e-IRG is the advisory layer, called "Stakeholders Forum", which should be straightforward to be included. e-IRG may examine whether an engagement in the "Working Groups" is appropriate. It was noted that the WGs will have more power than the advisory later.

It was agreed that e-IRG should react to the document, providing a statement. The e-IRG comments were to be provided in time for the Competitiveness Council meeting on 29th May, given the fact that on the 11th of June the European Commission would present the decision from 29th of May. Although there was little time for reaction, several delegates had read the document and expressed their satisfaction. Regarding governance, it is less overambitious and less top-down.

It was felt that e-IRG should be supporting it and make a statement that EOSC should be built on a federation of strong national building blocks. It would be good if e-IRG could also come up with a response on clear separation of funders, communities and providers. In other words, start with compliments and then make a statement.

Other suggestions were to be more open, beyond RTD and Connect and come up with a concise statement. Recommendations on the governing bodies were also to be included. More detailed comments referred to the frequency of the EOSC board meetings: 1 time per year is not enough, it should be at least 3 times. Also, more details about the Working Groups mode of work and meetings are needed. It was also stated that the above points were also relevant for the joint meeting between the Executive Boards of e-IRG and ESFRI.

In case some of the ministerial representatives who are also e-IRG delegates, are nominated in the EOSC Board, it should be clear that they represent their ministries or countries and not e-IRG.

A first quick and short reaction to the Strategy Working Group on the EOSC Roadmap Implementation had to be prepared by the e-IRG extended Executive Board. It had to be a positive reaction, providing a statement that EOSC should be built on a federation of strong national building blocks. More comments on the governance and other points had to be included. ❖

Full public summary of the meeting is available at:
<http://e-irg.eu/e-irg-meeting-mar-2018>

53th Delegates Meeting

European Open Science Cloud and EuroHPC developments highlighted and active work overview of National Nodes and Long-Term Preservation of Research Data documents

The closed 53rd e-IRG Delegates Meeting took place on May 16, 2018 at the Grand Hotel Sofia in Sofia, Bulgaria, during the Bulgarian EU Presidency.

e-IRG Chair Gabriele von Voigt, Germany, opened the meeting and welcomed the delegates. She thanked Aneta Karaivanova and the rest of the Bulgarian team for arranging the events in Sofia.

The Chair stated that there were no new nominations or re-nominations since the last Delegates Meeting in March. However, she reminded that Leif Laaksonen has left e-IRG as he will be retiring this year, and that Irina Kupiainen has been nominated to replace his seat. The Chair expressed her appreciation for Leif Laaksonen's work for e-IRG, including his term as a Chair and the Delegates acclaimed.

The Chair welcomed Augusto Burgueño-Arjona from the European Commission (EC). Augusto Burgueño-Arjona had envisaged a conversation about the Staff Working Document (SWD) on the Implementation Roadmap on the European Open Science Cloud (EOSC). As e-IRG has already commented on the SWD, Augusto Burgueño-Arjona said that he would be happy to answer any questions or remarks, following the latest development, which he did. The Chair thanked Augusto Burgueño-Arjona for his presence and detailed answers. Summing up, the Chair stated that there is a lot of work to be done, much more than just the technical business, rather on policies.

Update on EuroHPC

The Chair stated that two main points would be covered, namely a report on the Sofia April HPC event and the Member States consultation on EuroHPC regulations.

Report from the April HPC event in Sofia

Peter Brönnimann, Switzerland, stated that the title of the event was "Shaping Europe's Digital Future: HPC for Extreme Scale Scientific and Industrial Applications". The event took place in Sofia on 19 April 2018. Peter Brönnimann noted that for once the focus was on applications and how HPC can provide solutions for a wide range of societal and economic challenges. Mariya Gabriel, European Commissioner for Digital Economy and Society, presented the EuroHPC Joint Undertaking as a strategic initiative for Europe. She defended EuroHPC, setting goals

on the HPC infrastructure to achieve their full potential.

Aneta Karaivanova, Bulgaria, added that it was a very inspiring event, including the future developments on HPC with quantum accelerators, which after several years will be a reality. She referred to the panel on South Eastern Europe, which was staffed with mostly the same persons as the ones in the e-IRG workshop and that there was a second panel on biomedical applications.

Norbert Meyer presented an update from the developments in the last quarter of the current state of the EuroHPC Joint Undertaking including the latest statutes and the proposed Council regulation.

The Joint Undertaking is tasked to select the hosting entity of the petascale and pre-exascale supercomputers, in a fair, open and transparent manner. HPC resources will be available for commercial purposes but not for free. International access is also possible. Regarding the voting rights of members, initially they will be equally distributed to participating states and later based on the annual contribution of the public members. Regarding governance, the European Commission, as part of its role in the Governing Board, shall seek to ensure coordination between the activities of the Joint Undertaking and the relevant activities of Union Funding Programmes. The envisaged Advisory Groups are the Research and Innovation Advisory Group and the Infrastructure Advisory Group. Other groups of experts can be referenced.

Peter Brönnimann stated that the Joint Undertaking, composed of the EC, Member States and industry, will act as a pillar on Research & Innovation that will take some activities under H2020 before being worked out in other programmes.

Report from the OverArching Working Group (OA WG)

Arjen van Rijn introduced the topic by reminding the delegates of the OA WG Report entitled "e-IRG's Reflections on the ESFRI Roadmap 2018 process", a two-page document that was already sent to the ESFRI Executive Board. He reminded of the e-IRG members in the ESFRI Strategic Working Groups (SWGs) who did the work. Overall, the OA WG members were happy with the process and even happier than with the one before on the Roadmap 2016. Only the last question on EOSC would need some rework for the next round. Arjen van Rijn stated that the report is only critical on the new ESFRI DIGIT, as there is the risk of doubling the work of e-IRG and this needs to be clarified in the next round, even though e-IRG OA Working Group members involved in the ESFRI SWGs are by definition involved. He clarified that this does not have anything to

do with the Chair of the ESFRI DIGIT WG. Overall, he remarked that it is very relevant to interact with the ESFRI domains, as this interaction brings value to both generic e-Infrastructures and domain-specific ones, getting to know better each other.

Questions and comments:

In the draft Council Conclusions, both ESFRI and e-IRG are mentioned under item 8. Do you expect that they are aligned?

Arjen van Rijn stated that ESFRI and e-IRG are operating very well in this way, with the OA WG and e-IRG Chair. e-IRG has a broad remit there, not only to support the Research Infrastructures. The collaboration should not only be maintained but increased further, if possible. One concrete idea already suggested in the answer to the EC EOSC Roadmap Staff Working Document is to study use cases for ESFRI Landmarks. e-IRG's relation to ESFRI is getting better and better.

Rosette Vandenbroucke also reported on the status of the Landscape Analysis and the e-IRG contribution, acting as the editor from the e-IRG side. The e-IRG input has been included a few months ago already. The DIGIT WG might also add some points on the new ESFRI e-Infrastructure proposals. No comments were received on the e-IRG text. Ana Proykova sent a message praising the e-IRG work. However, it is possible that in the next round the DIGIT WG might undertake the e-Infrastructure landscape analysis and do it on their own.

The Chair thanked the OA WG for its work and the good collaboration with ESFRI.

Report from the Working Group on the Long-Term Preservation (LTP) of Research Data

Rosette Vandenbroucke thanked all delegates who commented on the document. The comments were quite wide-spread: from not good enough to fine. Several had more detailed comments. She reminded of the contributors in the new round of the document, namely Rosette Vandenbroucke, Marcelo Maggi from the University of Bari, and Jamie Shiers from CERN, supported by Marcin Lawenda and Michalis Maragakis from the Support Project. She stated that the document must be supported by all delegates, including concrete guidelines and not stay at general level. It should be more of a "niche" - i.e. specialised - document about long term preservation of research data, considering the FAIR principles.

The Long Term Preservation of Research Data document will be further enhanced, integrating the comments received with further work on the recommendations, and related text to support them. The paper should be concise and more targeted. References to other documents should be

put in areas that are not of focus or thinly covered, so that the reader can be redirected for more info.

Working Group on National Nodes

Group co-Chair Arjen van Rijn introduced the topic. He presented the picture of the e-Infrastructure Commons that is now embraced by EOSC and stated that the "C" in EOSC becomes more and more a "Commons". He reminded that in the e-IRG Roadmap 2016 two recommendations are directed at national governments and funding agencies, to embrace e-Infrastructure coordination at the national level and build strong national e-Infrastructure building blocks, enabling coherent and efficient participation in European efforts; and together analyse and evaluate their national e-Infrastructure funding and governance mechanisms, identify best practices, and provide input to the development of the

country with regard to consulting the ministry, the e-Infrastructure and e-Science committees, etc.

Non-Conflict of Interest (Col)

Arjen van Rijn presented the current status of the document. He explained the current state of the package that is available as an on-line document. The main parts of the package are the following:

- Non-Conflict of Interest policy and process
- Rules for participating in e-IRG
- Non-Col policy, addressing rules of publications and participation in e-IRG bodies
- Rules of representation of e-IRG (talk on behalf of e-IRG)
- e-IRG delegates e-infrastructure

used. For the policy documents Col information, including authors and statistics will be also included. A template is already included in the package that is a part of the policy document Terms of Reference (ToR).

It was remarked to double-check the security of the website, where all the information will be stored and also that the Delegates Meeting agendas are not publicly accessible.

e-IRG Liaisons

The Chair introduced the topic saying that it should be discussed with which groups and entities e-IRG should strive for an active liaison with defined contact persons and with which groups and entities observers should be exchanged.

Besides ESFRI, where there is an established collaboration and observers are already exchanged in the meetings as well as in the SWGs, the following other bodies were



European e-Infrastructure landscape.

He brought to mind the two-stage process agreed in the previous meeting, namely a template with headlines for countries including networking, computing, data, other e-Infrastructure services, as well as main domain-specific nodes, and whether they use the horizontal e-Infrastructure services or their own. The template should also cover aspects around funding streams; organisation, management and governance; legal restrictions; and access policies. The template consists of one page per country with a deadline in May. The second stage includes the analysis of the filled-in templates and a more structured questionnaire, consisting of national-specific questions but also questions about the links at European Union level.

Regarding responsibility, each e-IRG country delegate is responsible for the answer. It is up to her/him to do what is needed inside the

involvement template

- e-IRG Non-Conflict of Interest Policy
- Authorising a Conflict of Interest
- Declaration of No Conflict of Interests
- Terms of Reference template including Col statement

It was reminded that the principles of the Non-Col policy were already approved in the previous meeting. The only new contribution in the package by the Col group is the "Rules of representation". A related comment and revision that was not included in the document was updated during the meeting.

All delegates will be asked to fill in the "e-Infrastructure involvement template". The collected information will be kept inside the delegates space. It will be explained why this is done and how the info will be stored and

initially included in the agenda, namely GEANT, LIBER, PLAN-E and the Research Data Alliance (RDA). It was agreed to have a general discussion about the concept and then to go into different initiatives.

Activities during the Austrian EU Presidency

Paolo Budroni, Austria presented the updated plans on events that will take place. The EOSC Launch event will take place on 23 November 2018. It is organised by the EC together with the Austrian cabinet, probably in the library of the University of Vienna. The EC will be sending the invitations. Ministers of Germany, France, and the Netherlands expressed their interest. ❖

Full public summary of the meeting is available at:

<http://e-irg.eu/e-irg-meeting-may-2018>

European Open Science Cloud's federating core will give researchers access to e-Infrastructures and services

After the e-IRG Open Workshop held in Sofia, Bulgaria, we had the opportunity to speak with Augusto Burgueño-Arjona, Head of the e-Infrastructures and Science Cloud Unit in DG Connect at the European Commission, for an update of the European Open Science Cloud's status.

Augusto Burgueño-Arjona gave a presentation during the Vi-SEEM Workshop on e-Infrastructures at the Grand Hotel Sofia, co-located with the e-IRG Workshop. Augusto Burgueño-Arjona is responsible for the funding of the e-Infrastructure projects that are supposed to be the foundation of the European Open Science Cloud.

The European Science Cloud is about connecting researchers and about data infrastructures. It is not about the data infrastructure itself, it is not about EuroHPC, for instance.

No, EuroHPC is a different project. This is not about the deployment of new infrastructure. This is about making the existing infrastructure accessible to all European researchers.

How many researchers are we talking about?

In the slide of my presentation, I put 1,7 million researchers in Europe.

The European Open Science Cloud vision is already there since a number of years. What is the status today?

We have been using the concept of enabling access to existing infrastructures already for many years. In the past, we used to have initiatives focused on different technologies. For example, we used to have distributed computing and we had the EGI as an initiative

or data infrastructures like EUDAT, or GÉANT for networking. That has always been our approach to enable access to existing infrastructures. The European Open Science Cloud concept was put forward two years ago and we have just recently published a staff working document where we explain what we have done so far and what will come in the next few years.

The status of development is that we have one major project called EOSChub, which is a 30 million euro project with 100 participants.



Augusto Burgueño Arjona / EC

They are supposed to develop what we refer to in the staff working document as the federating core. They are going to develop all the enabling services that will allow researchers to access existing infrastructures and services. We also have OpenAIRE-Advance which is a 10 million euro project to cover the services that are required for the publishing part of research and the publishing of data as well. Another initiative is called FREYA. This is a project for persistent identifiers. There are also complementary activities like the Research Data Alliance that is supposed to facilitate the development of standards and recommendations for data sharing.

There are other initiatives which are also important like the EOSCpilot. This is a project which is looking at different governance models and also contributing to the overall concept. Another project is eInfraCentral

which is developing a catalogue of services that will be reused in the European Open Science Cloud.

So, there are already a number of projects but if you look at the total amount of money, it is still quite modest, you could say. What are the total plans for the funding until the end of Horizon2020?

It's difficult to pinpoint exactly how much money will be provided for this initiative because this initiative builds on existing e-Infrastructures. The projects that I mentioned

are about 100 million euro if you include GÉANT. There is more budget coming in 2020 in the order of 80 million euro. The idea is to complement the funding that is already invested at Member State level for the actual deployment and maintenance of research infrastructures. Overall, when you look at the

federating core and the federated entities which will be provided by Member States, regions and institutions, we are talking about hundreds of millions of euros, of course.

It will be a federated infrastructure and you mentioned entities. What will be the entities that will be federated?

I can think of all the research infrastructures that are already existing like for example ELIXIR. We expect that an infrastructure like ELIXIR will make available resources and services through the European Open Science Cloud, as well as the resources that are now being made available through infrastructures like EUDAT, EGI, and GÉANT. We also expect that the National Research and Education Networks (NRENS) will provide services to the European Open Science Cloud. So, there are indeed a large number of federated entities.

So it is not about federating countries like, for instance, GÉANT does, that federates NRENS. It is about federating at a more complex level?

It is about enabling access to services. The concept of federation reflects how the different allocations of responsibilities will be allocated. When you look at what we call the federating core, necessarily there must be somebody who provides identity and access management services, who takes care of security, who has accounting services to know how much resources have been used and so forth. We consider that the key element of the European Open Science Cloud is the provision of services and enabling access to those services.

*What will be the plans for the coming years?
What will be the focal point?*

Still in the timeframe of Horizon2020, there will be another Call in 2020. The budget that we have, is about 80 million euro. At least, this was announced in the Draft Work Programme that we presented in the last Programme Committee. The idea is to continue supporting the development of the federation core and also to support service providers to cover the cost of access, the cost that they incur when new users access their services. In the Ninth Framework Programme, we will see. ❖

ESFRI's 5th Strategy Report & Roadmap 2018

in Vienna with the presentation of six new Research Infrastructures

On September 11, 2018, the European Strategy Forum on Research Infrastructures (ESFRI) presented the 2018 ESFRI Roadmap on Large Scale Research Infrastructures during a half-day conference in Vienna, Austria. This conference was part of a series of events marking the Austrian Presidency of the Council of the European Union. It was organised by the Federal Ministry of Education, Science and Research, in

close cooperation with ESFRI and the European Commission.

ESFRI Chair; Giorgio Rossi presented the already fifth Strategy Report at the invitation-only meeting in the Aula der Wissenschaften. ESFRI periodically selects proposals of Research Infrastructures that become ESFRI Projects and have up to ten years to reach implementation, and identifies successfully implemented Research Infrastructures in the class of ESFRI Landmarks. The ESFRI Roadmap currently includes 18 ESFRI projects and 37 ESFRI Landmarks grouped in six thematic domains: Energy; Environment; Health and Food; Physical Science and Engineering; Social and Cultural Innovation; and Digit.

The ESFRI 2018 Roadmap identified six new Research Infrastructures (RI) of pan-European priority: eLTER; DiSSCo; EHRI; METROFOOD; IBISBA; and IFMIP-DONES. These new research infrastructures will be designed to meet the long term needs of European research communities. During the event, invited participants discussed about the prospects and challenges posed by the new Roadmap.

The coordinator of the eLTER Research Infrastructure, Michael Mirtl, introduced the basic outline of eLTER. eLTER Research Infrastructure's inclusion on the ESFRI Roadmap is a significant achievement for the European ecosystem, critical zone and socio-ecological research communities. It paves the way for the further development and formalization of the eLTER Research Infrastructure, the Integrated European Long-Term Ecosystem, Critical Zone and Socio-Ecological Research Infrastructure.

When fully operational, eLTER RI will be a permanently funded and managed, distributed infrastructure of field sites covering European environmental zones. These field sites will be designed to support excellent science on the functioning of life supporting systems, a whole-system approach, including studies of human-environment interactions and earth's critical zone.

The Research Infrastructure will provide a range of services to end users, including access to highly instrumented and expertly staffed sites and to their long-term environmental observation data. Additional services, such as data synthesis and modelling, support for research and technological development, and training programmes, will be provided via Topic Centres. The Research Infrastructure will also provide a range of outputs supporting policy and science users, such as information on European environmental trends.

The next step in the process is to prepare a proposal for a Preparatory Phase Project (PPP), which will develop key components of

the eLTER Research Infrastructure such as the legal entity, governance structure, Head Office, Topic Centres, the standard observation programme, the physical network of LTER Sites and LTSER Platforms included in the Research Infrastructure and the services that the Research Infrastructure will offer to users. The requirements of a wide range of stakeholders will be considered in this process, for example, to take account of emerging strategies in European environmental monitoring.

The eLTER ESFRI process is led by the German Helmholtz Centre for Environmental Research (UFZ) in close cooperation with the Environment Agency Austria (EAA) and with support of the German Federal Ministry of Education and Research. 17 countries supported the eLTER Research Infrastructure ESFRI application politically and 161 research institutions in 27 countries signed the scientific Memorandum of Understanding of the eLTER Research Infrastructure.

DiSSCo has also formally entered the European roadmap. The project was presented by its coordinator Dimitris Koureas. DiSSCo is a new pan-European Research Infrastructure initiative of 21 European countries with a vision to position European natural science collections at the centre of data-driven scientific excellence and innovation in environmental research, climate change, food security, one health and the bioeconomy. Its mission is to mobilise, unify and deliver bio- and geo-diversity information at the scale, form and precision required by scientific communities; transforming a fragmented landscape into a coherent and responsive Research Infrastructure.

European natural science collections are now, for the first time at European level, recognised as a priority Research Infrastructure for the European Research Area. This constitutes a step-change in the way natural science collections will continue develop and operate. As DiSSCo enters the European roadmap for Research Infrastructures, the project embarks on a Preparatory Phase to improve the DiSSCo technical, scientific, technical and financial readiness levels, and deliver the final blueprints for the construction phase, starting in 2020.

The application for placing the European Holocaust Research Infrastructure (EHRI) on the roadmap produced by ESFRI was submitted by the Dutch Minister of Education, Culture and Science, Ingrid van Engelshoven. With the announcement in Vienna, EHRI is placed on the ESFRI Roadmap, which provides a plan for the further scientific integration of Europe in many research areas. In this way the EHRI project will move to a new phase in which its

results will be secured in a long-term European organisation. Supported by individual Member States, this organisation will continue to support research into the Holocaust. In this year's update of the ESFRI roadmap, EHRI is the only initiative in the humanities that has been given this opportunity. According to Dr. Conny Kristel, Senior Researcher at NIOD and Director of EHRI since its inception in 2010, who sadly passed away on October 6, 2018, the great importance of the subject and the necessity of the renewal of historical research beautifully reinforce each other.

Research into the Holocaust depends on a diversity of archives and sources that are dispersed across the globe. EHRI facilitates online and physical access to these sources as well as methods and techniques that enable researchers and archivists to collaborate and share expertise. EHRI is currently funded by Horizon 2020, the largest European research and innovation programme ever. By becoming a long-term European Research Infrastructure, EHRI's dependence on European subsidies will be replaced by a business model in which member states jointly finance the organisation. EHRI is coordinated by NIOD, Institute for War, Holocaust and Genocide Studies in Amsterdam.

When EHRI started the creation of an inventory of archival material in 2010, it soon became clear that its scientific efforts also fulfill an important political wish: namely that the Holocaust is a comprehensive lesson from history that must not be forgotten, especially at a time when eyewitnesses are dwindling in numbers. If 'never again' is the call, 'modernize research' provides the answer. Whenever people look for their common history, however horrific it may be, insight and mutual understanding can develop. In recent years, new opportunities for Holocaust research have been created in countries behind the former Iron Curtain, enabling us to arrive at a better view of our shared past.

Minister van Engelshoven characterized the importance of EHRI as follows: "The unique thing about EHRI is that it links all European Holocaust archives to each other. As a result, new research can be done and more knowledge can be gathered about this dark period in our shared past. It is therefore very important that Europe decides to give EHRI a permanent status as a Research Infrastructure, so that historical research can continue and be further improved."

A common good to meet European Union research goals and improve food production and consumption, Giovanna Zappa and Claudia Zoani from the Italian National

Agency for new Technologies, Energy and Sustainable Economic Development (ENEA) and the Joint Research Unit METROFOOD-IT presented the Infrastructure for promoting metrology in food and nutrition (METROFOOD-RI), the first ESFRI Research Infrastructure coordinated by Italy for the domain "Health & Food", designed to support research and develop concrete solutions and synergies in the area of food quality, safety and authenticity.

METROFOOD-RI is a new distributed Research Infrastructure aimed to promote scientific excellence in the field of food quality and safety. It provides high-quality metrology services in food and nutrition, comprising an important cross-section of highly interdisciplinary and interconnected fields throughout the food value chain, including



agri-food, sustainable development, food safety, quality, traceability and authenticity, environmental safety, and human health. METROFOOD-RI consists of a physical infrastructure (P-RI) and an electronic infrastructure (e-RI) to coordinate and integrate existing networks of plants, laboratories, experimental fields and farms for crop production and animal breeding, small-scale plants for food processing and storage, kitchen-labs for food preparation. The e-RI will make available an access platform to share and integrate knowledge and data on metrological tools for food analysis, focusing on food composition, nutritional contents, levels of contaminants and markers.

The METROFOOD-RI consortium includes 48 partners from 18 European countries, with

more than 2200 researchers involved. In Italy, 12 research institutions set up a devoted JRU and ENEA, which acts as coordinator, engaged several research units belonging to the different disciplinary fields. Quality, safety, traceability and food authenticity, sustainability and nutrition are the keywords around which researches and studies focus, with the aim to promote measurement reliability and the adoption of the FAIR (Findable, Accessible, Interoperable, Re-usable) approach, favouring harmonisation and knowledge integration and sharing.

The ESFRI committee also included IBISBA in the ESFRI 2018 Roadmap as Research Infrastructure project. Michael O'Donohue, the project coordinator, had the honour to officialise the presence of IBISBA on the ESFRI Roadmap. On December 1, 2017, the IBISBA 1.0 project was launched. This project entitled Industrial Biotechnology Innovation and Synthetic Biology Accelerator 1.0, aims to create a coordinated network of research infrastructure facilities that support industrial biotechnology. The work of the network is to open access to first-rate research facilities and lay the foundations for a permanent Research Infrastructure that will address some of the main challenges that are currently hampering the development of industrial biotechnology, a key enabling technology (KET) of Europe's bioeconomy.

The proposal to implement DONES in the ESFRI Roadmap 2018 is the missing link to develop nuclear fusion as an electricity producing source while opening a new Research Infrastructure to the wide scientific community, stated University of Granada-based project representative Carlos Alejandre. The mission is to generate materials irradiation test data for design, licensing, construction and safe operation of nuclear fusion.

IFMIF-DONES, which stands for International Fusion Materials Irradiation Facility - DEMO-oriented Neutron Source - is currently being developed in the frame of the EUROfusion Early Neutron Source work package (WPENS), based on the results achieved in the ongoing IFMIF/EVEDA (Engineering Validation and Engineering Design Activities) project by Japan and Europe in the frame of the Broader Approach (BA) agreement. The neutron source's limited product of "irradiation volume x neutron fluence" on one hand and the large potential test matrix - defined by number of material grades, test types, irradiation and test temperature levels and damage dose levels - on the other hand, require a careful selection of test conditions by collaborative effort of the communities of DEMO designers, fusion materials science and the irradiation facility designers. ❖

Interview Simon McIntosh- Smith

World's first ARM-based Isambard supercomputer is ready for science

At the Workshops Day at ISC'18 in Frankfurt, Germany, we talked with e-IRG delegate Simon McIntosh-Smith, a professor of high-performance computing at the University of Bristol about the first ARM-based supercomputer.

With some of his colleagues, Simon McIntosh-Smith designed and built Isambard, the world's first production-ready ARM-based supercomputer. The team worked together with Cray. The machine is actually a Cray XC50. It is a real production machine. It is not going to be just a testbed. One is going to run Isambard as a real science cluster. People in the UK will apply for time on this machine and go through a science review process. If the development team likes what they want to do, they will be allowed time on the Isambard machine. Hopefully, they will log in and use it like they would any other big high-performance machine.

The difference though is under the hood. Isambard will actually be a first generation of high-performance computing optimized ARM-based CPUs. The machine was due to arrive at the end of July 2018 and Simon McIntosh-Smith hoped to open up for science this Summer.

What is the size of the machine?

It is a full cabinet of the Cray XC50 hardware. It has 168 dual socket nodes. The Cavium ThunderX2 ARM CPUs that are used are 32-core each at 64 cores per node. So, this is something like just over 10,700 cores for the whole machine. Isambard uses the Cray Aries interconnect between all those 168 nodes.

If you position Isambard in the European supercomputing scale, it is tier 0?

In the UK, one would call it tier 2. The PRACE machines are tier 0. The national machines like Archer in the UK are the tier 1 machines. Isambard is a tier 2 class machine, it is a regional machine. It is operated by four universities together, called the GW4 Alliance - the Great Western four university of Bath, Bristol, Cardiff and Exeter, all in the southwest of the UK.

This is not a machine you can buy from a catalogue?

Actually, you pretty much can. That is the other exciting thing. It is just going to be an XC50 standard product. You can go to Cray and buy one of these off their list. Cray is not the only company and not the only system vendor doing ARM-based machines now. HPE and Bull/Atos are doing them too. Fujitsu is having one coming soon. Most of the main vendors now have this on their list. It is



becoming more of a commodity standard thing that is just starting to happen now.

This is a good sign that the technology is catching on. At the ARM workshop at ISC'18, you presented some benchmarks?

Isambard is focused on real science and it is part of the science ecosystem in the UK. So, the team took ten of the most heavily used science codes on the tier 1 machine, Archer. The team has been porting and optimizing those on Isambard. The developers had some early access nodes which are basically a sort of white boxes with the right Cavium ThunderX2 CPUs in them and a pre-release of the Cray software stack. They got all of the Cray compilers, libraries and profilers, as well as all of the open source tools. They had things like GNU compilers and Clang-based compilers as well. They went through all the important science codes. They are all already compiled and running well for the prototype early access. Those are lots of chemistry codes, things like VASP and CASTEP, as well as

molecular dynamics codes like GROMACS and NAMD. They have got CFD codes in the top 10 like OpenFOAM and lots of atmosphere and weather codes. In the UK, the unified model from the Met Office is one of the key codes. They also got NEMO, which is an ocean code. All those are real science codes which are already running on this prototype machine.

The team has been presenting itself across the board. Performance was very competitive with the latest x86 and Skylake from Intel. It was really a great fact. It all pretty much worked out of the box. It was good to learn that the software ecosystems are mature. The performance was good. This is really important because that means it is a genuine option for people buying new machines. It is really exciting to see what this means for the team's options when they will be buying their new machines. They will have a bit more variety and more competition. This is good for the team and the users.

As for the machine itself, the processors are ARM?

Yes. They are Cavium ThunderX2 processors which are based on ARM.

What is the network connect of this system?

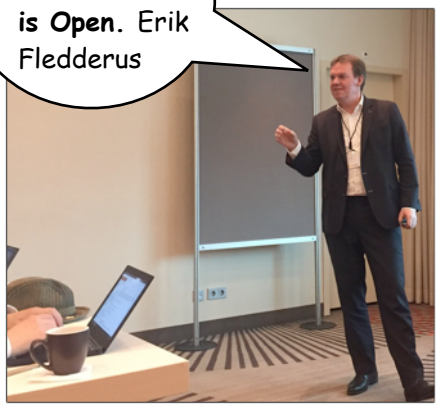
Since Isambard is a Cray XC50, it has the Aries interconnect. That is what they have been using with Intel-based systems for a long time. Also, for the team, interestingly, that is

the same interconnect as within the Archer tier 1 machine. When some comparisons are being performed, there are already real results on Archer with Aries and the Cray tool chain. Now, the team can run some tests on Isambard with Aries and the Cray tool chain. The only real difference is the processor. So, the team has gone from an x86 to an ARM-based Cavium ThunderX2 processor and almost everything else is the same. The team is really seeing the difference, going just from one to the other, they can use some sort of apples-to-apples comparison which is very useful for the team.

Surely, you are looking forward to the scientific results of the machine?

It should be really good, hopefully, if the acceptance goes well. Isambard is bound to be open for science this Summer. Hopefully, people will very quickly get used to the fact that ARM works just like anything else that they use and just start using it for real science. ❖

The future of Science is Open. Erik Fledderus



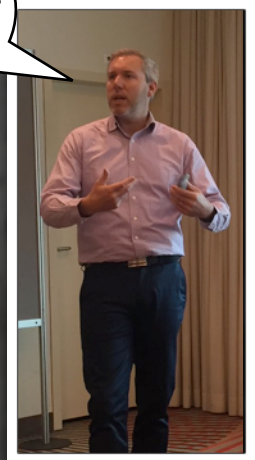
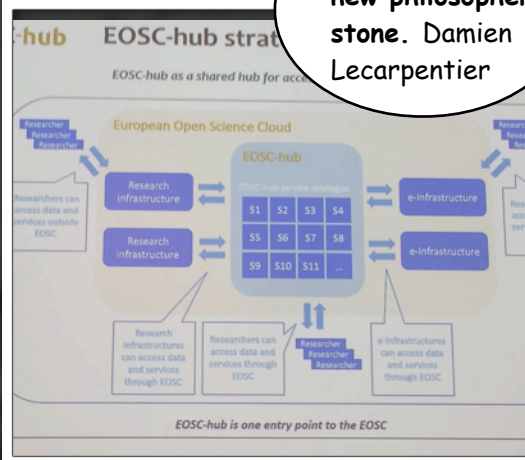
e-Infrastructure Commons beginning of EOSC Arjen van Rijn



e-IRG info element in EOSC governance discussions Arjen van Rijn



EOSC is not the new philosopher's stone. Damien Lecarpentier



Here it is, the PLAN-E workshop venue



Creating seal of trust for FAIRness is complex Peter Doorn



FAIR enough checklist COMING SOON!

Is your data FAIR enough? **Sneak Peek**

Checklist to evaluate FAIRness of data(sets)

by Eliane Fankhauser @ DANS

- Quiz/checklist format
- Easy to understand and fill out by any researcher
- Focus on awareness raising rather than on FAIR orthodoxy

SPOILER ALERT!



FAIR Enough checklist that is not that rigid & more usable.

Why should researchers care?

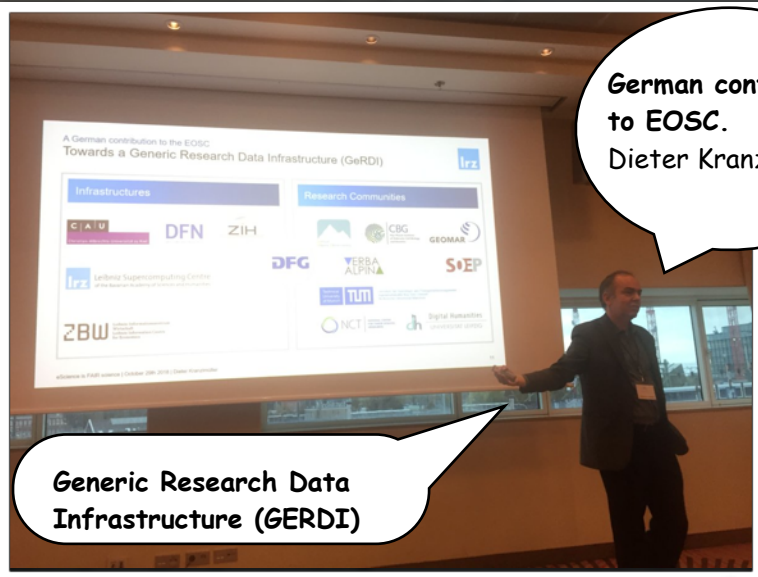
Opportunities/responsibilities for research:

- Increased responsibility:
 - Transparency and reproducibility
 - Reduce risk of fraud
- Increased efficiency:
 - (Rapid) sharing of resources
 - Reduce risk of duplication

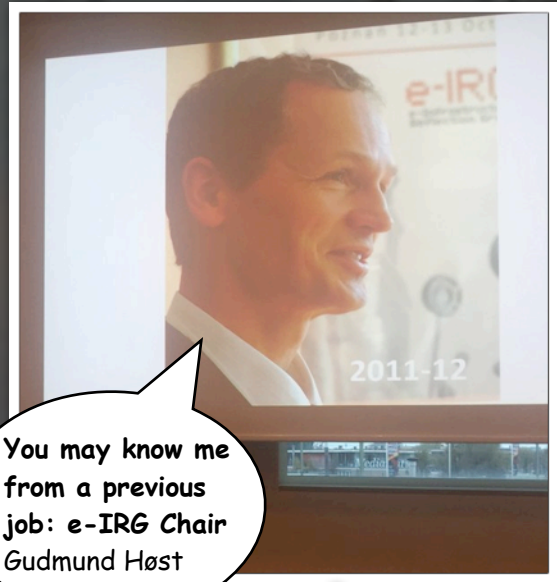
Is "openness" needed for this?

Political statements: Research results from public funding should be openly accessible and possible to use by everybody.

Why care about Open Science enabled by e-Infrastructures? Sverker Holmgren



German contribution to EOSC. Dieter Kranzlmüller

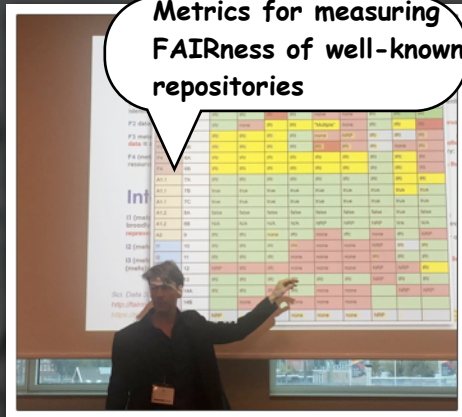


You may know me from a previous job: e-IRG Chair Gudmund Høst

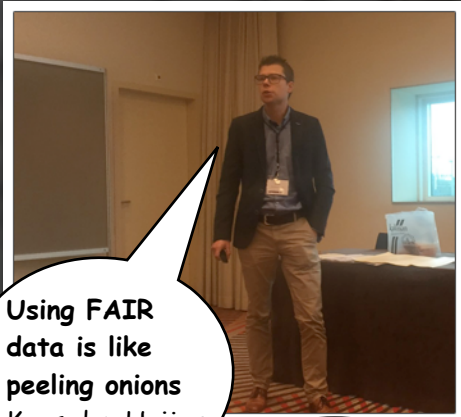
Generic Research Data Infrastructure (GERDI)



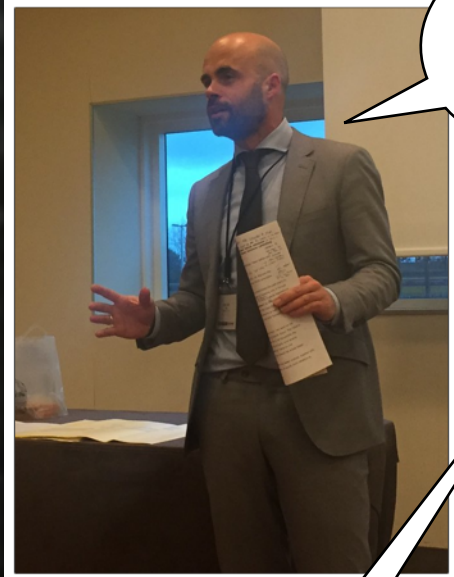
GO FAIR explained Erik Schulthes



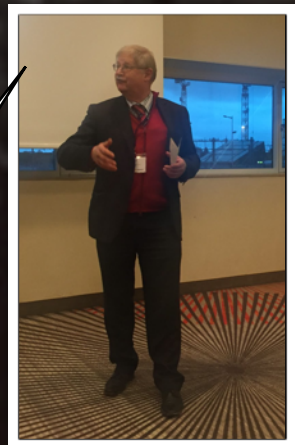
Metrics for measuring FAIRness of well-known repositories



Using FAIR data is like peeling onions Kees den Heijer



Dutch government supports FAIR Open Science Aart Jan Bette



That's it for this PLAN-E meeting Patrick Aerts



Let's go to the eScience Conference Wilco Hazeleger

European Open Science Cloud to be officially launched in Vienna

The newly appointed representatives of both Governance Board and Executive Board will present the Governance Structure. The first version of the EOSC Portal, the entry point of the Cloud, will be demonstrated and its functionalities explained. ❖

<https://eosc-launch.univie.ac.at/home/>

Interview Ognjen Prnjat

Vi-SEEM virtual research environment emerges as a mini-EOSC for Southeast European and Eastern Mediterranean scientists

The e-IRG Open Workshop in Sofia, Bulgaria was followed by the International Conference on e-Infrastructures for excellent science in Southeast Europe and Eastern Mediterranean, organized by the Vi-SEEM project. We were able to talk with Ognjen Prnjat, Vi-SEEM project coordinator and responsible for the conference. Ognjen Prnjat works for the Greek Research and Technology Network. This is the Greek National Research and Education Network (NREN) as well as the supercomputing and Cloud provider for Greece and beyond.

Ognjen Prnjat has been leading regional infrastructure projects for the last 15 years, focusing on the region of Southeast Europe, also including the Western Balkan and Eastern



Mediterranean countries. This has been a big partnership of about 20 countries in the region in the fields of networking, grid computing, Cloud computing, high performance computing, and now a unified virtual research environment for the scientists of the region.

What is the Vi-SEEM project all about?

Vi-SEEM is a three-year European Commission project. It is a virtual research environment for the scientific communities from the Southeast and Eastern Mediterranean region. It is unifying different types of e-Infrastructure resources for the benefit of scientific communities and digital cultural heritage, life sciences, and climate prediction and modelling.

Who are the partners?

There are 15 partners from the region, including the EU Member States Romania, Bulgaria, Greece, Cyprus, Hungary, the Western Balkan countries, and some Southern Caucasian countries like Georgia and Armenia, and the countries from the Eastern Mediterranean region such as Egypt, Jordan, and Israel. It is a wide geographical area, a wide geographical spread of countries that are covering a lot of scientific resources. There are also a lot of communities who are active in that region.

What are the main results of the project thus far?

The project has a mini European Open Science Cloud (EOSC) running which unifies different types of e-Infrastructure resources in an open Cloud environment. This gives access to different scientific communities. There are hundreds of different applications running. There are hundreds of publications resulting from these applications and scientific collaborations. There are quite a few tangible results, not only in terms of infrastructure, but also in terms of actual research.

How are you working together with the other stakeholders or groups in Europe?

Vi-SEEM is very much involved in the policy dialogue as well. There is a close collaboration with other policy makers from the region. In this particular conference that is organized in Sofia, there is also co-location with e-IRG and with the Western Balkan Summit of the Western Balkan leaders who are meeting with the European Union leaders later on. So, there is a very close policy level collaboration, as well as a full blown technical level collaboration, in terms of working together in pan-European e-Infrastructures. Both the policy level and technical level collaborations are in place.

At this moment of the interview, the Vi-SEEM conference is not yet halfway but can you already tell a little bit about the contents?

On the first day, there is the policy afternoon. The policy makers of the region are joining the project team and the scientists from the region. e-IRG delegates from Western Europe are also present. The second day is a full day of scientific presentations about the tangible results of the use of the virtual research environment. This is a dedicated day for the scientists.

What is your favourite application?

I feel very strong about climate change, so climate modelling definitely is very important. But digital cultural heritage is very interesting too. In the Eastern Mediterranean region and Greece, we have a long history and there are some very nice applications in visualisation of monuments and cultural heritage collections. There are very diverse application fields.

Does the project run until the end of this year? And what will come after that?

Vi-SEEM runs until the end of September 2018. We have developed a sustainability model. There will be services provided beyond the term of the project. We are also looking forward to joining other pan-European initiatives, like for instance EuroHPC. There will be other things coming up. ❖



e-IRG Workshop Sofia

Big Data for HPC, European Open Science Cloud progress, and European investment for e-Infrastructures in the Southeast and Mediterranean region addressed at the e-IRG Workshop, May 14-15, 2018 in Sofia, Bulgaria

Nearly 100 participants attended the e-IRG workshop organised in the framework of the Bulgarian EU Presidency on 14-15 May, 2018 in Sofia, Bulgaria. The workshop had three main topics: challenges resulting from Big

Data for HPC; progress on the development of the European Open Science Cloud; and use of European Structural Investment Funds for e-Infrastructures in the Southeast European and Mediterranean region.

e-IRG chair Prof. Gabriele von Voigt, on behalf of e-IRG, and local host Prof. Aneta Karaivanova, Bulgarian Academy of Sciences (BAS), opened the workshop and welcomed the audience.

Keynote

In his keynote speech Ivan Dimov, deputy minister of Bulgarian Ministry of Education and Science highlighted the importance of education, training and the developments of digital skills in the ongoing digital transformation to secure the competitiveness not only of the region but of the European Union as a whole. Furthermore, he underlined the importance of regional connectivity to improve the regional business climate, research and innovation ecosystem. Addressing the first topic of the workshop Ivan Dimov said that focus should be put on the development of the next generation of key HPC technologies and systems towards exascale. In this context he emphasized the efficiency of the implementation and the importance of suitable algorithms for supercomputing.

The session was completed by two presentations from the European flagship projects, Quantum Computing and Human Brain Project.

Dr. László Bacsárdi from the Budapest University of Technology and Economics

presented an overview on the possibilities of quantum computing and displayed a timeline towards the realisation of quantum computing, which comprises innovation in various areas like communication, computation, simulation and sensing on top of the basic science. He also mentioned that although true large scale quantum computers will not be constructed soon enough, they can and will be used as accelerators in several cases in the near future.

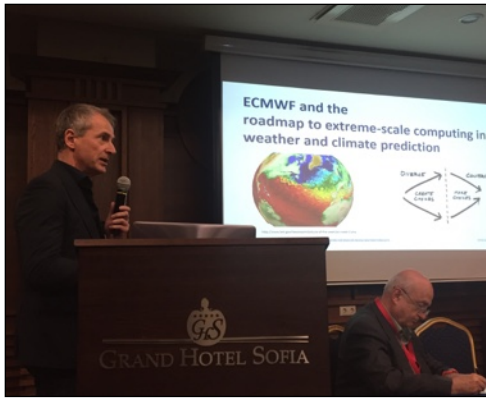
Colin McMurtrie from the Swiss National Supercomputing Centre (CSCS) presented the FENIX e-Infrastructure, a framework that provides elastic HPC resources for research communities. Members of the FENIX consortium are large European HPC sites, including the Barcelona Supercomputing Centre (BSC), CEA, CINECA, CSCS and the Jülich Supercomputing Centre (JSC), that are aiming to establish HPC and data infrastructure services for multiple research communities. The Interactive Computing e-Infrastructure (ICEI) is the first instantiation of FENIX, which provides a federated data and computing infrastructure for the Human Brain Project. He also explained the idea behind Fenix credits that will be used by

communities in order to acquire many different types of resources such as CPU, memory, network IPs, etc.

Second session

The second session of the workshop comprised two presentations from the foreseen private partners of the EuroHPC Joint Undertaking, ETP4HPC and Big Data Value Association (BDVA), and a presentation from the user perspective from the European Centre for Medium-Range Weather Forecasts (ECMWF).

Marcin Ostasz from the Barcelona Supercomputing Center represented the European Technology Platform for High-Performance Computing (ETP4HPC), a contractual Public-Private-Partnership. One key activity of ETP4HPC is the production of the Strategic Research Agenda (SRA), a multi-annual HPC technology roadmap that provides input to the Commission's Work Programmes for the development in the area of HPC. The EuroHPC Joint Undertaking (JU) was initiated to encounter shortcomings in the HPC area in the development as well as in the supply of HPC resources. A two-phase



approach is envisaged. The first phase until 2020 will deal with pre-exascale and petascale machines, and applications and technologies for the exascale. This is done by projects that are using the concept of Extreme-Scale Demonstrators - prototypes of future European exascale machines. In the second phase, which will last until 2028, exascale and post-exascale machines, the first hybrid HPC/Quantum infrastructure, and applications and technologies for post-exascale will be addressed. The JU will be a three-party partnership between the European Commission, Participating States and Private Members.

Yannick Legré from the EGI Foundation, as representative of the Big Data Value Association (BDVA) presented BDVA as the second foreseen private member in the EuroHPC JU. BDVA is a contractual Public-Private-Partnership that contributes to the Digital Single Market Strategy Implementation and provides data innovation recommendations to the European Commission. BDVA is present in 28 countries and currently has 194 members. The published BDVA Roadmap provides recommendations towards cross-sectorial and cross-lingual data integration to support the data-driven digital transformation. For the next Framework Programme, BDVA believes that the impact of the cPPP approach can be further strengthened by the stronger involvement of Member States and regions, and a stronger collaboration between the private side and the EC to maximise success.

Peter Bauer from the European Centre for Medium Weather Forecast (ECMWF)



presented the European Centre for Medium Weather Forecast, which is an independent intergovernmental organisation managed by 34 Member States and cooperating states. He described the basic features of the ECMWF weather forecast model that is the leading model in the world in terms of verification statistics, and how it has contributed in predicting weather extremes that have also a social impact. He stressed the importance of high spatial resolution and the role of model ensembles, in order to model the weather complexity. He highlighted the role of HPC and data management for achieving true extreme-scale performance and the Extreme Earth Flagship proposal. The Flagship proposal among others can contribute to better weather and climate modeling, help foresee localised extreme weather events such as tropical cyclones and even relate weather prediction to health related outbreaks of epidemics.

In the panel discussion with representatives from the flagships, the EuroHPC private members and the ECMWF, the issue regarding the risks of the recently launched big "digital" initiatives was discussed. Yannick Legré saw a risk for overlap and called for coordination across the Member States, while Peter Bauer saw a risk in spending too much money in one topic since the flagships are already 1 billion euro projects and difficult to manage. Colin McMurtrie saw the risk in the plethora of HPC systems and the related abstraction.

Second day

The second day of the workshop started with a session on the progress of the development of the European Open Science Cloud (EOSC).

Athanasios Karalopoulos from the European Commission, DG RTD, summarized how the EOSC has developed until now and presented the six action lines - architecture, data, services, access & interface, rules and governance - of the EOSC Implementation Roadmap of the first phase. The second phase of the implementation of the European



Open Science Cloud will start after 2020 after a thorough evaluation. The timeline for the implementation of the EOSC would start with the adoption through the Competitiveness Council at the end of May 2018, followed by the second EOSC Summit on 11 June, 2018 during which consultations on the rules for participation and the data action plan would be launched. The Member State agreement on the EOSC governance structure is planned for November, so that by the end of 2020 the operational phase of the European Open Science Cloud should start. The EOSC provides chances, but includes also risks for the implementation, such as the support from the Member States to back the bottom up approach.

Matthew Dovey from JISC provided an update from the EOSCpilot project, which is already half way through. Three types of challenges – scientific, technical and cultural - have been identified throughout the course of the project: how to deploy EOSC; what are the services needed; and how to convince people to participate. The scientific challenges are considered as true challenges, while the other two are considered more like barriers which must be overcome. The scientific challenges are addressed through the scientific demonstrators to determine the scientific needs. The technical challenges are addressed by the determination of core services and the definition of rules of engagement. Matthew Dovey reported that the outcome of the EOSCpilot project will be carried over into the implementation projects and thus sustain the work that has been done.

Yannick Legré from the EGI Foundation presented the EOSC-hub project, which is aiming to create a federated integration and management system for the European Open Science Cloud comprising 20 major digital research infrastructures, EGI, EUDAT CDI and INDIGO-DataCloud jointly offering services, software and data for data-driven research. During the presentation details of the project and the main involved science domains were presented. Furthermore, Yannick Legré presented the concept and value proposition of the HUB, a European-level contact point



for researchers to discover, access, use and re-use a broad spectrum of resources for research. Taking up the six action lines from the EOSC Implementation Roadmap it was shown how the HUB contributes to each of these lines.

The presentation "Turning FAIR Data into Reality" by Françoise Genova from the Observatoire Astronomique de Stasbourg, member of the Commission's FAIR Data Expert Group, was part of the series of presentations that the members of the expert group held to get input on the FAIR Data Action Plan. This plan, which is under development, will propose a list of concrete actions on how to make data FAIR. Françoise Genova reported that the notion of data includes metadata, code and many other aspects and that trust is important in the ecosystem of data sharing. Thus, two suggested frameworks have been merged into a single framework: CoreTrustSeal.

The expert group conducted together with OpenAIRE a survey on the mandatory H2020 Data Management Plan. From the responses to the questionnaire, recommendations like the need for clarified requirements, a revised template structure or the need for discipline-specific guidelines, have been drawn. A broader consultation on the FAIR Data Action Plan would be launched during the EOSC Summit on 11 June, 2018. The final report will be published in November during the official launch and formal communication under the Austrian EU Presidency.

The session was completed by three reports from Spain, Poland and the Czech Republic providing national views on EOSC.

Ignacio Blanquer from the Universitat Politècnica de València reported that Open Science is in development in Spain and Portugal, which are closely related countries because of the existence of Ibergrid. Ibergrid, a result of the collaboration of both countries, provides services and training for



users and is focusing on a minimum viable ecosystem towards EOSC.

Norbert Meyer from the Poznan Supercomputing Center reported that in Poland EOSC services are introduced and interconnected with existing local services. Poland is interested in participating in best practices projects like EOSCpilot and EOSC-hub and other infrastructure related projects. Both countries, Spain and Poland, see ESFRI projects as an important pillar.

Miroslav Ruda from the Czech Research and Education Network (CESNET) said the federated compute resources in the Czech Republic comprise a system of systems, which are part of the European system. Regarding the uptake of the FAIR principles Miroslav Ruda reported that the Czech research communities are deciding about the specifications and tempo of the FAIR implementation.

The last session of the e-IRG workshop under the Bulgarian EU Presidency addressed how European Structural Investment Funds (ESIF) are, and can be used to finance e-Infrastructure developments in the Southeast European and Mediterranean area. The session was chaired by Prof. Kostadin Kostadinov, an advisor to the Bulgarian Ministry of Education and Science, and included presentations from Bulgaria, Cyprus, Greece and Hungary.

Prof. Svetozar Margenov from the Institute of Information and Communication Technologies of the Bulgarian Academy of Sciences reported on the development of Bulgarian supercomputers and the synergies between advanced digitalization tools and advanced computing.

From Cyprus Prof. Constantia Alexandrou provided an overview on the Cyprus Institute activities towards a Center of Excellence in simulation and data science, a collaboration with the Jülich Supercomputing Center. Prof. Alexandrou concluded that European funding



is essential for Cyprus since national funding is not adequate.

The Greek perspective was presented by Prof. Panos Argyrakis who presented collaborative activities of the Greek NREN GRNET. It was also explained that Greece is implementing a national Roadmap for Research Infrastructures, which especially supports e-Infrastructures addressing the increasing horizontal needs for big computing and networking as well as Big Data. Panos Argyrakis closed his presentation reporting that Greece has signed as 12th country the European declaration on high-performance computing, joining the EuroHPC initiative.

Eventually, Tamas Máray from the National Information Infrastructure Development (NIIIF) addressed the utilisation of Structural Funds for e-Infrastructure development and reported about the experience in Hungary. He pointed to the compatibility dilemma using Structural Funds, which should be used to narrow the development gaps among regions and Member States, and the purpose of Research Infrastructures, which should offer unique research services and aim to solve grand societal challenges. Furthermore, Tamas Máray reported that Research Infrastructure and ESIF matters are often under different ministries. Closing the presentation Tamas Máray pointed out a number of challenges and issues that need to be addressed to be able to successfully use European Structural Investment Funds for the development of national Research infrastructures including e-Infrastructures. ❖

The full workshop minutes, slides and video recordings of the presentations can be found on the e-IRG website:

<http://e-irg.eu/e-irg-workshop-may-2018>



Interview Ivan Dimov

Exchange of experience and focus on specialisation are key to Europe regaining status in EuroHPC and e-Infrastructures

After the e-IRG and Vi-SEEM Workshops in May 2018, we were invited by Ivan Dimov, Deputy Minister of Education and Science at his cabinet in Sofia, Bulgaria for an interview. Ivan Dimov gave a keynote at both workshops about e-Infrastructures and supercomputing. He is responsible for research in Bulgaria. Before that, Ivan Dimov was a scientist, working in computer science. He has been Director of the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences. He has also been working in some other places in Europe and in America and was for some period of time in England at the University of Reading, in Germany and in France.

And now, you are very busy with the Presidency of the European Union, currently being carried out by Bulgaria. A few weeks ago, there was a conference on EuroHPC which was organized by Bulgaria in collaboration with the European Commission. This conference presented a very nice overview of the state of HPC in Europe and EuroHPC. Can you comment a little bit on that meeting?

This was a very important conference for Bulgaria but also for Europe. In the past years, Europe was not among the ten biggest users of supercomputing, so it was necessary to decide what one should do in order to be again a big player in this area. Without supercomputing, without this tool for researchers, you cannot achieve big results.

Even if you deal with theoretical physics, chemistry or biology, you often need such a big machine to perform your computations in order to discover new knowledge. Data mining means dealing with Big Data you are trying to extract knowledge from. This is a big challenge for Bulgaria where there is quite a strong community dealing with high performance computing. Unfortunately, Bulgaria has no big computational power. There is a supercomputer but we need to unite our effort with all the countries in Europe. Commissioner Mariya Gabriel is Bulgarian and responsible for this part of the Commission. It was very important to be able to see what Europe should do to

perform better in this important area.

The conference brought this a little bit further. And now, there was a workshop about e-Infrastructures which is also a strong component for high performance computing. How do you see the relationship between high performance computing and e-Infrastructures?

The research infrastructure is extremely important for scientists because without infrastructure it is not possible to deliver high quality results. The normal way is that you work in research projects. You produce innovations and the innovations are going to the industry in order to make the European citizens happy. This is a very important chain of events: from research to innovation to production.

If you compare Europe to the United States of America or China, Europe is not a big player in this area. That is why we need to unite the effort of all the European countries. This is an area where each country has a specific experience and combining this experience for researchers, we reach the best in this area. In Bulgaria, there is quite a strong community dealing with algorithms, software, and middleware. Bulgaria is more specialized in software dealing with algorithms and software which is extremely important for high performance computing.

Very often, you have computers with high peak performance but low computational efficiency. One should care about specific algorithms that should be prepared to work on parallel machines. It is not trivial to distribute your tasks or your jobs on a large number of processors and have an efficient implementation. This is a sort of niche in Bulgaria where we focus on and we may contribute to the HPC process. Of course, we are working together with colleagues from other countries. It is a matter of policy of how to organize all this business in order to be more efficient.

It was decided that the European Commission already allocated one billion

euro to buy two large supercomputers and also planned to reserve about 9 billion euro for the next Framework Programme in this very important area. At the same time, this infrastructure is also related to electronic infrastructure which is very important for the region of Southeast Europe. This was actually the focus of one of the conferences. If we want to work with countries in the Western Balkans, and this is one of the priorities of our presidency, we need to have good electronic infrastructure. We are putting a lot of effort to fill this gap. This priority is also important for the whole of Europe because you will have smaller discrepancy when you compare research infrastructures in different parts of Europe.

So, basically, it is good to specialize and small countries can also contribute important components but you need to collaborate to achieve results on a European scale.

Absolutely. Modern research areas are very broad. Usually, to compete with big players in this area, you need to specialize in a smaller area to be focused. At the same time, you need to work very closely with colleagues from other countries. Bulgaria has a good experience working with colleagues from Central and Western Europe and also with colleagues from the Balkan Peninsula and Eastern Mediterranean. We need this collaboration and this infrastructure. We can see that people are very active and you have a sort of synergy when you combine different specialists from different countries. This is helpful for the whole research infrastructure in Europe.

Currently, you are working very hard and the focus is on things that are happening during the Bulgarian Presidency. What do you expect in the next year will be important?

The European Commission is about to announce the next Ninth Framework Programme which is a very important task. We are working very hard, also with our colleagues from Austria, who will be in charge of the next Presidency. Europe is not a big



Ivan Dimov (r)

player but we know our place and our areas in which we are strong and where we can contribute. We are trying by formulating our national priorities in the field of research to identify a small number of areas where we are strong enough and areas in which we may cooperate. We are focusing on those areas. We have created a number of Centres of Excellence funded by the government in some specific areas where we know that we have good expertise. One of these areas is information and communication technologies. We are very much focused on some specific priorities where we can achieve good results. In such a way, we can say that Bulgarian scientists are good partners in some areas.

There should be a sort of harmony or synergy between national research policy and European research policy. One should identify areas in which it is good to collaborate with big and strong teams. We are also collaborating with teams from America and China. This is absolutely needed. At the same time, we can identify areas that are local areas, connected to heritage and language. We are running projects connected to the

digitalization of heritage and texts because we should care about this. At the same time, we have identified areas in which it is good to collaborate on the European level. It is important to have a focused policy to target areas of research where we can expect results.

How we measure this is another important issue. There are metrics which are more or less accepted: using data from databases, in science and other sources to identify areas in which you have papers published in the top one or 10 percent. In such a way, you identify areas in which you can collaborate on a European level and also identify areas that are important for this country. They are local but they also should be stimulated.

In the Netherlands, where I come from, water is an important issue because most of the country is below sea level. Do you have scientific issues which are specific for Bulgaria?

We are focused on genome studies which are connected to food and agriculture because Bulgaria has quite a strong community in this area, also because of some hazards. In the

previous year we had a strong snowfall which was a challenge during the month of January. We have very precise models to predict different climate scenarios and weather forecasts, very precise models to be able to predict which part of the highway from Sofia to Burgas should be closed. For this purpose, the supercomputer in Sofia was used, again with very good results. It was convincing politicians that you should put money into this because then you are able to predict with an accuracy of about 300 meters and this was achieved for the first time using such a fine mesh. The mathematical technology is known but the problem was computational power. We had this power at that time. The Bulgarian society understood that ICT is something important and it is good to reserve several millions of euro to buy a good supercomputer to solve this problem instead of becoming stranded in the snow. This was very convincing and after this, we joined the EuroHPC initiative. ❖

The complete interview can be found at <http://primeurmagazine.com/weekly/AE-PR-07-18-31.html>

Around Europe

Country news

Country descriptions can be found in the e-IRG Knowledge Base. The country description in the e-IRG Knowledge Base has been actualized and updated with new information on the national infrastructure, policy and documents.

<http://knowledgebase.e-irg.eu/countries>

Austria

Austria has decided to join PRACE via Aconet for 3 years.

Because of the EU presidency Austria is gathering information about Research Data infrastructures and has created the Research Data Alliance (RDA) Austria. Austria has invited Jean-Claude Burgelman to know more about the European Open Science Cloud (EOSC) in the RDA Austria launch event.

With regards to Research Data Management (RDM), there is an initiative among 21 universities to issue an RDM policy.

Norway

The Government has issued a national strategy on the use and access of Research Data. The Research Council now requires

projects to have a Data Management Plan to get funding. The content is not being checked. The project teams are free to select what they want. With regard to EuroHPC, there is a level of coordination among the Nordic countries. A letter has been written to the European Commission and PRACE. The national e-Infrastructure provider has received a funding of 12 million for 2018-2019 to upgrade the e-infrastructure. On a question regarding the type of data relevant to Norway, a concrete example was environmental data.

Sweden

Universities and Research Councils have set up a group to look into the coordination of e-Infrastructures including networking, computing, and data. The idea is to create a strong national node and to coordinate among these.

The Netherlands

The Dutch government has recently decided to structurally invest an additional 20 million euro in the Dutch national digital infrastructure. Details will have to be discussed but it is expected that a large fraction of these funds will be channeled through SURF.

Turkey

The Ministry has issued a tender of 20 million euro on e-Infrastructures. For Research Data, a pilot portal has been installed. ❖

The screenshot shows the e-IRG Knowledge Base interface. At the top, there is a navigation bar with 'Start', 'Projects', 'Documents', 'Calendar', 'National Nodes', 'Countries', and 'e-IRG'. Below this, a sidebar on the left lists various categories: Knowledge Base, Organisations, All Projects, Countries, HPC Analysis, HPC Projects, HPC Organisations, EuroHPC members, e-Infrastructure Analysis, e-Infrastructure Projects, e-Infrastructure Organisations, National Nodes, KPI Analysis, KPI Classification, KPI Project Summary, H2020 Topics, and Contact. The main content area is titled 'Countries in Europe' and displays a grid of 15 country flags with their names: Austria, Belgium, Cyprus, Czech Republic, FY Republic of Macedonia, Finland, Greece, Hungary, Italy, Latvia, Malta, and Montenegro.

ECMWF's scalability programme and co-design effort result in proposal for "Extreme Earth" Flagship Initiative

During the e-IRG Open Workshop in Sofia, Bulgaria, Dr. Peter Bauer, Deputy Director of Research at the European Centre for Medium-Range Weather Forecasts (ECMWF), gave a talk on the need for co-design in supercomputers for weather prediction.

Afterwards, we met with Peter Bauer who was happy to join the e-IRG Workshop in Sofia, which he described as a well-organized event bringing together national representatives involved in high performance computing (HPC) strategy and how to handle Big Data in the future, together with application owners like ECMWF.

ECMWF is an international and independent organisation that sits in Reading, in the United Kingdom and was founded by its Member States in 1975 to centralize scientific excellence but also resources in a single place to provide medium-range forecasts that are much better than any national weather service could perform. Since then, this concept has been proven successful because ECMWF is leading medium-range global weather forecasts.

Despite this success over 40 years ECMWF is reaching now a period where realizing the ever increasing forecast skill requirements by Member States of the global community on future computers becomes increasingly difficult. This is because the codes become more and more complex. ECMWF wants to invest in spatial resolution that allows the ECMWF researchers to represent physical processes with more and more accuracy. ECMWF wants to invest in complexity because eventually the scientists have to represent the entire complexity of Earth system processes in a single forecasting

system from atmosphere processes but also from atmosphere chemistry - not only physics - ocean processes, sea ice, land surface processes and the end direction of all these components in the system.

ECMWF also invests a great deal in ensemble prediction. Ensemble prediction has been founded by the meteorological community to also forecast predictional certainties, not only a prediction of a temperature or precipitation in a place, but also the uncertainty of a prediction that changes a lot depending on how good the model and the initial conditions are, of course, but also on regimes. Certain regimes forecasting temperatures in Sofia for the next ten days is easier than in other situations and the researchers need to assess that to provide their users with a feel for the probability or the likelihood of events and give them confidence in their decision making.

All these investments in resolution, complexity and ensembles require enhanced computing. The models become more and more



complex and more and more difficult to run on the current HPC infrastructures. This is why ECMWF needs a concerted investment in proving the models, at the same time investing in computer code development, and together with high performance computing industry and experts to also prove the hardware side of this coin. This is a real co-design effort that ECMWF needs to invest in to be able to run its models with much enhanced prediction skill but at an affordable cost, even in ten years.

How are you working on that? Are you starting new projects?

ECMWF understands the challenge as something that cannot only rely on individual, small research projects but that needs a comprehensive effort actually. For this purpose, ECMWF has founded a so-called scalability programme about five years ago where the researchers said that it is not just a small-scale effect looking at a particular aspect of the forecasting model to make it run faster

but to look at every individual aspect of the entire chain between receiving observational data, dealing with increasing volumes of satellite observations, for example, making the model faster but also producing the forecast products faster for the Member States and users to use.

This entire chain needs to be improved. It requires mathematical changes, coding changes. It requires workflow changes and a tight collaboration between mathematical or algorithmic developments and workflows but in the end the researchers also have to put this on computers that are changing architectures right now. So, ECMWF is aiming for moving targets. ECMWF needs to make sure that it can run these codes on today's technology but also on technologies that one is expecting in 10 or 20 years.

Do you expect that there will be different kinds of architectures for different kinds of applications in the future?

There is right now this discussion between general-purpose computing and domain-specific computing. General-purpose computing is what we had in the past. There were single-type architectures like x86 CPUs that served everybody sufficiently well, that scaled according to the increasing requirements of most applications following Moore's Law. That appears to be coming to a halt now. New technologies appear. There are processors coming for different types of users like GPUs, for

example, FPGAs, other types of CPUs. We need to make sure that we can exploit those. Some of them promise the same type of computational performance, or even better, at much lower power cost. This is something that ECMWF needs to exploit given the specific requirements. The investments that are needed for making ECMWF's codes work and these types of architecture require again investing research in programming models.

One thing is to develop the physics that you put in code and that you want to run on any type of machine but then, some of these architectures require specific programming models to be able to exploit the specific features of that architecture and we need to invest in that. This is not something an application owner can do on his own. It needs to be done in a concerted effort, a so-called co-design.

The scalability programme started in 2013. It has gained significant momentum since then. It

is strongly supported by ECMWF's Member States. ECMWF has succeeded in acquiring a number of externally funded projects, that are participating but also coordinating. These projects are funded by the European Commission's Horizon 2020 programme. They focus on many different aspects like algorithms, programming models, workflows, and these things.

What the researchers realized when they discussed this entire complex within their community is that to really redesign forecasting systems within an entire community requires a much larger effort. ECMWF has formulated this in a proposal to the European Commission for a future Flagship Initiative. Flagships are very ambitious science projects that the European Commission is supporting. These are 10-year projects that try to unite entire communities to perform fundamental scientific research and perform structural changes in Europe serving the European society. Examples are the Human Brain Project, the Graphene Project, and the Quantum Technology Project, for example.

Recently, the European Commission has drafted another Call for ideas for a new generation of flagships. The weather forecasting community has responded to that and formulated an idea that is called "Extreme Earth" which focuses on the entire revision and renovation of forecasting systems, focusing on extremes in the Earth's system, and hence the name. It focuses on extremes forecasting because extremes are very impactful. Just think of the hurricanes that affected last Summer and Autumn season large parts of the southern United States. Just two of them, Harvey and Irma, caused 300 billion worth of damages. Extremes like this, but also droughts, earthquakes, volcanoes are extremes of the Earth system that we need to predict with much enhanced reliability in the future. The Earth system aspect of Extreme Earth brings together what we do in the atmosphere, ocean and sea ice with the solid Earth community. This reflects on the volcano and earthquake aspects of it.

This idea brings together these types of applications with high performance computing as the main enabling technology. With high performance computing is meant the actual computing aspects and the technology mentioned earlier but also the data issues dealing with large volumes and very diverse types of data in the future, from the entire Earth system, how to manage that, how to make this information available so that users across the community but also downstream communities like the energy and food sectors, risk assessment, risk management, national agencies in that context can use the data in a much more efficient way. ❖

National Nodes

as the foundation for the European Open Science Cloud

The European Open Science Cloud (EOSC) will only be as strong as the building blocks it is made of. The building blocks are its national counterparts. The Competitiveness Council conclusions taken on 28-29 May 2018 *state that Member States are encouraged to invite their relevant communities, such as e-infrastructures, research infrastructures, research funding organisations (RFO's) and research performing organisations (RPO's), to get organized so as to prepare them for connection to the EOSC and that the Commission is called on to make optimal use of ongoing projects, existing expertise and knowledge available via existing initiatives, such as ESFRI, e-IRG, GO FAIR and others.* Therefore, e-IRG has put together a Working Group (WG) to gather from the national delegates of all Member States (MS) and Associated Countries (AC) information critical for EOSC in the European Union (EU) MS and AC. The intent is to analyze the current status in detail and develop recommendations and/or name good practices towards the national e-Infra Commons, to ease integration at the EU level.

To achieve this, a specific process has been decided and is in motion. A questionnaire was formed by the WG and sent to the e-IRG delegates who provided information on the organisation of national e-Infrastructures dealing with network, computing, data and services. The requested information concerned issues such as the coordination of national horizontal e-Infrastructures, their governance model, their funding, and the access policies currently in place. There were additional questions on the existence of domain-specific national nodes, and their coordination with national horizontal e-Infrastructures.

The response rate by the national delegates

was satisfactory: 27 countries out of the 32 countries participating officially in e-IRG, have replied to the questionnaires.

With a more detailed study under way, we can clearly see, even with the preliminary analysis made so far, some trends at a European level. Regarding the number of e-Infrastructure providers that exist on these countries, we see that although most countries have more than one provider, there still are a few that have only one single coordinating e-Infrastructure service entity. It is also evident that some countries have ongoing processes aiming at enhancing the coordination of e-Infrastructure initiatives, whilst altogether most countries show a good level of coordination amongst their existing ones.

There are also examples of bottom-up multi-stakeholder coordination of both the generic (network, computing, data), and the domain-specific ones in several counties. As for governance, there is a variety of solution models applied at a European level, with ministries, research councils, universities and research communities having roles in some, but not all, countries. A general trend is obvious in Network e-Infrastructures, which appear to have a different governance model than all other (computing, data, services). Funding is described in interesting ways with mechanisms that channel funds to horizontal e-Infrastructures. When it comes down to access, most countries only allow for national access to resources developed or purchased usually with state funds.

It should be noted that more results will be presented in November Vienna during the e-IRG workshop and in the upcoming document. ❖

<http://e-irg.eu/e-irg-special-edition-nov-2018>

Michalis Maragakis – e-IRGSP5 project.



Interview Michael Resch

German HPC and EuroHPC, a question of competitive collaboration benefiting science

After his presentation at ISC'18 in the session "German HPC in Context", we were able to talk with Michael Resch, Chairman of the Gauss Centre for Supercomputing. His talk covered the German HPC developments, also in connection with the European developments in EuroHPC. Michael Resch told us that all these things fit nicely together because Germany has been very active. It was among the first countries to sign the agreement for the EuroHPC initiative. This fits nicely, both to the German strategy and the German strategy fits nicely into EuroHPC. These are two very complementary activities.

You also mentioned that in Germany the local structure of high-performance computing has been changed as well?

No, we don't have to change the regional or the federal structure in HPC in Germany. It's rather that we have this layered concept. At the very low layer you have universities, locally responsible for HPC at a lower level. Then you have these regional or statewide centres which we call tier two. We work closely with these two levels - tier two and tier three - in Germany. Working together with them guarantees that researchers can actually move in this pyramid of performance from level 3 up to level 0, and maybe even down to level 3 when they do their big simulations and then, while they have to do the evaluation of their results, going back to their local university. The local structure, this federated structure that we have in Germany, very well supports the researchers. It actually allows us to have this hierarchical approach of the European, national, statewide, local level.

There are three big supercomputing centres in Germany. During most of the history of the TOP500, these centres were somewhere in the top 10 or top 15. Now, they are a little bit behind. Is that their natural place?

There is a gap now because we had a project in which we had the funding for systems from 2007 to 2019. We recently started a new project which gives us funding from 2017 until

2025, maybe 2026. The Jülich system came in slightly later than planned, so basically the big system in Jülich is probably delayed by one year. It is partially a delay in the funding and partially a delay in technological developments that we expected to happen. This is nothing that actually frightens us. The fastest system is ranked number 23. That is not good but we will be back in the game in November with the Munich system and in June 2019 with the Jülich system. Then you will see German systems in the top 10. This is a normal situation. Every big centre goes down and the good thing about us is that we usually don't go further down than probably rank 20. 2018 is very much an exceptional year.

You also need exceptional years.

Yes, sometimes you have exceptional years because you are better; sometimes you have exceptional years because it is like the worst but that's okay.

In the plans for Germany, you also aim to have exascale systems around 2023.

Absolutely. Professor Lucas from the Federal Ministry of Science announced this last year. We are working on a strategy where all the three centres move up to exascale level. We want one of the centres, namely Jülich, to be an exaflop centre for EuroHPC. This shows how we can merge the two strategies, the European strategy and the German strategy, Europe having the ability to fund two exaflop systems, Germany having the ability to fund much more than that. So, it will be a Jülich system with an exaflop and then Munich and Stuttgart will follow to reach this level probably slightly later. We will also be using different technologies to support different communities.

Being part of EuroHPC means that from a German point of view, you could get much quicker access to exascale systems than you would have normally?

I don't think it speeds up the process too much but there is more funding because you pull the resources in that case for Jülich. You pull the resources of Germany and Europe and by bringing the two of them together you reach the exaflop level much earlier than Germany alone could do. That is an advantage.

In EuroHPC, they are talking about two things: firstly, about making machines available for users and applications – and of course, that is a very important part which a lot of people tend to forget – and the other thing is to have exascale systems based on European technology. What do you think about these things: is one more important than the other?

It is obvious that we have to support European science and European society. One



approach is to provide the best technology. This means that you just look at the level of performance that scientists need, you look at the kind of support that they need for their simulations. That is sort of the number, the exaflop number. This is what you would like to reach to be competitive internationally when it comes to science and also when it comes to industrial usage. The other thing is this: Europe is a country or Europe is a continent - I call it a country, so I assume I am a good European - where we don't have big national IT players in the sense that the United States have an IBM or Japan has NEC. That puts us in a position where we have an advantage of competing companies but we have also a certain disadvantage of not having a local player.

Developing European technology for us as a centre means that we have a local player and there is more diversity in the market which is positive. These two things go together but it doesn't mean that if you have European technology we need to have exactly one exaflop or if we have exactly one exaflop it has to be only European technology. The target is to support science and society. This is an interplay of a variety of activities that are going on in HPC.

The advantage of having European technology to some level is that also the industry can benefit of course.

I don't know whether industrial people would see this the same way but I think they would be happy - as far as I work with industry - to have more choice. It is certainly easier for them to work with a European vendor in the same way as it is easier for us because of time zone differences and travel times if you have problems with things like that. However, I think that industry is less local than the European supercomputer centres are.

Take an organisation or a company like Daimler. They have their development, research, production, manufacturing, and marketing whatsoever worldwide. They are not that much national anymore. This is something we see with big companies but also with small and medium-sized companies.

In the state of Baden-Württemberg we have companies with some 600 people working in



25 different countries all around the globe. Industry has a different view of this. For them it is not something like: it is either Europe or it is not Europe because they are anyway playing on several fields.

Interview Thomas Lippert

to propose modular way of thinking to optimize both supercomputing technology and HPC policy

At the ISC'18 Conference in Frankfurt, Germany, we had the opportunity to talk with Thomas Lippert, who has recently become Chair of PRACE. Apart from that, he is also Director of the Jülich Supercomputing Centre where the first module of a new HPC system has been installed. This system is procured in the framework of the Gauss Centre for Supercomputing in Germany. The machine of about 10 Petaflops CPU performance and 2 Petaflops of GPU performance is called JUWELS.

This first module is part of a modular supercomputer. Does this mean that the modular supercomputer will have different kinds of hardware in it?

It means that the modular supercomputer has entities which are supercomputers by themselves, called modules: Let's say a system composed of CPUs mainly and another system that is concentrating on the use of GPUs. These modules are not brought together just by a standard connection or through storage technology but they are brought together directly on a joint network,

Does this also hold for SMEs? You have a number of consulting companies, for instance?

It depends. We work with medium-sized companies with 600, 800 or 1000 people. They have subsidiaries in some 25 countries. We also have small and medium sized companies that are only located around Stuttgart, for example. For them it makes a difference. The others are more international.

You mentioned that Germany, and especially Jülich, would be a candidate to have one of the first EuroHPC machines. Barcelona would also be a candidate.

Yes, this is good. It is a process of competition. Science is about competing ideas, science is about competition in a sense of having better technology, better ideas,

maybe with a network bridge between two different network technologies or on the same network type. This connection between both modules is substantial. It is possible that a code that consists of a workflow can be run primarily on the cluster part and then maybe on the second part, the second module still to come which will be highly scalable.

You can also think of codes that have different parts, that have different scalability properties, so that the code can be divided according to these different scalability properties to parts that run on the cluster and parts that run on the booster, the other module. This is a way through which one can simplify the technology of the scaling module and concentrate a lot of technology on the cluster module. This will enable us to do the high scalable parts more cheaply than on a full scale cluster system.

The system is now up and running, at least the new part?

The first part is up and running. The second part will come next year. But the system is already in the TOP500. This is the cluster-booster system called JURECA.

It is very good that you mention the TOP500 because Erich Strohmaier in his presentation showed a comparison of HPC centres. The Jülich Supercomputing Centre was at rank 8. He compared 11 centres over the years. The only European centre that was consistently in the ranks was the Jülich Supercomputing Centre.

In some sense, this is by chance because the way Jülich does computing is never to try to be number 1 or to be very far up in the TOP500 list. Jülich has consistently been in the TOP500 over the years because it has two or three machines in the TOP500 list. Jülich tries to concentrate its budget on optimizing the computing for its portfolio. That is the reason why the Jülich Centre might be appearing in this list of 11 centres.

better concepts, better strategies. I expect other candidates to come forward and hand in their candidacy. We will just see. The Gauss Center is in a strong position. We do excellent science, we have excellent systems, we have big centres, very good centres, so we are not afraid of a competition. In science, you have collaboration and competition. You could call it collaborative competition or competitive collaboration. That is normal.

I think this is one of the strengths of Europe. There is not a single government that probably makes a decision about this organisation being good and this organisation being not good. We are competing at the European level and we are competing, each of us, at the national level. That makes us better and that is positive. ❖

One wanted to show that centres are more than just one or two machines.

Exactly. Of course, with this experience where Jülich usually had a system which was more or less a cluster system and another one that was more or less something like a scalable system, like the BlueGene, we found out it would be better to combine both technologies into one computer because this is what we need in the future where we will have more heterogeneous types of computing, a more workflow type of computing and different concurrencies in one highly scalable code. These different concurrencies might fit better to different technologies.

Can we switch to PRACE of which you are now the Chairman? What are the main topics for PRACE in the future?

Let's look a little bit back to PRACE. PRACE has developed very much into a high-quality provider; high quality in the sense of the science provider of supercomputing. The organisation is qualified as THE provision entity for supercomputing cycles. Given the future developments of supercomputing in EuroHPC, one must see how we can utilize those virtues of PRACE, which are not only the cycles but which are also things like training people, going into code improvement, giving support in the so-called High Level Support Teams (HLSTs). We can utilize all those virtues also further for the upcoming pre-exascale and exascale technologies to be provided by EuroHPC.

This would be maximizing the optimal engagement of all parties, including the users, the provisioners through PRACE, and of course the funders and builders of technology. I want to bring PRACE together with EuroHPC so that both activities in a maximal way will profit and benefit, for the benefit of science, the users and industry. ❖



In this issue

- ✓ **Special Edition of Open e-IRG Workshop 2**
- ✓ **e-IRG Knowledge Base 3**
- ✓ **DI4R2018 Conference report 3**
- ✓ **Highlights from the delegates meetings 4**
- ✓ **European Open Science Cloud 8**
- ✓ **ESFRI's 5th Strategy Report & Roadmap 2018 9**
- ✓ **Interview Simon McIntosh-Smith 11**
- ✓ **PLAN-E Workshop @ IEEE eScience Conference 2018 12**
- ✓ **Interview Ognjen Prnjat 14**
- ✓ **e-IRG Workshop Sofia 15**
- ✓ **Interview Ivan Dimov 18**
- ✓ **Around Europe 19**
- ✓ **ECMWF's scalability programme results in proposal for "Extreme Earth" Flagship Initiative 20**
- ✓ **National Nodes 21**
- ✓ **Interview Michael Resch 22**
- ✓ **Interview Thomas Lippert 23**

Upcoming events

- 11-16 November 2018 - *International Conference on High-Performance Computing, Networking, Storage and Analysis - SC18, Dallas, Texas, USA*
<http://sc18.supercomputing.org/>
- 12-14 November 2018 - *Second European Big Data Value Forum, Vienna, Austria*
<http://www.european-big-data-value-forum.eu/>
- 20-21 November 2018 - *e-IRG Open Workshop - Special Edition, Vienna, Austria*
<http://e-irg.eu/e-irg-special-edition-nov-2018>
- 21-22 November 2018 - *Second EOSC Stakeholders Forum, Vienna, Austria* <https://eoscipilot.eu/events/second-eosc-stakeholders-forum>
- 23 November 2018 - *Launch of the European Open Science Cloud (EOSC) Governance Structure, Vienna, Austria*
<https://www.eu2018.at/calendar-events/political-events/BMBWF-2018-11-23-EOSC-Launch.html>
- 4-6 December 2018 - *Imagine Digital ICT 2018: Connecting Europe Conference and Exhibition, Vienna, Austria*
<https://ec.europa.eu/digital-single-market/en/events/ict-2018-imagine-digital-connect-europe>

About e-IRG

e-IRG is a strategic body to facilitate integration in the area of European e-Infrastructures and connected services, within and between member states, at the European level and globally.

The mission of e-IRG is to support both coherent, innovative and strategic European e-Infrastructure policymaking and the development of convergent and sustainable e-Infrastructure services.

The e-IRG produces white papers, roadmaps and recommendations, and analyses the foundations of the European Open Science Cloud and European Data Infrastructure.

<http://e-irg.eu>

e-IRG
e-Infrastructure
Reflection Group

e-IRG secretariat

c/o The Netherlands Organisation for Scientific Research (NWO)
P.O. Box 93460
NL-2509 AL The Hague
The Netherlands

Phone: +31 70 344 0526
Mobile: +31 6 303 699 04
email: secretariat@e-irg.eu

Visiting address:

Java Building,
Laan van Nieuw Oost-Indië 300
NL-2593 CE The Hague
The Netherlands

e-IRGSP5

This Magazine is produced by the e-IRGSP5 project. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730954.

<http://e-irgsp5.e-irg.eu>

