



2ND EUDAT CONFERENCE

28 OCTOBER 2013 - 30 OCTOBER 2013 - ROME, ITALY

DOCUMENTATION

The 2nd EUDAT conference addresses new data challenges – the case for cross-disciplinary science and services – in view of the European Commission’s imminent Horizon 2020 programme. International cooperation and collaboration aspects of e-infrastructures, with a particular focus on the recently launched Research Data Alliance (RDA), are also high on the agenda together with input from the community-driven initiatives collaborating with EUDAT. The essential role and contribution of industry – as a user and service provider, and hence key stakeholder in this infrastructure – is also covered.

The Eternal City sets the backdrop for this gathering of international data infrastructure providers and practitioners from academia and industry to discuss the opportunities offered through the Collaborative Data Infrastructure (CDI) supported by EUDAT.

Follow EUDAT on Twitter @eudat_eu and use the #EUDATC2 to follow the conference via Twitter

V2.1 dated 29th October 2013



EUDAT is co-funded by the EC's Seventh Framework Programme (FP7) under grant agreement n° 283304

TABLE OF CONTENTS

WELCOME	4
LOGISTICS.....	5
MEDIA PARTNERS.....	6
TRAINING TRACK 1.1: Data Staging, Replication and Storage: Integrating with EUDAT's Building Blocks	9
TRAINING TRACK 1.2: Implementation of Staging, Replication and Storage: Services and Tools	10
TRAINING TRACK 2.1: Persistent Identifiers, Handles, Type Registries, EPIC.....	11
TRAINING TRACK 2.2: METADATA.....	12
WORKSHOP: CumuloNimbo - THE SKY IS NO LONGER THE LIMIT FOR BIG DATA!	13
WORKSHOP: Data Sharing and Enrichment: iMarine and Lifewatch Solutions & Experiences.....	14
WORKSHOP: SIM4RDM Cross-Stakeholder Research Data Management	15
WORKSHOP: Social Sciences and Humanities (SSH) tackle the big data challenge.....	17
PLENARY SESSION I - Addressing the new data challenges – the case for cross-disciplinary science and services.....	21
PLENARY SESSION II - Life and Earth Sciences at a cross-road: community driven flagship initiatives in the EU and USA.....	23
PARALLEL SESSION: TRACK 1 - EUDAT Services -1.1 Safe Replication and Data Staging.....	25
PARALLEL SESSION: TRACK 1 - EUDAT Services -1.2 SIMPLE STORE	27
PARALLEL SESSION: TRACK 2 - INTEROPERABILITIES - 2.1 FEDERATED AAI	29
PARALLEL SESSION: TRACK 2 - INTEROPERABILITIES - 2.2 PUBLIC-PRIVATE PARTNERSHIPS	31
PARALLEL SESSION: TRACK 3 - Policy and Sustainability issues - 3.1 Access and Re-Use Policies	34
PARALLEL SESSION: TRACK 3 - Policy and Sustainability issues - 3.2 Data Management Plans and Certification	36
PARALLEL SESSION: TRACK 4 – NEW SERVICES - 4.1 SEMANTIC ANNOTATION.....	38
PARALLEL SESSION: TRACK 4 – NEW SERVICES - 4.2 DYNAMIC & REAL-TIME DATA	40
PARALLEL SESSION: TRACK 1 - EUDAT Services -1.3 METADATA	43
PARALLEL SESSION: TRACK 2 - INTEROPERABILITIES - 2.3 IDENTIFIERS	45
PARALLEL SESSION: TRACK 3 - Policy and Sustainability issues - 3.3 COST & FUNDING MODELS	47
PARALLEL SESSION: TRACK 4 – NEW SERVICES - 4.3 WORKFLOWS	49
PLENARY SESSION III - Towards Global Data Infrastructure Components	50



PLENARY SESSION: WRAP-UP AND CONCLUSIONS	53
WORKSHOP: PID Information Types WG Meeting	55
WORKSHOP: DIGITAL PRESERVATION OF CULTURAL DATA.....	56
PARTICIPANTS LIST	58

WELCOME

For its second edition, the EUDAT conference is bringing together data infrastructure providers and practitioners from across the globe to discuss current data infrastructure challenges and solutions, with a focus on interoperability, cross-disciplinarity and cross-border collaboration.

International e-infrastructure cooperation and collaboration benefits are high on the agenda, with keynotes by **Richard Frackowiak** (on the Human Brain flagship project), **Ewan Birney** (on Annotating the Human Genome), **Maryline Lengert** (on Helix-Nebula the Science Cloud), and a view from across the Atlantic from **Bill Michener** (on DataONE). Kostas Glinos will present the Commission's plans for Horizon 2020 and progress reports on a series of **Research Data Alliance (RDA) activities** and working groups will round up the event.

Participants will also be offered a 360° view of EUDAT's existing and future services and how they will benefit both researchers and research organisations in accessing and preserving their research data. Parallel tracks will give a complete presentation of the first set of EUDAT services (B2SAFE to replicate research data safely, B2SHARE to store and share long-tail research data, B2FIND to find research data, B2STAGE – to get data to computation) and new services that are in the pipeline (such as Dynamic Data and Semantic Annotation). Federated AAI, persistent identifiers, and public-private partnerships will be discussed in a track focusing on interoperability, while a fourth track on Policy and Sustainability issues will address data management plans, data access and reuse policies, and cost and funding models.



We would like to give you the opportunity to have a say in the selection and development of the next set of EUDAT services. As we are now discussing the next set of services to be designed, we are inviting you as data infrastructure users and funders to express your interests and requirements. Please take some time to complete this survey and share your views on which types of data services would be beneficial for your research community. To thank you for your participation, everyone who completes and submits this short survey (<http://www.webropolsurveys.com/S/22DD04AFBE3B3533.par>) will be entered in a

draw to win a prize – the results will be announced at the Wrap-up & Conclusions Session on Wednesday 30th October.

This year we wanted to provide you again with a great opportunity to engage with EUDAT members and other data experts through a rich three-day programme interspersed with many networking opportunities.

We hope you enjoy the conference!

Kimmo Koski

Peter Wittenburg

Damien Lecarpentier

EUDAT Project Coordinator

EUDAT Scientific Coordinator

EUDAT Project Manager

LOGISTICS

CONFERENCE VENUE

HOTEL NH VITTORIO VENETO
Corso d'Italia 1
00198 Rome (Italy)
Nearest Metro Stop: Barberini

USEFUL TELEPHONE NUMBERS

- 112 General Emergency, gendarmerie (Carabinieri)
- 113 General Emergency, Police (Polizia di Stato)
- 115 Fire and rescue service (Vigili del Fuoco)
- 117 Italian Customs/Financial/Border Police Guardia di Finanza
- 118 Ambulance and first aid service (Italian Red Cross included)

COCKTAIL RECEPTION MONDAY 28TH OCTOBER 2013 @ 18:30

A networking cocktail reception will take place from 18:30 to 20:00 at the NH Hotel, all participants are welcome to attend.

CONFERENCE DINNER TUESDAY 29TH OCTOBER 2013 @ 20:00

The conference dinner takes place at Ristorante da Meo Patacca, Piazza dei Mercanti 30, 00153 Roma
The restaurant is located in the Trastevere area of Roma and take a approx 20 minutes to walk there from the NH Hotel. Transport will be provided for all those who indicated they would attend the dinner on the registration form. The bus leaves the NH Hotel at 19:30 sharp, buses will return to the NH Hotel following dinner. Special dietary requirements indicated at the time of registration will be taken into consideration for the dinner. For those wishing to make their own way to the restaurant, from NH Hotel, walk for 300 m to San Paolo del Brasile, then take the 160 bus (to Rufino) and get off at Greca (after 12 stops). Cross the river Tevere on Ponte Palatino, keep left and take the second one on your right (via Peretti Pietro) and go straight on until you get to Piazza dei Mercanti 30.

MOBILE PHONES

As a courtesy to participants and speakers, please turn off your mobile phones or switch to silent mode during the conference sessions.

TRAINING SESSION VIDEO & AUDIO RECORDING

The EUDAT training sessions are being video & audio recorded live so we ask you to pay particular attention to switching off or to silent mode your mobile phones before entering the room and to make as little noise as possible while entering. Thank you for your co-operation.

MEDIA PARTNERS



INTERNATIONAL SCIENCE GRID THIS WEEK

[iSGTW](#) is a weekly online publication that covers distributed computing and the science it enables. It reports on the latest research across all scientific fields and provides in-depth analysis of all aspects of distributed computing technology, including grids, clouds, HPC, and more. The publication covers many of the top events related to e-infrastructures and regularly publishes contributions from leading experts in the field.

iSGTW is also an online destination where you can find information about events, deadlines, and jobs. With editors based in both Europe and the US, as well as contributors from around the world, the publication has a truly global focus. It is read in over 190 countries and has received over half a million page views over the last year. Now eight years old, the publication celebrated its 400th issue in June.

You can join our growing community of over 9,000 email subscribers here: www.isgtw.org/subscribe. Please feel free to follow us on our [Facebook](#) and [Google+](#) pages. We're also on Twitter as '@isgtw'.



DIGITAL MEETS CULTURE

An interactive online magazine where digital technology and culture collide. Join in the collection and sharing of information and events, on a truly global scale.

digitalmeetsculture.net is a communication platform run by company Promoter, and it was designed as an on-line magazine about the digital culture for collecting and sharing information and events, in a global dimension. Currently, the magazine counts over 7.000 unique visitors per month, and more than 22.000 page visualizations.

digitalmeetsculture.net is addressed to people belonging to different sectors as the cultural heritage (museums and libraries managers, researchers, technicians, who are interested in the digitization process and technologies), and the technical and scientific area (universities, professionals, IT experts, e-infrastructure providers and researchers); to people interested in the creative uses of the digital art (artists, photographers, performers, art critics), to people belonging to the educational sector (teachers, students, vocational trainers) and also to general users who want to be informed and up-to-date on these matters.

It represents a valuable information tool as well as an advertising showcase towards a very selected and high-profiled audience, providing extensive information about international projects and initiatives for digitization, preservation, data management and enjoyment of digital cultural heritage.

Interactivity is also a key-point in **digitalmeetsculture.net**: like any modern communication website, users can join the portal, thus being allowed not only to comment the articles, but more importantly to play an active part by sending events and articles, right up to the point of joining the staff if they wish.

Beside general contents about culture, art and technologies, the portal offers interesting opportunities for the dissemination and web-presence powering of scientific projects: by giving wide visibility to the project's activities and achievements, by deepening their topics through interviews and related articles, and by highlighting news and advertising events as media partner.

In this light, **EUDAT** and **digitalmeetsculture.net** recently started a collaboration that is not limited to the media partnership for the 2nd EUDAT conference. The activities of **EUDAT** and the hot topic of e-infrastructures can count on **digitalmeetsculture.net** to reach and be disseminated to an even wider audience, especially related to the digital cultural heritage community that is actively looking for research and storage solutions, including the use of e-infrastructures, as witnessed for example by the achievements of EU project DCH-RP (also present at the 2nd EUDAT conference with a dedicated workshop "Digital Preservation of cultural data")



Monday 28th October 2013

EUDAT Training & Associated Workshops

ROOM	RAFFAELLO	PUCINI	TIZIANO	TINTORETTO	BOTTICELLI	MASCAGNI	VERDI
09:00 - 11:00	EUDAT Training Track 1.1: Data Staging, Replication and Storage: Integrating with EUDAT's Building Blocks	EUDAT Training Track 2.1: Persistent Identifiers, Handles, Type Registries, EPIC		CumuloNimbo – The sky is no longer the limit for Big Data!			
11:00 - 11:30			<i>Coffee</i>				
11:30 - 13:00	EUDAT Training Track 1.1: Data Staging, Replication and Storage: Integrating with EUDAT's Building Blocks	EUDAT Training Track 2.1: Persistent Identifiers, Handles, Type Registries, EPIC		CumuloNimbo – The sky is no longer the limit for Big Data!	Social Sciences and Humanities (SSH) tackle the Big Data Challenge	<i>Knowledge Exchange - Research Data working group (Closed)</i>	
13:00 - 14:00			<i>Networking Lunch</i>				
14:00 - 16:00	EUDAT Training Track 1.2: Implementation of Staging, Replication and Storage: Services and Tools	EUDAT Training Track 2.2: Metadata	SIM4RDM - Cross-Stakeholder Research Data Management	Data Sharing and Enrichment: iMarine and Lifewatch Solutions & Experiences	Social Sciences and Humanities (SSH) tackle the Big Data Challenge	<i>Knowledge Exchange - Research Data working group (Closed)</i>	
16:00 - 16:30			<i>Coffee</i>				
16:30 - 17:00	EUDAT Training Track 1.2: Implementation of Staging, Replication and Storage: Services and Tools	EUDAT Training Track 2.2: Metadata	SIM4RDM - Cross-Stakeholder Research Data Management	Data Sharing and Enrichment: iMarine and Lifewatch Solutions & Experiences	Social Sciences and Humanities (SSH) tackle the Big Data Challenge	<i>Knowledge Exchange - Research Data working group (Closed)</i>	
17:00 - 18:00	EUDAT Training Track 1.2: Implementation of Staging, Replication and Storage: Services and Tools	EUDAT Training Track 2.2: Metadata	SIM4RDM - Cross-Stakeholder Research Data Management	Data Sharing and Enrichment: iMarine and Lifewatch Solutions & Experiences			<i>EUDAT GENERAL COUNCIL (Closed)</i>
18:00 - 19:00							<i>EUDAT GENERAL COUNCIL (Closed)</i>
18:30 - 20:00	NETWORKING COCKTAIL @ HN VITTORIO VENETO						

EUDAT TRAINING COURSES - MONDAY 28TH OCTOBER

TRAINING TRACK 1.1: DATA STAGING, REPLICATION AND STORAGE: INTEGRATING WITH EUDAT'S BUILDING BLOCKS

DATE & TIME: MONDAY 28TH OCT – 09:00 – 13:00

ROOM: RAFFAELLO

OVERVIEW:

A key benefit of a collaborative data infrastructure is the opportunities it offers for moving (replicating, uploading) data more easily for a variety of purposes such as preservation, access optimisation, to improve sharing, and to prepare data for processing. Data is precious: particularly that which corresponds to observations of real-life systems (such as seismic activity, ocean temperatures, observations about species and languages/cultures) that may not be reproducible. In these cases, having reliable, off-site copies of your data can be invaluable. Also, sharing of data between scientists, organisations and disciplines is becoming more common, and provides possibilities of new approaches to problems by bringing together data from multiple sources. Finally, in order to have data readily available for computation or analysis it needs to be moved from where it is stored, to where it is processed (and back again).

All of these use cases require that data can be moved efficiently and reliably. This session covers some of the technologies and techniques that can be used to do this, and will also discuss these in the context of EUDAT. After attending this session, you will know which technologies uses to provide its Data staging, Safe replication and SimpleStore services and how they are used. You will know the benefits of using these services, and have learned what the underlying technologies provide to help with moving of data. You will know how these EUDAT building blocks could be integrated with existing services.

AGENDA:

TRAINING TRACK 1.1: DATA STAGING, REPLICATION AND STORAGE: INTEGRATING WITH EUDAT'S BUILDING BLOCKS	
09:00-09:15	Arrival & Registration
09:15-10:00	Introduction
10:00-10:45	Data Replication & Data Staging Services
10:45-11:15	Break
11:15-12:00	SimpleStore: Introduction
12:00-12:45	iRODS: What it is, what it can do

CONTACT PERSON: ADAM CARTER, EPCC - ADAM.CARTER@ED.AC.UK

EUDAT TRAINING COURSES - MONDAY 28TH OCTOBER

TRAINING TRACK 1.2: IMPLEMENTATION OF STAGING, REPLICATION AND STORAGE: SERVICES AND TOOLS

DATE & TIME: MONDAY 28TH OCT – 14:00 – 18:00

ROOM: RAFFAELLO

OVERVIEW:

In this session we will go into more depth regarding the technologies used in EUDAT's data services. We'll show you how iRODS has been set up and configured, and in particular how it has been integrated with services like EPIC to obtain PIDs. We will describe how you can use GridFTP to stage your data on to the EUDAT infrastructure or between EUDAT and HPC resources and we'll show you how EUDAT's data staging script can be used to help you make use of services like Globus Online. We'll talk about how repositories like those based on Fedora Commons can be integrated with the Safe Replication service and show a concrete example of how this has been done.

AGENDA:

TRAINING TRACK 1.2: IMPLEMENTATION OF STAGING, REPLICATION AND STORAGE: SERVICES AND TOOLS	
14:00-14:45	Practical : iRODS Hands On
14:45-15:30	Practical : Safe Replication
15:30-16:00	Coffee
16:00-16:15	GridFTP
16:30-17:00	GridFTP hands-on
17:00-17:30	Safe Replication & Repositories
1730	Close

CONTACT PERSON: ADAM CARTER, EPCC - ADAM.CARTER@ED.AC.UK

TRAINING TRACK 2.1: PERSISTENT IDENTIFIERS, HANDLES, TYPE REGISTRIES, EPIC**DATE & TIME: MONDAY 28TH OCT – 09:00 – 13:00****ROOM: PUCCINI****OVERVIEW:**

In this session we'll review how PIDs can be used, and we will discuss how the handle system can be used with a type registry to provide functionality beyond basic handle resolution. The session will also include a hands-on tutorial on the use of the EPIC API which can be used to construct or interact with the EPIC Handle service.

AGENDA:

TRAINING TRACK 2.1: PERSISTENT IDENTIFIERS, HANDLES, TYPE REGISTRIES, EPIC	
09:00-09:15	Arrival & Registration
09:15-09:45	Very brief introduction to PIDs.
09:45-10:45	How the handle system can be used with a type registry to provide functionality beyond basic handle resolution.
10:45-11:15	Break
11:15-12:00	EPIC and Handles in EUDAT
12:00-12:45	EPIC/iRODS practical

CONTACT PERSON: ADAM CARTER, EPCC - ADAM.CARTER@ED.AC.UK

TRAINING TRACK 2.2: METADATA**DATE & TIME: MONDAY 28TH OCT – 14:00 – 18:00****ROOM: PUCCINI****OVERVIEW:**

Metadata is going to serve increasingly more functions. Traditionally it is used for finding useful data and tools. In future it will be increasingly often used for scientific purposes and for orchestrating scientific workflows. It is therefore not surprising that metadata remains the focus of discussions and deservedly affords the attention of the Research Data Alliance. Data is really only of use in research when its context and provenance is understood. Context such as when it was created, who it was created by, to what it relates, how it was created and how it can be used is even more vital to understand once data has been stored or shared. Metadata provides this context and is key to being able to make the most of data by making it findable, understandable and reusable.

This session covers the principles of Metadata, introduces important metadata efforts such as Dublin Core, ISO standards and in particular community based solutions to foster science and compares approaches such as the use of fixed schema and meta models. It will also discuss what the intentions and architectures of major initiatives such as EUDAT, DataONE, Europeana are, how the work in this area is carried out, how one can participate in the EUDAT services and how MD services can be used. We'll also show how Metadata is used in EUDAT's SimpleStore and how this has been built on Invenio.

AGENDA:

TRAINING TRACK 2.2: METADATA	
14:00-14:30	Metadata Standards and Interoperability
14:30-15:00	Metadata Dissemination
15:00-15:45	How YOU can publish your metadata
15:45-16:15	Break
16:15-17:00	Metadata and SimpleStore / Invenio
17:00-17:30	Metadata in EUDAT / DataONE and Europeana

CONTACT PERSON: ADAM CARTER, EPCC - ADAM.CARTER@ED.AC.UK

DATE & TIME: MONDAY 28TH OCT – 09:00 – 13:00**ROOM: TINTORETTO**

Big Data: Only for analytical processing? These days that idea is ancient history. Now, with CumuloNimbo, Big Data can deal with any workload!

So far, Big Data has been a synonym for performing large analytical queries on massive amounts of data. However, in the CumuloNimbo workshop at the 2nd EUDAT conference, we will see how Big Data is getting even bigger, thanks to the new advances brought about by the CumuloNimbo project in ultra-scalable transactional processing.

Big Data is no longer constrained to just the analytical processing world. Unlike today, transactional workloads in the near future will scale out linearly without having to resort to sharding (which frequently requires that substantial modifications are made to applications). An extra benefit is that users will not have to resort to NoSQL (which currently implies sharding) in order to scale. If SQL is needed, it can be used and will scale out. If NoSQL is enough, it can be used and fewer resources will be consumed for query processing. If both SQL and NoSQL are needed, the two interfaces will be available to access the same data store in a coherent manner. Additionally, since updates will now scale as much as is needed, it will be possible to have Big Data combined with complex event processing (CEP) on massive event streams. This approach is based on enabling CEP queries to correlate these massive event streams over stored Big Data at network rates.

With these new advances, the Big Data world will be transformed to incorporate any workload from online analytical processing (OLAP) to online transaction processing (OLTP) and even to complex event processing. The new Big Data milieu will also support all the different data management paradigms available – from NoSQL technologies to CEP and SQL. Additionally, it will be possible to integrate specialized data stores (such as graph databases or document-oriented data stores) with SQL and thus applications will be able to exploit the best of the SQL and NoSQL worlds.

To learn more about the exciting advances in Big Data that are being brought about by the CumuloNimbo project, please join us at the CumuloNimbo workshop at the 2nd EUDAT conference.

Agenda

- CumuloNimbo Overview.
- Ultra scalable transactional processing
- SQL support over NoSQL
- Beyond HBase
- CoherentPaaS and LeanBigData
- Discussion

For more information, see <http://cumulonimbo.eu>.

CONTACT PERSON:

Ricardo Jimenez, Technical University of Madrid (UPM)

WORKSHOP: DATA SHARING AND ENRICHMENT: IMARINE AND LIFEWATCH SOLUTIONS & EXPERIENCES

DATE & TIME: MONDAY 28TH OCT – 14:00 – 18:00

ROOM: TINTORETTO

OVERVIEW:

Data and Interoperability, in themselves challenging subjects, are also closely tied to research communities. Data sharing across the boundaries of communities and organizations, in the sense that data is usable in other contexts than those for which it has been generated, is a pressing need to guarantee richer, better quality and timely science-based knowledge creation. Addressing this challenge requires approaches that not only facilitate access to data produced by others but are also able to transform and enrich the shared data in a way that makes them suitable for being consumed in different contexts.

The goal of this workshop is to highlight these major interoperability issues as well as to showcase and compare the diverse yet complementary approaches and solutions developed by different stakeholders to promote data interoperability on a large scale. Experiences discussed will include:

- the ones gained by the iMarine (www.i-marine.eu) consortium developing a cutting-edge e-infrastructure and fostering a collaborative approach to open data access and interoperability, by supporting marine and biodiversity specialist communities to unlock knowledge and support science policy decision making.
- Lifewatch's (www.lifewatch.eu) showcase "Patterns of ecosystem fragility to alien and invasive species in Europe", its tools and data services. The showcase has collected and shared data from existing research provided by some of the participating universities and institutions. The data includes about 11570 species from 314 different terrestrial and marine ecosystems.

AGENDA:

14:00 - 14:20	Welcome and Workshop Objective : D. Castelli - ISTI-CNR
14:20 - 15:10	LifeWatch: Standards required/used by the community: H. Schentz - Umweltbundesamt GmbH, A. Oggioni - LTER
15:10 - 16:00	iMarine: Accessing and managing Biodiversity data - P. Pagano - ISTI-CNR
16:00 - 16:15	Coffee Break
16:15 - 17:00	LIFEWATCH ICT CORE bricks: Integration of preparatory phase projects developments for the coordination and management of the distributed e-Infrastructure construction : Antonio José Sáenz Albanés - LW ICT-Core, Daniel Fuentes - LW ICT-Core
17:00 - 17:45	iMarine: Analyzing and processing Biodiversity data: A. Manzi - CERN, G. Coro- ISTI-CNR, P. Pagano- ISTI-CNR
17:45 - 18:00	Conclusions : N. Fiore - LW Service Centre

CONTACTS:

iMarine – Donatella Castelli, ISTI CNR - donatella.castelli@isti.cnr.it

Lifewatch – Nicola Fiore, Univeristà di Salento - nicola.fiore@unisalento.it

WORKSHOP: SIM4RDM CROSS-STAKEHOLDER RESEARCH DATA MANAGEMENT**DATE & TIME: MONDAY 28TH OCT – 14:00 -18:00****ROOM: TIZIANO**

The SIM4RDM project aims to improve current policies in the area of managing research data. The project has surveyed a selection of research data management (RDM) stakeholders across the EU and is currently building a self-assessment tool to help them evaluate and improve their RDM maturity.

This tool is based on an online questionnaire and produces overall maturity scores and detailed results measured against five scales, thus allowing users to decide what RDM areas they wish to improve first. The tool will suggest sets of concrete activities designed to increase the user's RDM maturity, and users are also pointed to online resources and case studies to help them complete these activities.

The process of developing the self-assessment tool revealed that there are overlaps in the activities suggested to funding bodies, infrastructure providers, research institutions, publishers and researchers. However RDM policies and recommendations are usually developed in cultural and sectorial isolation: senior executives in these stakeholder groups have very few opportunities to discuss RDM issues of common interest with each other.

SIM4RDM has therefore planned a workshop at the 2nd EUDAT Conference to offer senior executives (from funding bodies, infrastructure providers, research institutions, publishers and research centres across Europe) a unique opportunity to provide feedback on the functionality of the self-assessment tool, and to explore opportunities for cross-stakeholder collaboration on RDM issues of common interest.

The workshop is structured in two parts. In the first part, the participants will highlight RDM issues that are crucial to their activities and which they would like to explore through cross-stakeholder collaboration. In the second part of the workshop, the participants will engage in small-group exercises to provide feedback on the functionality of the self-assessment tool and to explore some of the opportunities for cross-stakeholder action highlighted earlier in the workshop.

AGENDA:

13.00 Lunch

14.00 Project overview and event outline - Matthew Dovey and Gabriel Hanganu, Jisc

14.15 RDM collaboration pitches

- Simon Hodson, CODATA
- Ingrid Dillo, DANS
- Jonathan Tedds, Leicester University
- Rebecca Lawrence, F1000 Research Ltd.
- Leif Laaksonen, CSC

15.30 Coffee break



15.50 Interactive hands-on session (attendees divided in small groups work on the collaboration pitches and their potential integration in the sim4rdm framework)

17.40 Reports from groups

17.55 Wrap-up and close

Further information about SIM4RDM is available on the project website (www.sim4rdm.eu).

CONTACT:

Gabriel Hanganu - Jisc Programme Manager SIM4RDM

Email: g.hanganu@jisc.ac.uk

WORKSHOP: SOCIAL SCIENCES AND HUMANITIES (SSH) TACKLE THE BIG DATA CHALLENGE

DATE & TIME: MONDAY 28TH OCT – 11:00 -17:00

ROOM: BOTTICELLI

Social Sciences and Humanities (SSH) are in general not known for Big Data challenges since the term often is reduced to processing large volumes of data. However, there are more aspects such as data to be worked on being highly distributed or having complex relationships with each other that need to be exploited etc. Partly, it has also to do with the fact that new trends in SSH are completely ignored such as studying the human brain based on brain imaging methodologies and genetic insights, the human language capacity and society based on massive crowdsourcing, the dynamics of diversities in cultures and languages based on extensive observations and phylogenetic methods, understanding the challenges of aging societies by collecting a wide range of markers including physiological ones, and many more. These questions address, amongst others, the interest to better understand the principles of maintaining stable minds and societies - thus grand challenges in the area of SSH.

The workshop will see contributions from experts addressing these new challenges in SSH that combine large and complex data with new computational challenges. One aspect that is similar to all these initiatives is that the investigators often do not have the facilities to develop large software packages, to manage appropriate storage and computer systems and to centralize the required facilities. We want to see what kind of directions are currently being worked on, what their challenges are and what kind of infrastructures are required to allow SSH researchers to carry out this kind of new research. The workshop targets experts from humanities departments who tackle data issues of a certain scale and complexity that clearly crosses normal situations and can describe the challenges and the opportunities. It also targets technologists that have an interest in helping to find methodological and technological solutions for the described challenges.

AGENDA:

11:30 Introduction
11:35 Peter Doorn - How to tackle the challenge of Long-term Access to Big Data in the Humanities and Social Sciences?
12:15 Nicola Masini - Satellite digital data for Cultural heritage: new strategies to share and extract information. The Virtual laboratory Italy-China
12:45 Riccardo Pozzo - From Data Science to Data Humanities
13:15 Lunch Break
14:15 Nanna Floor Clausen - Dealing with Danish Cencuses - problems and opportunities
14:45 Binyam Gebre - Massive Crowdsourcing: changing humanities.
15:15 Luca Pezzati - Digitizing Cultural Heritage
15:45 William C. Block - Freedom on the move
16:15 Coffee Break
16:30 DH and H2020 - General Discussion
17:00 End

CONTACT:

- Peter Wittenburg, MPI for Psycholinguistics, peter.wittenburg@mpi.nl
- Sebastian Drude, MPI for Psycholinguistics, sebastian.drude@mpi.nl
- Steven Krauwer, U Utrecht
- Erhard Hinrichs, U Tübingen
- Tobias Blanke, King's College, tobias.blanke@kcl.ac.uk
- Heike Neuroth, SUB Göttingen, neuroth@sub.uni-goettingen.de
- Hans Jørgen Marker, SND Gothenburg
- Jürgen Quandt, Gesis Cologne



Tuesday 29th October 2013

EUDAT Conference Plenary & Parallel Sessions



EUDAT 2ND CONFERENCE PROGRAMME OVERVIEW – TUESDAY 29TH OCTOBER 2013

RAFFAELLO + TINTORETTO + BOTTICELLI				
09:30-11:00	Plenary I: Addressing the new data challenges – the case for cross-disciplinary science and services. Chair: Damien Lecarpentier, EUDAT Project Manager, CSC-IT Center for Science			
09:30-10:15	Brain disease and big data - a way towards personalised medicine - Richard Frackowiak , Director, Department of Clinical Neuroscience, Head of Service of Neurology, CHUV University Hospital, Lausanne			
10:15-10:35	Data Infrastructures in Horizon 2020 - Kostas Glinos , European Commission, Head e Infrastructure unit - Directorate General for Communications Networks, Content and Technology			
10:35-11:00	EUDAT: a cross-disciplinary data infrastructure in Horizon 2020 - Kimmo Koski , Managing Director CSC - IT Center for Science			
11:00-11:30	Coffee Break			
11:30-13:00	Plenary II: Life and Earth Sciences at a cross-road: community driven flagship initiatives in the EU and USA. Chair: Alberto Michelini, Director of the National Earthquake Center, Istituto Nazionale di Geofisica e Vulcanologia (INGV)			
11:30-12:00	DataONE: Data Observation Network for Earth - William (Bill) Michener , Professor and Director of e-Science Initiatives for University Libraries, University of New Mexico & DataONE Principal Investigator			
12:00-12:30	Annotating the Human Genome - Ewan Birney , Associate Director of the EMBL-European Bioinformatics Institute (EMBL-EBI)			
12:30-13:00	Helix-Nebula: The Science Cloud - Maryline Lengert , Senior Advisor, ESRIN, ESA & Helix Nebula			
13:00-14:00	Networking Lunch			
14:00-18:30	Parallel Tracks			
	RAFFAELLO + TINTORETTO + BOTTICELLI	PUCCINI	TIZIANO	VERDI + MASCAGNI
14:00-16:00	Track 1: EUDAT Services - Chair: <i>Johannes Reetz</i>	Track 2: Interoperabilities Chair: <i>Morris Riedel</i>	Track 3: Policy and Sustainability issues Chair: <i>Damien Lecarpentier</i>	Track 4: New Services Chair: <i>Herbert Schentz</i>
	Safe Replication and Data Staging	Federated AAI	Access and Re-Use Policies	Semantic Annotation
16:00-16:30	Coffee Break			
16:30-18:30	Track 1: EUDAT Services - Chair: <i>Mark van de Sanden</i>	Track 2: Interoperabilities Chair: <i>Raphael Ritz</i>	Track 3: Policy and Sustainability issues Chair: <i>Rob Baxter</i>	Track 4: New Services Chair: <i>Alberto Michelini</i>
	Simple Store	Public-private Partnerships	Data Management Plans and Certification	Dynamic and Real-Time Data
20:00-22:30	Conference Dinner at Ristorante da Meo Patacca (Transport departs NH Vittorio Veneto @ 19:30 sharp)			

PLENARY SESSION I - TUESDAY 29TH OCTOBER

PLENARY SESSION I - ADDRESSING THE NEW DATA CHALLENGES – THE CASE FOR CROSS-DISCIPLINARY SCIENCE AND SERVICES

CHAIR: DAMIEN LECARPENTIER, EUDAT PROJECT MANAGER, CSC-IT CENTER FOR SCIENCE, EUDAT PROJECT MANAGER, CSC-IT CENTER FOR SCIENCE


DATE & TIME: TUESDAY 29TH OCT – 09:30 -11:00



PLENARY ROOM:RAFFAELLO + TINTORETTO + BOTTICELLI

OVERVIEW:

The Addressing new data challenges – the case for cross-disciplinary science and services – plenary session chaired by Damien Lecarpentier, EUDAT Project Manager, sets the scene of the conference and places a spotlight on the new data challenges affecting research communities and their funders with an introductory keynote on the Human Brain Project (HBP) and the data challenges this new FET Flagship project will face while achieving full simulation of the human brain. Richard Frackowiak will be followed by Kostas Glinos, EC DG Connect, Head of e-Infrastructure Unit who will give an overview of the European Commission’s strategy on data infrastructures in view of the imminent launch of the EC framework programme Horizon 2020. Kimmo Koski will present EUDAT and its work and achievements in laying the foundation of a pan-European Collaborative Data Infrastructure. The participants will learn more about its various components in the successive parallel sessions.

AGENDA:

09:30-10:15	Brain disease and big data - a way towards personalised medicine, Richard Frackowiak, Director, Department of Clinical Neuroscience, Head of Service of Neurology, CHUV University Hospital, Lausanne
	Prof. Richard Frackowiak’s research interest has been the functional and structural architecture of the human brain in health and disease. He has pioneered the development and introduction of positron emission tomography and magnetic resonance imaging and prosecuted a research programme dedicated to understanding the organisation of human brain functions, but his focus has been on plasticity and mechanisms for functional recuperation after brain injury and the patho-physiology of cerebral neurodegenerations. He became interested in the use of MR-based morphometry especially in the study of genetic influences on brain disease and in a search for biomarkers and endophenotypes of neurodegenerative disorders. Most recently he introduced computerised image classification for diagnosis and treatment monitoring into clinical science
10:15-10:35	Data Infrastructures in Horizon 2020, Kostas Glinos, European Commission, Head e Infrastructure unit - Directorate General for Communications Networks, Content and Technology

	<p>Kostas Glinos has been with the European Commission since 1992. He leads the e Infrastructure unit of the Directorate General for Communications Networks, Content and Technology since 1 January 2009. From 2003 to 2008 he was Head of the Embedded Systems and Control unit and interim Executive Director of the ARTEMIS Joint Undertaking. Previously he was deputy head of Future and Emerging Technologies. Before joining the Commission Kostas worked with multinational companies and research institutes in the U.S., Greece and Belgium. He holds a diploma in Chemical Engineering from the University of Thessaloniki, a PhD from the University of Massachusetts and a MBA in investment management from Drexel University.</p>
<p>10:35-11:00</p>	<p>EUDAT: a cross-disciplinary data infrastructure in Horizon 2020, Kimmo Koski, Managing Director CSC - IT Centre for Science</p>
	<p>Kimmo Koski started in his current position as Managing Director of the Finnish IT center for science, CSC, in August 2004. Prior to his present position, Koski spent 4.5 years in Nokia Research Center and Nokia Technology Platform. Earlier work experience includes 10 years at CSC in various positions and a one-year visiting period in CERN in Switzerland. Koski received his doctorate from Helsinki University of Technology in January 1996. His dissertation was on Metacomputing Technology. During the recent years Koski has been active in building European HPC and data infrastructure through involvement in major EU initiatives, such as PRACE and EUDAT. Koski has acted as the vice chair of PRACE council in 2010-2013 and the coordinator of EUDAT since October 2011.</p>



PLENARY SESSION II - TUESDAY 29TH OCTOBER

PLENARY SESSION II - LIFE AND EARTH SCIENCES AT A CROSS-ROAD: COMMUNITY DRIVEN FLAGSHIP INITIATIVES IN THE EU AND USA

CHAIR: ALBERTO MICHELINI, DIRECTOR OF THE NATIONAL EARTHQUAKE CENTER, ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA (INGV)


DATE & TIME: TUESDAY 29TH OCT – 11:30 -13:00



PLENARY ROOM:RAFFAELLO + TINTORETTO + BOTTICELLI

OVERVIEW:

The Life and Earth Sciences at a cross-road: community driven flagship initiatives in the EU and USA session chaired by Alberto Micheleni INGV will showcase major initiatives to build community driven infrastructures in the area of biomedical and earth science. Bill Michener will present DataONE, a very successful US initiative looking at making access to environmental data easier for both the scientist and the public. He will be followed by Ewan Birney from EBI / ELIXIR who will address the issue of how to safeguard the results of research in life science and highlight recent developments driven by the European Bioinformatics Institute and in ELIXIR. Maryline Lengert will provide insights on the new initiatives at the European Space Agency with a particular focus on Helix-Nebula: The Science Cloud and its synergies with EUDAT. Speakers will outline how EUDAT can & will contribute to supporting the emerging infrastructures in these different areas.

AGENDA:

<p>11:30-12:00</p>	<p>DataONE: Data Observation Network for Earth, William (Bill) Michener, Professor and Director of e-Science Initiatives for University Libraries, University of New Mexico & DataONE Principal Investigator</p>
	<p>Bill Michener is Professor and Director of e-Science Initiatives at the University of New Mexico’s University Libraries. He serves as Project Director for two large National Science Foundation supported projects: (1) Data Observation Network for Earth (DataONE)—a large DataNet project that supports cyberinfrastructure development and community engagement for the biological, environmental, and Earth sciences; and (2) the New Mexico Experimental Program to Stimulate Competitive Research. He is actively involved in research related to creating information technologies supporting data-intensive science, development of federated data systems, and community engagement and education. He has a PhD in Biological Oceanography from the University of South Carolina and has published extensively in marine science, as well as the ecological and information sciences.</p> <p>Bill has authored five books and more than 100 journal articles and book chapters. He has a strong background and training in organizational sustainability and governance, project management, and meeting facilitation. Presently, he serves on the Board of Directors (or Administrative Board) for Dryad, Inc., the Organization for Tropical Studies, and the Cornell Lab of Ornithology, as well as the Governance Committee for an emerging international organization that seeks to nurture a network of organizations that are involved in Public Participation in Scientific Research (i.e., citizen science). He serves as editor of the Ecological Society of America’s Ecological Archives, Associate Editor for Ecological Informatics, and as a member of the Ecology Editorial Board, and was recently appointed to the Technical Advisory Board for the Research Data Alliance</p>

12:00-12:30	<p>Annotating the human genome, Ewan Birney, Associate Director of the EMBL-European Bioinformatics Institute (EMBL-EBI)</p>
	<p>Dr Ewan Birney is Associate Director of the EMBL-European Bioinformatics Institute (EMBL-EBI). He is one of the founders of the Ensembl genome browser and other databases, and has played a key role in many large-scale genomics projects, notably the sequencing of the Human Genome in 2000 and the analysis of genome function in the ENCODE project. He has been Lead Analysis Coordinator for ENCODE since 2007; he also coordinated data analysis in the "1% Pilot". Dr Birney has played a vital role in annotating the genome sequences of the human, mouse, chicken and several other organisms; this work has had a profound impact on our understanding of genomic biology. His research group currently focuses on genomic algorithms and inter-individual differences in human and other species.</p> <p>As Associate Director of EMBL-EBI, Dr Birney shares strategic oversight of EBI services with Rolf Apweiler (co-Associate Director). EMBL-EBI hosts some of the world's most important collections of biological data, including DNA sequences, the genomes of animals and plants , three-dimensional molecular structures, data from gene expression experiments, Protein mass spectroscopy, small molecules of biological interest, their drug/protein target interactions and pathways. Many of these resources are developed in close collaboration with our partners at Swiss Institute of Bioinformatics, Sanger Institute, and Ontario Institute for Cancer Research. The EBI leads the European Elixir project which aims to broaden and deepen these relationships in Europe to provide an stable bioinformatics infrastructure that works well in the national, transnational and international contexts.</p> <p>As well as these strategic aspects, Ewan Birney still does research, working on aspects such as DNA Compression, Functional genomics analysis (eg, a recent paper with the Furlong lab) and using inter individual differences to understand basic biology (eg, a paper of CTCF binding in two families)</p>
12:30-13:00	<p>Helix-Nebula: The Science Cloud, Maryline Lengert, Senior Advisor, ESRIN, ESA & Helix Nebula</p>
	<p>Maryline Lengert, Senior Advisor in the IT department of ESA, started the Cloud Computing Strategic Plan initiative end of 2009 - now called Helix Nebula, the Science Cloud - to catalyse the creation of a European owned public cloud, serving initially the Science Research Area. Maryline has 20+ years of experience in the IT domain within an international environment. Her previous positions include Technical Manager and Operations Director for Hospital Integrated Information System in S.A.I.C. France, Quality Manager and Head of the Account Management and Requirement Analysis Division in the ESA IT Department. Maryline holds a PhD in Physics and a Master in International and European Studies.</p>



PARALLEL SESSION: TRACK 1 - EUDAT SERVICES -1.1 SAFE REPLICATION AND DATA STAGING

CHAIR: JOHANNES REETZ, RZG / MPG, GERMANY

DATE & TIME: TUESDAY 29TH OCTOBER – 14:00 - 16:00

PLENARY ROOM:RAFFAELLO + TINTORETTO + BOTTICELLI


OVERVIEW:

Safe Replication (SR) and **Data Staging (DS)** are EUDAT services for moving data between sites and storage systems for two kinds of purposes. The **purpose of SR** is to keep the data from a repository safe by replicating it across different geographical and administrative zones according to a set of well-defined policies. It is also a way to store larger volumes of data permanently at those sites which are providing powerful on-demand data analysis facilities. In particular, SR operates on the domain of registered data where data objects are referable via persistent identifiers (PIDs). SR is more than just copying data because the PIDs must be carefully managed when data objects are moved or replicated.

The general **purpose of DS** is to move data between storage systems, specifically, DS transfers data from the domain of registered data into a temporary storage space, and vice versa, it moves data from any storage space into the domain of registered data, typically a repository.

The session will introduce the problem space of SR and DS, presents the achievements that have been made during the last year for enabling communities to make use of the SR service as well as DS, demonstrates a few use cases, outlines the commonalities and differences between the policies for SR, presents new developments towards a common service layer interface and a data policy management framework.

AGENDA:

14:00 - 14:10	B2SAFE and B2STAGE: Two Core Services of the EUDAT CDI, Johannes Reetz, RZG / MPG, Germany
	Johannes Reetz is Operations Manager of EUDAT and project manager at the Garching Computing Centre of the Max Planck Society (RZG). He received a diploma in physics in 1991 and a PhD in natural sciences from the University of Munich in 1999. His scientific background is observational and theoretical astrophysics in the domain of cool stellar atmospheres and the chemical evolution of the Galaxy. After a working stay at the ESO Data Management Division in 1999, he joined the XDV team at the Max Planck Institute for Plasma Physics developing data acquisition software for the W7-X nuclear fusion experiment. In the European HPC project DEISA, he worked on grid middleware, the development of the accounting facilities, and was involved in DEISA operations and task leader for data-related services.
14:10 - 14:20	B2SAFE adoption in the EPOS community, Claudio Cacciari, CINECA, Italy
5min Q&A	

14:25 - 14:35 :	Utilisation of B2SAFE by the MPI-TLA CLARIN Center, Willem Elbers, Max-Planck-Institute for Psycholinguistics, The Netherlands
	Willem Elbers has been working as a software developer at The Language Archive, a unit of the Max Planck Institute for Psycholinguistics, since 2009. Prior to this position he studied computer science with a master in artificial intelligence at the Radboud University in Nijmegen. His specific areas of interest include data grid middleware, single sign on solutions, authentication and authorization infrastructures and the management and access of digital collections of linguistic data.
5min Q&A	
14:40 - 14:50	CLARIN-CUNI, Pavel Stranak, Charles University, Czech Republic
5min Q&A	
14:55 - 15:10	VPH and the Biomedical Scientific Case for EUDAT, Peter Coveney, University College London, UK
15:10 - 15:30	Community integrated B2STAGE tools, Stefan Zasada, University College London, UK and Giuseppe Fiameni, CINECA, Italy
	Giuseppe Fiameni holds a degree in Computer Science from the University of Bologna and joined the SuperComputing Department of CINECA (largest Italian supercomputing centre) as technology consultant on 2004. Over the years he has been contributing to many European (PRACE, DEISA, EMI, etc.) and National projects maturing a strong experience in High Performance Computing infrastructure and large data management and analysis systems. He has extensive experience in the field of computational sciences, parallel architectures, parallel programming models, scaling applications and system performance evaluation. G.F. is currently leading the "Middleware for HPC Services" group of the SuperComputing Applications and Innovation department which is responsible for the implementation of data and computational services and contribute to the evolution of the CINECA computational and data infrastructure. He is actively contributing to the EUDAT and the Human Brain project, member of the Research Data Alliance and of the ICT board of the EPOS project (www.epos-eu.org).
5min Q&A	
15:35 - 14:45	B2SAFE Policies, Willem Elbers, Max-Planck-Institute for Psycholinguistics, The Netherlands
15:45 - 15:55	B2SAFE - Data Policy Manager Service Case, Maria Francesca Iozzi, SIGMA/University of Oslo, Norway
	Maria Francesca Iozzi is presently employed as senior software engineer at the University Center of Information Technology, University of Oslo. She majored (with honors) at the University of Pisa in theoretical chemistry and obtained the PhD Degree in Physical Chemistry at the University of Napoli, Italy, specializing in molecular modeling and theoretical methods applied to material science. After many year of academic research she entered the HPC world as support for end users dealing with chemistry and material science software. She has been part of PRACE (Partnership for Advanced Computing in Europe) where she acted as coordinator of the effort of different European groups to enable common software to petascale computing. More recently she joined the EUDAT (European Data Infrastructure) as leader of the task force developing the Data Policy Manager Service.

CHAIR: MARK VAN DE SANDEN, SURFSARA


DATE & TIME: TUESDAY 29TH OCTOBER – 16:30 - 18:30

PLENARY ROOM:RAFFAELLO + TINTORETTO + BOTTICELLI

OVERVIEW:

Many researchers are collaborating across institutional boundaries and are facing the problem of finding an easy way of storing and sharing their data. But even single researchers should want to safeguard their own results and data in a safe storage place that will be kept open and accessible for other researchers in the future. They often have large numbers of small files, for example, files containing derived data in the form of spread sheets or analysis results. Although the information in these small files is important, they usually do not belong to large research collaborations with well-defined data storage plans. These kind of files are known as the “long tail data”, which are often stored locally on laptops and departmental storage devices with the risk of losing valuable scientific data, either because other researchers do not have easy access to the data or because such storage systems are often not adequately secure. To solve this problem EUDAT is working on a service called Simple Store which combines a shared storage space with the possibility to add metadata and to assign persistent identifiers. This session will present the current work and plans on the Simple Store service, related work from other projects and there will be time for discussion on the subject of “long tail data”. Options to provide cloud storage as a temporary low-barrier storage solution will be discussed.

AGENDA:

16:30 - 18:30	Welcome & Introduction, Mark Van De Sanden, SURFsara
	<p>Mark van de Sanden has a Bachelor degree in computer engineering from the technical college's Hertogenbosch and started work at the National Aerospace Laboratory (NLR) as system administrator. In 1997 he joined SURFsara as an UNIX system administrator of supercomputing environments. In 2002 he started working on the SURFsara mass storage infrastructures and is currently team leader of the Data Services group. He is involved in large scale data projects like the Large Hadron Collider (LHC) as part of the WLCG NL-T1 and Long Term Archiving (LTA) facilities of the LOFAR low frequency telescope. In EUDAT he is coordinating the service building work package.</p>
	Meeting DRIHM Citizen Scientist needs with SimpleStore, Alberto Parodi – CIMA Research Foundation



Expert in atmospheric modelling and statistical analysis of extreme events, in the development of simplified models of dry and moist convection and the study of the main sources of uncertainty in the high resolution numerical modelling of deep moist convective processes. Awarded with a CNR-MIT grant in 2002 in the framework of the bilateral USA-Italy investigations on climate change and hydrogeological disasters. Since 2003 has developed teaching activities at the University of Genova in the following fields: Hydraulics, Fluid Mechanics, Dynamics of Atmosphere and Computational methods in Environmental Engineering. Coordinator of FP7 project DRIHMS (Distributed Research Infrastructure for Hydro-Meteorology Study, www.drihms.eu, 2009-2011) and DRIHM (Distributed Research Infrastructure for Hydro-Meteorology Study, www.drihm.eu, 2011-2015). Coordinator of FP7 project DRIHMS (Distributed Research Infrastructure for Hydro-Meteorology Study, www.drihms.eu, 2009-2011), DRIHM (Distributed Research Infrastructure for Hydro-Meteorology Study, www.drihm.eu, 2011-2015) and DRIHM2US (Distributed Research Infrastructure for Hydro-Meteorology Study to United States of America, www.drihm2us.eu, 2012-2014). Antonio Parodi is author and co-author of 30 papers published in international peer-reviewed and referred journals.



CHAIR: MORRIS RIEDEL, JUELICH SUPERCOMPUTING CENTRE IN GERMANY, UNIVERSITY OF ICELAND

DATE & TIME: TUESDAY 29TH OCTOBER – 14:00 - 16:00

ROOM: PUCCINI

AGENDA:

14:00 - 14:10	Welcome & Introduction, Morris Riedel, Juelich Supercomputing Centre in Germany, University of Iceland
	<p>Dr. - Ing. Morris Riedel is an Adjunct Associate Professor at the School of Engineering and Natural Sciences of the University of Iceland. He received his PhD from the Karlsruhe Institute of Technology (KIT) and started the work in parallel and distributed systems in the field of scientific visualization and computational steering of e-science applications on large-scale HPC resources. He is also the deputy division leader of the division “Federated Systems and Data” of the Juelich Supercomputing Centre in Germany. At this institute, he is also the head of a specific scientific research group focussed on “Interoperability and Applications”. Lectures given in universities such as the University of Applied Sciences of Cologne and University of Technology Aachen (RWTH Aachen) include ‘Handling of large datasets’ and ‘Scientific and Grid computing’. His current research focusses on 'high productivity processing of big data' in the context of scientific computing applications.</p>
14:10 – 14:30	Federating User Identities in the EUDAT File manager Today and Tomorrow, Rion Dooley, Texas Advanced Computing Center
	<p>Rion Dooley is a research associate at the Texas Advanced Computing where he leads the Web and Cloud Services group. He earned his Ph.D. in CS from LSU in 2004 with the support of a Board of Regents Fellowship. He worked at LSU's Center of Computation and Technology for 2 years before moving to UT. Past projects include the development of science gateways such as GridChem, the TeraGrid mobile user portal, and the XSEDE user portal. Rion is currently the lead architect of the Agave API, and co-PI of the Distributed Web Security for Science Gateways project. He is also a senior participant in the Science Gateway Institute planning project. Rion's primary research interests include distributed systems, cloud infrastructure, and data management.</p>
14:30 – 14:35	Q&A

14:35 - 14:55	eduGAIN: making the infrastructure work for users Nicole Harris, Project Development Officer, TERENA
	<p>Nicole Harris joined TERENA in March 2013 as one of the Project Development Officers to support TERENA's Task Forces and contribute to technical projects. She works primarily in the security and middleware areas. Nicole is responsible for TF-CSIRT, TERENA's forum for Computer Security Incident Response Teams, and its associated service Trusted Introducer. Nicole has been involved in REFEDS as joint coordinator with Licia Florio since 2010. REFEDS (Research and Education FEDerations) is the voice that articulates the mutual needs of research and education identity federations worldwide, and provides a range of working groups, advisories and services to enhance interaction and interoperability between identity federations, including working with services such as eduGAIN on interederation. She will also be providing support to the TERENA Certificate Service, which is a natural complement to both REFEDS and TF-CSIRT.</p>
14:55 – 15:00	Q&A
15:00 - 15:20	Bottom-up community solution with EDUGain and REMS, Mikael Linden, CSC – IT Centre for Science
15:20 – 15:25	Q&A
15:25 - 15:45	Federated Identity Management for Research Collaborations, Bob Jones - Head of CERN openlab
	<p>Bob Jones is the Head of CERN openlab, as well as a member of the IT department head office with responsibilities in EC co-funded projects. Following a B.Sc. (Hons) in Computer Science from Staffordshire University, Bob joined CERN in 1986 as a software developer with the information technology department providing support for the physics experiments running on the Large Electron Positron (LEP) particle accelerator. He completed his PhD thesis in Computer Science at Sunderland University while working at CERN. He has been involved in several research projects for the Large Hadron Collider (LHC) accelerator and has contributed to online software system for the ATLAS experiment (http://www.atlas.ch/) at the LHC. Before joining openlab, Bob was a leader of the European Commission funded series of EGEE (Enabling Grids for ESciencE http://www.eu-egee.org) projects that helped develop the world-wide grid system that is used to process LHC data. Bob is the coordinator for the Helix Nebula EC project (http://www.helix-nebula.eu/)</p>
15:45 – 15:50	Q&A
15:50 – 16:00	Session Conclusions, Morris Riedel, Juelich Supercomputing Centre in Germany, University of Iceland



CHAIR: RAPHAEL RITZ, HEAD OF DATA SCIENCE GROUP AT RZG, MAX PLANCK SOCIETY, GERMANY**DATE & TIME: TUESDAY 29TH OCTOBER – 16:30 – 18:30****ROOM: PUCCINI****OVERVIEW:**

EUDAT recently showcased at the ISC'13 conference in Leipzig, Germany. While the conference primarily focused on supercomputing, visitors at the booth demonstrated significant interest in EUDAT services, in particular in relation to EUDAT's open access Collaborative Data Infrastructure (CDI). We were asked basic questions such as "What is the added value of EUDAT versus services like Google Drive and Dropbox?", "Is EUDAT a business, and, if so, what is the business model behind it?", and even "Does EUDAT plan to deliver any analytics services to explore and exploit the data it is hosting?".

These questions naturally lead to debates on the sustainability of e-infrastructures and the largely unexplored role of industry, which is a common challenge in this area. Indeed, given the complex but federating nature of EUDAT and its potential impact on society, it is crucial for EUDAT to involve industrial partners at an early enough stage in the project development. This is vital for fostering the adoption and uptake of the EUDAT services, and also for devising an appropriate economical model for sustaining the e-infrastructure in the longer term. EUDAT was born in an ecosystem of infrastructure operators and service providers at regional, European and international levels. While developing and delivering an innovative product/service, EUDAT is also integrating its services with the portfolios of other existing e-infrastructures so as to offer custom-tailored services to targeted user communities. It is therefore important that e-infrastructure and service providers understand, and ultimately contribute, to the sustainable strategy of EUDAT. To bring this about, EUDAT has started to interact with actors from different sectors, including service providers in Cloud and HPC computing, system integrators and independent software vendors in Big Data and data analytics.




With the objective of making further progress on these key issues, EUDAT has organized two industry sessions at the 2nd annual conference.




- The first industry session took place on October 28th from 9am to 1pm and included contributions from major actors in the Cloud computing and Big Data areas, collocating with the CUMULONIMBO project final workshop.
- During this session, the hot topic of Public-Private Partnerships will be addressed with input from the pioneering HELIX NEBULA initiative.

This session addresses a range of important industry-related issues, from simple questions like "What can EUDAT offer in terms of products and services to industry?" to more elaborate considerations such as "What role should industry take in supporting/benefitting from EUDAT?", "What commitments will various industries make?", and "Will EUDAT consider the secondary use of data as a possible business model?". These topics will be discussed openly by a panel of industrial representatives, together with the audience.

With recent international discussions on data protection and liabilities, EUDAT clearly has a role to play as the European Collaborative Data Infrastructure, supporting not only European researchers and citizens, but also the growing Big Data industry. Thus, if you want to know more and contribute to this ongoing debate, come and attend the industry sessions

AGENDA:

16:30 – 18:30	(1) Public-Private Partnerships & e-Infrastructures - Chairman Bob Jones / CERN
	EUDAT - Morris Riedel, Juelich Supercomputing Centre in Germany, University of Iceland
	<p>Dr. - Ing. Morris Riedel is an Adjunct Associate Professor at the School of Engineering and Natural Sciences of the University of Iceland. He received his PhD from the Karlsruhe Institute of Technology (KIT) and started the work in parallel and distributed systems in the field of scientific visualization and computational steering of e-science applications on large-scale HPC resources. He is also the deputy division leader of the division “Federated Systems and Data” of the Juelich Supercomputing Centre in Germany. At this institute, he is also the head of a specific scientific research group focussed on “Interoperability and Applications”. Lectures given in universities such as the University of Applied Sciences of Cologne and University of Technology Aachen (RWTH Aachen) include ‘Handling of large datasets’ and ‘Scientific and Grid computing’. His current research focusses on 'high productivity processing of big data' in the context of scientific computing applications.</p>
	HELIX NEBULA - Bob Jones, CERN
	<p>Bob Jones is the Head of CERN openlab, as well as a member of the IT department head office with responsibilities in EC co-funded projects. Following a B.Sc. (Hons) in Computer Science from Staffordshire University, Bob joined CERN in 1986 as a software developer with the information technology department providing support for the physics experiments running on the Large Electron Positron (LEP) particle accelerator. He completed his PhD thesis in Computer Science at Sunderland University while working at CERN. He has been involved in several research projects for the Large Hadron Collider (LHC) accelerator and has contributed to online software system for the ATLAS experiment (http://www.atlas.ch/) at the LHC. Before joining openlab, Bob was a leader of the European Commission funded series of EGEE (Enabling Grids for ESciencE http://www.eu-egee.org) projects that helped develop the world-wide grid system that is used to process LHC data. Bob is the coordinator for the Helix Nebula EC project (http://www.helix-nebula.eu/)</p>
	(2) Open Discussion on PPP - Chairman Raphael Ritz / RZG
	<p>Raphael Ritz leads the Data Science and Services group within the computing center of the Max Planck Society. In addition, he represents the MPG in international contexts focusing on data handling.</p>

	PRESENTATIONS & DISCUSSIONS – PANELISTS:
	Monica Marinucci / ORACLE
	<p>Monica Marinucci is Director for Research at Oracle for Europe, Middle East and Africa.</p> <p>Monica joined Oracle Corporation in June 2004 after almost 10 years in the R&D/Academic world. She spent several years at CERN and previously she held different positions in other R&D organisations. She is a graduate of the La Sapienza University, Rome.</p>
	Maryline Lengert / ESA
	<p>Maryline Lengert, Senior Advisor in the IT department of ESA, started the Cloud Computing Strategic Plan initiative end of 2009 - now called Helix Nebula, the Science Cloud - to catalyse the creation of a European owned public cloud, serving initially the Science Research Area. Maryline has 20+ years of experience in the IT domain within an international environment. Her previous positions include Technical Manager and Operations Director for Hospital Integrated Information System in S.A.I.C. France, Quality Manager and Head of the Account Management and Requirement Analysis Division in the ESA IT Department. Maryline holds a PhD in Physics and a Master in International and European Studies.</p>
	Jurry De La Mar / T-Systems
	<p>Since 2004 Jurry de la Mar has started and lead various strategic projects with European Institutions for T-Systems in Europe. In 2008 he took over the sales and account management responsibility for European Institutions in Germany as well as the Galileo and Earth Observation programmes within T Systems' Public Sector. In 2011 he was one of the initiators to create Cloud Computing services for European Science which has now led to Helix Nebula – The Science Cloud consortium, in which T-Systems actively participates. And he is Member of the Supervisory Board of cesah GmbH, the Centre for Satellite Navigation in the State of Hesse, Germany.</p> <p>In his previous role Jurry de la Mar was located from 2000 till 2003 to Singapore and was Head of T-Systems and Deutsche Telekom's sales and operations in the region South East Asia and India where he joined the major accounts division in Frankfurt.</p>
	Pawel Kamocki / IDS Mannheim / University Paris V / University of Münster
	<p>Paweł Kamocki (Institut für Deutsche Sprache, Mannheim, CLARIN ERIC): Legal expert with background in language science; member of the CLARIN ERIC Legal Issues Committee, PhD candidate at Paris Descartes University and at Westfälische Wilhelms-Universität Münster (thesis subject: Legal issues in e-Research and e-Learning). He also works with the RDA/CODATA Interest Group on Legal Interoperability. He has presented on legal issues in many fora, including eIRG, CLARIN and EUDAT, and published in peer-reviewed journals.</p>

PARALLEL SESSION: TRACK 3 - POLICY AND SUSTAINABILITY ISSUES - 3.1 ACCESS AND RE-USE POLICIES

CHAIR: DAMIEN LECARPENTIER, EUDAT PROJECT MANAGER, CSC-IT



DATE & TIME: TUESDAY 29TH OCTOBER – 14:00 - 16:00


ROOM: TIZIANO

OVERVIEW:

Data access and reuse policies must be made clear to potential users of data infrastructures, both the “data providers” – who need to know how their data will be made accessible and who will be able to access it – and the “data customers” – who need to know about access or reuse restrictions before starting working with any of this data. This session will review the issues surrounding access and reuse from multiple angles and will report back on the outcomes of the recent EUDAT Working Group scoping workshop in this area.

AGENDA:

14:00 - 16:00	Report back from Barcelona workshop - Rob Baxter, EPCC/EUDAT
	Dr Rob Baxter graduated in 1989 with a BSc BSc (1st Class Hons) in Physics/Theoretical Physics from the University of St Andrews. He then spent a year in Cambridge, doing Part III of the Maths Tripos and falling off punts before coming to the University of Edinburgh in 1990 to join the Particle Physics Theory Group. He completed my PhD in lattice QCD in 1993 and subsequently joined EPCC. He currently co-manages the Software Development Group at EPCC, involved in commercial and scientific software development and working on projects such as SSI: the UK’s Software Sustainability Institute; ADMIRE: advanced data mining and Internet-scale data integration; Maxwell: how to build a supercomputer out of FPGAs; Condition-based Monitoring, with ITI Techmedia.
	An analysis of policies, norms and community agreements for research data - Simon Hodson - CODATA
	Simon Hodson is Executive Director of CODATA http://www.codata.org , an organisation whose mission is to strengthen international science for the benefit of society by promoting improved scientific and technical data management and use. He also sits on the Board of Directors of the Dryad data repository http://datadryad.org , a not-for-profit initiative to make the data underlying scientific publications discoverable, freely reusable, and citable. From 2009 to 2013, as Programme Manager, he led two successive phases of Jisc's innovative Managing Research Data programme http://researchdata.jiscinvolve.org/wp/ .

	Legal aspects of data sharing - Paweł Kamocki, Institut für Deutsche Sprache/Uni. Mannheim
	<p>Paweł Kamocki (Institut für Deutsche Sprache, Mannheim, CLARIN ERIC): Legal expert with background in language science; member of the CLARIN ERIC Legal Issues Committee, PhD candidate at Paris Descartes University and at Westfälische Wilhelms-Universität Münster (thesis subject: Legal issues in e-Research and e-Learning). He also works with the RDA/CODATA Interest Group on Legal Interoperability. He has presented on legal issues in many fora, including eIRG, CLARIN and EUDAT, and published in peer-reviewed journals.</p>
	Discussion



PARALLEL SESSION: TRACK 3 - POLICY AND SUSTAINABILITY ISSUES - 3.2 DATA MANAGEMENT PLANS AND CERTIFICATION

CHAIR: ROB BAXTER, EPCC - UNIVERSITY OF EDINBURGH



DATE & TIME: TUESDAY 29TH OCTOBER – 16:30 – 18:30



ROOM: TIZIANO

OVERVIEW:

With more and more research data born digital, and more and more emphasis being placed on open access to data (not least by the recent G8 Charter), effective data management planning has never been more relevant. This session will explore the ins and outs of this topic, including a look at some of the tools now available to support researchers in planning for the future of their data. It will also cover the issues of trust involved in handing over, or taking custody of, research data, and will include an interactive session on future directions for DMP in Europe.

AGENDA:

16:30 – 16:35	Welcome & Session Introduction, Rob Baxter, EPCC
	<p>Dr Rob Baxter graduated in 1989 with a BSc BSc (1st Class Hons) in Physics/Theoretical Physics from the University of St Andrews. He then spent a year in Cambridge, doing Part III of the Maths Tripos and falling off punts before coming to the University of Edinburgh in 1990 to join the Particle Physics Theory Group. He completed my PhD in lattice QCD in 1993 and subsequently joined EPCC. He currently co-manages the Software Development Group at EPCC, involved in commercial and scientific software development and working on projects such as SSI: the UK's Software Sustainability Institute; ADMIRE: advanced data mining and Internet-scale data integration; Maxwell: how to build a supercomputer out of FPGAs; Condition-based Monitoring, with ITI Techmedia.</p>
16:35 – 17:00	Motivation: increasing importance of DMP, Simon Lambert, STFC/APARSEN
	<p>Simon Lambert is a project manager in the Scientific Computing Department at the Science and Technology Facilities Council (STFC) in the UK. Since 2006 he has concentrated on projects in digital preservation and data management, particularly applied to scientific data such as that held by STFC's facilities. He is the joint coordinator of APARSEN, a large Network of Excellence partly funded by the European Commission's Seventh Framework Programme that brings together 31 diverse organisations across Europe to develop a common vision and a Virtual Centre of Excellence in digital preservation. He has also contributed to European studies of openness in scientific data (the ODE project) and open access publishing (SOAP).</p>

17:00 - 17:35	DMP Online tools – general, Sarah Jones, DCC/HATII
	<p>Sarah Jones works for the Digital Curation Centre, a UK national service to support universities with research data management. Her current work focuses on the DCC's institutional engagement programme. She is leading DCC support for several universities, helping them to scope, develop and implement Research Data Management services. Two key areas of Sarah's work are research data policy and data management planning. She has developed a number of guidance resources and training in these areas and is involved shaping DMPonline, the DCC's web-based tool to help researchers write Data Management Plans.</p>
17:35 – 18:00	DSA accreditation: principles, advantages, implementation, Marion Massol, CINES
	<p>Marion Massol is a research engineer who has been working on long-term preservation of digital data since March 2009. She is head of Digital Preservation Department at CINES (Centre Informatique National de l'Enseignement Supérieur), a French IT datacenter for the Higher Education and Research community. CINES is the national center for the preservation of French PhD theses, publishing (HAL), digitalized books and documents (Persée, university libraries, high schools, labs...), administrative data (labs, institutes, universities...). CINES is the French repository of EUDAT. In this project, Marion Massol is also the Quality Assurance task leader.</p> <p>Marion Massol is sharing her expertise in long-term preservation and quality assurance through national and international working groups (as pilot of the French working group called "PIN" on digital preservation and of the national initiative BSN6, member of the Data Seal of Approval board, APARSEN project, DARIAH project, national steering committee on the standard SEDA, expert in AFNOR and ISO standard working groups...) and trainings.</p>
18:00 – 18:25	DMP/tools in EUDAT – open discussion on what EUDAT should pursue?
18:25 – 18:30	Wrap Up & Conclusions



PARALLEL TRACKS – TRACK 4 – NEW SERVICES

During the preparation phase of the EUDAT proposal a number of scientific communities were engaged with and were involved in discussing the first EUDAT core service set. From this the well-known short list of 4 core services was extracted, the requirements were defined and basically much focus was put to get these 4 services running. Due to the ongoing interactions it became obvious that one of these services has 4 different flavors for different types of data providers. In parallel EUDAT defined two more services which are currently being worked on also a result of the EUDAT interactions (see the note on community engagement on the web: <http://www.eudat.eu/published-articles>).

Now it is time to begin further intensive discussions about what the candidates for future common services within a Collaborative Data Infrastructure could be. Already back in March 2013 at the EUDAT user forum in London some candidates were proposed. Currently EUDAT has launched a questionnaire covering many communities and individuals to understand their preferences and ideas. Furthermore, for some areas (dynamic/realtime data, workflow support, semantic services, access issues) EUDAT is organising special workshops in Mid-September to understand better what services could be useful as solutions for these areas. The EUDAT 2nd conference will be a further step, the presentation of an evaluation of all suggestions received to date and opportunity to openly discuss the outcome with all community experts and individuals that are interested. This will help EUDAT to find proper service cases for the coming phases that are based on the opinions of multiple communities together with the 5 core ones.

PARALLEL SESSION: TRACK 4 – NEW SERVICES - 4.1 SEMANTIC ANNOTATION

CHAIR: HERBERT SCHENTZ, ECOSYSTEM RESEARCH & MONITORING, UMWELTBUNDESAMT GMBH

DATE & TIME: TUESDAY 29TH OCTOBER – 14:00 - 16:00

ROOM: VERDI + MASCAGNI

Semantic services are of common interest, although it is not a straightforward process to identify exactly which services in the semantic domain could be seen as common services shared by some or many communities. Currently EUDAT is discussing one concrete service which we call “Semantic Annotation”. Semantic annotation can be applied to derived and typical long tail data, rather than to regular raw data created by machines. A typical example of the use of semantic annotation is a scenario where data is produced by humans and will therefore contain errors. Consequently scientists will want to annotate the errors and create references to accepted ontologies. This paradigm is becoming important to an increasing number of disciplines. At a certain level of abstraction, semantic annotation can be seen as a common service that can be applied to processes of data enrichment in many scientific disciplines. Such an annotation module could be used as plug-in for EUDAT core services, and also as a plug-in for community services. EUDAT will approach this work on two strands: on the one hand we want to implement a semantic annotation service as soon as possible, and, on the other hand, we want to start an elaborative discussion on other possible common services.

AGENDA

Semantic Annotation Chair: Herbert Schentz, Ecosystem Research & Monitoring, Umweltbundesamt GmbH	
Morris Riedel: General Overview New Services	
	Former suggestions, current survey, etc.
	First short Q+A
Herbert Schentz: Workshop Results	
	Start of EUON - European Ontology Network
Michael Mirtl, Environment Agency Austria & David Vicente Barcelona Supercomputing Center	
	Semantic Annotation Work in EUDAT
Other contributions and General discussion on Semantics in EUDAT	

CHAIR: ALBERTO MICHELINI, DIRECTOR OF THE NATIONAL EARTHQUAKE CENTER, ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA (INGV)

DATE & TIME: TUESDAY 29TH OCTOBER – 16:30 – 18:30

ROOM: VERDI + MASCAGNI

OVERVIEW:

Some dynamic data is generated by sensors which produce data streams that may be temporarily incomplete (owing to latencies or temporary interruptions of the transmission lines between the field sensors and the data acquisition centres) and that may consequently fill up over time (automatically or after manual intervention). Dynamic data can also be generated by massive crowd sourcing where, for example, experimental collections of data can be filled up at random moments. The nature of dynamic data makes it difficult to handle for various reasons: a) establishing valid policies that guide early replication for data preservation and access optimization is not trivial, b) identifying versions of such data – thus making it possible to check their integrity – and referencing the versions is also a challenging task, and c) performance issues are extremely important since all these activities must be performed fast enough to keep up with the incoming data stream. There is no doubt that both applications areas (namely data from sensors and crowdsourcing) are growing in their relevance for science, and that appropriate infrastructure support (by initiatives such as EUDAT) is vital to handle these challenges.

AGENDA

Dynamic Data Chair: Alberto Michelini, Director of the National Earthquake Center, Istituto Nazionale di Geofisica e Vulcanologia (INGV)
Alberto Michelini: Workshop Results
Herman Stehouwer: Crowd Sourcing Use case
Alberto Michelini: Sensor Data Use Case
Other contributions and General discussion on Dynamic Data in EUDAT
Alberto Michelini: Update New Services



Wednesday 30th October 2013

EUDAT Conference Plenary & Parallel Sessions

Associated Workshops & Meetings

EUDAT 2ND CONFERENCE PROGRAMME OVERVIEW – WEDNESDAY 30TH OCTOBER 2013

Parallel Tracks				
	RAFFAELLO + TINTORETTO + BOTTICELLI	PUCCINI	TIZIANO	VERDI + MASCAGNI
9:00-10:30	Track 1: EUDAT Services - <i>Chair: Daan Broeder</i>	Track 2: Interoperabilities <i>Chair: Ulrich Schwardmann</i>	Track 3: Policy and Sustainability issues <i>Chair: Marcin Ostasz & Ari Lukkarinen</i>	Track 4: New Services <i>Chair: Christian Pagé</i>
	Metadata	Identifiers	Cost and funding models	Workflows
10:30-11:00	Coffee Break			
RAFFAELLO + TINTORETTO + BOTTICELLI				
11:00-13:00	Plenary III: Towards Global Data Infrastructure Components. Chair: Leif Laaksonen, Collaboration Director, CSC-IT & RDA Europe Coordinator			
11:00-11:20	The Research Data Alliance: Status Update - John Wood , RDA Council Chair & Secretary General Association Commonwealth Universities (ACU)			
11:20-11:30	Data Foundation and Terminology – Raphael Ritz , Head of Data Science Group at RZG, Max Planck Society, Germany			
11:30-11:40	Metadata – William (Bill) Michener , Professor and Director of e-Science Initiatives for University Libraries, University of New Mexico & DataONE Principal Investigator			
11:40-11:50	PID Information Types – Tobias Weigel , German Climate Computing Center (DKRZ) / University of Hamburg			
11:50-12:00	Data Type Registry – Daan Broeder , Max Planck Institute For Psycholinguistics, Nijmegen			
12:00-12:10	RDA Working Group Practical Policy – Rainer Stotzka , Karlsruhe Institute of Technology (KIT), Institute for Data Processing & Electronics, Software Methods Group Head			
12:10-13:00	Panel discussion			
RAFFAELLO + TINTORETTO + BOTTICELLI				
13:00-13:30	Wrap-Up and Conclusions. Chair: Damien Lecarpentier, EUDAT Project Manager, CSC-IT Center for Science			
13:30-14:30	Networking Lunch			
	PUCCINI	TIZIANO	VERDI	
14:30-16:00	EPIC User Meeting (Closed)	PID Information Types WG Meeting	DIGITAL PRESERVATION OF CULTURAL DATA workshop	
16:00-16:30	Coffee Break			
16:30-18:30	EPIC User Meeting (Closed)	PID Information Types WG Meeting	DIGITAL PRESERVATION OF CULTURAL DATA workshop	

CHAIR: DAAN BROEDER, THE LANGUAGE ARCHIVE – MPI FOR PSYCHOLINGUISTICS

DATE & TIME: WEDNESDAY 30TH OCTOBER – 09:00 - 10:30


PLENARY ROOM:RAFFAELLO + TINTORETTO + BOTTICELLI

OVERVIEW:

The data subject is hot; it is considered the new oil of the digital area. Data is produced at staggering rates. The main questions are “Where store it?”, “How to find it?” and “How to make the most of it?”. In this context, metadata is the key to improving and ensuring the quality of data for current and future usage, and metadata catalogues are the place to find the interesting relevant and most valuable data sets. EUDAT is developing a joint metadata catalogue that combines the metadata, gathered from various repositories and sources, and hence bridges key information from research across science domains. This session will present current work and plans on the joint metadata service, the need for harvestable repositories and the challenge to understand and to bridge domain specific ontologies. It will also present related work to the dispersed metadata challenge. The session foresees ample discussion time to hear views and contributions from the participants.

AGENDA:

<p>09:00 - 09:30</p>	<p>Data interoperability in cultural heritage: the Europeana approach - Nuno Freire, The European Library, Europeana Foundation</p>
	<p>Nuno Freire is a Senior Researcher at The European Library. He holds a PhD in Informatics and Computer Engineering from the Instituto Superior Técnico of the Technical University of Lisbon. During his entire career he has been involved in research projects in the area of digital libraries. His areas of interest include information systems, information retrieval, information extraction, data quality, and knowledge representation, particularly in their application to digital libraries and bibliographic data.</p>
<p>09:30 – 10:00</p>	<p>Data Federation via Metadata – (Bill) Michener, University of New Mexico’s University Libraries</p>
	<p>Bill Michener is Professor and Director of e-Science Initiatives at the University of New Mexico’s University Libraries. He serves as Project Director for two large National Science Foundation supported projects: (1) Data Observation Network for Earth (DataONE)—a large DataNet project that supports cyberinfrastructure development and community engagement for the biological, environmental, and Earth sciences; and (2) the New Mexico Experimental Program to Stimulate Competitive Research. He is actively involved in research related to creating information technologies supporting data-intensive science, development of federated data systems, and community engagement and education. He has a PhD in Biological Oceanography from the University of South Carolina and has published extensively in marine science, as well as the ecological and information sciences.</p>

	<p>Bill has authored five books and more than 100 journal articles and book chapters. He has a strong background and training in organizational sustainability and governance, project management, and meeting facilitation. Presently, he serves on the Board of Directors (or Administrative Board) for Dryad, Inc., the Organization for Tropical Studies, and the Cornell Lab of Ornithology, as well as the Governance Committee for an emerging international organization that seeks to nurture a network of organizations that are involved in Public Participation in Scientific Research (i.e., citizen science). He serves as editor of the Ecological Society of America's Ecological Archives, Associate Editor for Ecological Informatics, and as a member of the Ecology Editorial Board, and was recently appointed to the Technical Advisory Board for the Research Data Alliance</p>
<p>10:00 – 10:30</p>	<p>B2find: The EUDAT Metadata Service - Daan Broeder, The Language Archive – MPI for Psycholinguistics</p>
	<p>Daan Broeder, for many years senior developer for archive and infrastructure solutions and now deputy director of TLA unit of the MPI for Psycholinguistics, is technologist (electronics and IT) and has a long record in leading development tasks in international projects. Currently he is a member of the executive board of the Dutch CLARIN project and participating several EU projects concerned with research infrastructure development as DASISH and EUDAT.</p> <p>In addition he is convener of new ISO standards in the linguistic domain (TC37/SC4: Component Metadata Infrastructure "CMD" and persistent identification "PISA").</p>






CHAIR: ULRICH SCHWARDMANN, GWDG, GERMANY

DATE & TIME: WEDNESDAY 30TH OCTOBER – 09:00 - 10:30

ROOM: PUCCINI

AGENDA:

09:00 - 09:05 :	Welcome & Session Introduction, Morris Riedel, JUELICH
	Dr. - Ing. Morris Riedel is an Adjunct Associate Professor at the School of Engineering and Natural Sciences of the University of Iceland. He received his PhD from the Karlsruhe Institute of Technology (KIT) and started the work in parallel and distributed systems in the field of scientific visualization and computational steering of e-science applications on large-scale HPC resources. He is also the deputy division leader of the division “Federated Systems and Data” of the Juelich Supercomputing Centre in Germany. At this institute, he is also the head of a specific scientific research group focused on “Interoperability and Applications”. Lectures given in universities such as the University of Applied Sciences of Cologne and University of Technology Aachen (RWTH Aachen) include ‘Handling of large datasets’ and ‘Scientific and Grid computing’. His current research focusses on 'high productivity processing of big data' in the context of scientific computing applications.
09:05 - 09:10	Q&A
09:10 - 09:30	On Working with EPIC Persistent Identifiers, Ulrich Schwardmann, GWDG
	Ulrich Schwardmann studied mathematics and informatics and made his PhD in mathematics in 1987. Afterwards he became employee at GWDG with main activities in parallel and scientific computing and optimization. In 2007 he started to set up a persistent identification service for the Max-Planck-Society at GWDG as a mayor building block for the data management activities of GWDG, that he is leading since then.
09:30 - 09:35	Q&A
09:35 - 09:45	DataCite – global infrastructure for making data citable based on the DOI system - Jan Brase DataCite
	Jan Brase has a degree in Mathematics and a PhD in computer science from the University of Hannover. He was head of the DOI (Digital Object Identifier) registration agency of the German National Library of Science and Technology and has been Executive Director of DataCite since its founding in 2009. He is vice president of the International Council of Scientific and Technical Information (ICSTI), co-chair of the International DOI Foundation (IDF) and a co- chair of the CODATA-ICSTI (Committee on Data for Science and Technology/International Council on Scientific and Technical Information) Task Group on Data Citation.

09:45 - 10:00	Handle System Evolution and Governance, Larry Lannom, Director of Information Services, Vice President at the Corporation for National Research Initiatives (CNRI)
	<p>Larry Lannom is Director of Information Services and Vice President at the Corporation for National Research Initiatives (CNRI), where he works with organizations in both the public and private sectors to develop experimental and pilot applications of advanced networking and information management technologies. His current work is focused on CNRI's Digital Object Architecture, which is based on the concept of the digital object, a uniform approach to representing digital information across computing and application environments, both now and into the future. He is responsible for the development and ongoing evolution of a series of infrastructure components needed to implement the architecture. This includes a high performance resolution system, known as the Handle System that maps identifiers into current state information about the digital objects being identified, repositories for storing digital objects and from which they may be accessed, and metadata registries for managing collections of digital objects across one or more repositories.</p>
10:00 - 10:20	Exploring Persistent Identifiers for Open Time Series, Robert Huber – MARUM
	<p>Dr. Robert Huber, Geologist and Information Specialist holding a PhD in Marine Geology. He worked several years as information system architect for the aerospace industry and the renewable energy industry. Since 2002 he is employed at the Centre for Marine Environmental Sciences (MARUM) at the University Bremen and responsible for projects in scientific data management and IT development especially in the fields of marine observatory networks at the PANGAEA working group. He was/is leading the data management work package for ESONET, FIXO3 and COOPEUS and was/is involved in the related EU projects EMSO, HYPOX, SIOS and ENVRI.</p>
10:20 - 10:25	Q&A
10:25 - 10:30	Session Conclusions, Morris Riedel, JUELICH



PARALLEL TRACKS – TRACK 3 - POLICY AND SUSTAINABILITY ISSUES

PARALLEL SESSION: TRACK 3 - POLICY AND SUSTAINABILITY ISSUES - 3.3 COST & FUNDING MODELS

CHAIR: MARCIN OSTASZ & ARI LUKKARINEN

DATE & TIME: WEDNESDAY 30TH OCTOBER – 09:00 - 10:30

ROOM: TIZIANO

OVERVIEW:

This workshop aims at providing an overview of the process of building a funding, cost and potentially a pricing model EUDAT’s services. We are also looking to collect feedback on those activities from EUDAT’s customers, communities and potential operators.


EUDAT is now in a position to define a process to obtain funding for its operation. Funds to operate EUDAT can be provided by the EC, Member States, regional funds or on a community level. It is also assumed that some of the contributions will be in-kind. Our work in this area aims at analysing different scenarios that will suit EUDAT and its service portfolio.



EUDAT will also need to look at the cost of operating its sites. Initially, this information will provided by each site based on their own models, with a few to apply a more consistent approach at a later stage. Two existing projects in the area of data cost models will be presented (DANS and 4C) with a view to relate them to the work of EUDAT.

The main questions to be tackled during the workshop are:

- What funding models can suit EUDAT, reflecting its diversity and federated structure? What models ensure long-term sustainability and inter-operability? **Who should pay for what, to whom, from what source and how?**
- What cost models could be applied on the site level?

AGENDA:

09:00 - 10:30	EUDAT Services and Current EUDAT Funding and Cost Model activity - Marcin Ostasz, EUDAT/BSC
	<p>Barcelona Supercomputing Centre - Marcin Ostasz graduated from the Technical University of Budapest at the Faculty of Electronics with an MSc degree and he also holds a Master of Business Administration (MBA) degree awarded by Oxford Brookes University in the UK. Marcin has over 13 years of combined experience gained at various technical, project management, operations management, business analysis and process improvement positions with organisations such as Nokia, American Power Conversion, Dell, GE and Barclays Bank. Marcin is currently working at Barcelona Supercomputing Centre as a business analyst. His tasks</p>

	include supporting projects and organisations such as PRACE, the European Technology Platform, EUDAT and Mont-Blanc. He specialises in managing industrial relations, road-mapping, workshop management and business analysis.
	National archive view - Peter Doorn, DANS
	Peter Doorn studied Human Geography at Utrecht University and received his PhD there. He taught Computing for Historians at Leiden University from 1985 to 1997. Later, he was the director of the Netherlands Historical Data Archive and a head of department at the Netherlands Institute for Scientific Information Services (NIWI). He acquired and directed a considerable number of externally funded digitisation projects and other projects in the field of humanities computing.
	Overview of Cost Models for Digital Curation - Kevin Ashley, C4 Project/DCC UK
	As Director, Kevin maps new territory as the DCC embarks on its third phase of evolution (2010 - 2013), where the accent is on enabling capacity and capability amongst the research community in matters of digital curation. Previously, as Head of Digital Archives at the University of London Computer Centre (1997 - 2010), he was responsible for a multi-disciplinary group that provided services related to the preservation and reusability of digital resources on behalf of other organisations, as well as programmes of research, development and training. His group operated NDAD (the National Digital Archive of Datasets) for The National Archives of the UK for over twelve years, capturing, preserving, describing and releasing government data as part of TNA's collections. As a past or present member of numerous advisory and steering groups, including JISC's Infrastructure and Resources Committee, JIIE, the Repositories and Preservation Advisory Group, the Advisory Council for ERPANET and the Archives Hub Steering Committee, Kevin has contributed widely to the research information community.
	Discussion



CHAIR: CHRISTIAN PAGÉ, CERFACS**DATE & TIME: WEDNESDAY 30TH OCTOBER – 09:00 - 10:30****ROOM: VERDI + MASCAGNI**

Well-described and documented scientific workflows that can be executed to achieve new results are becoming more and more important in all scientific disciplines to cope with the increasing amount of data in appropriate ways and to increase the reproducibility of scientific results. This is true both for raw data generated by sensors and software systems and processed in regular ways, and also for many areas of derived data - the long-tail data. As we move towards “data fabric solutions”, workflow support for manipulating data will be essential. Data infrastructure initiatives such as EUDAT and DataONE are already working on workflow systems and building up expertise, while large institutions such as LANL and SDSC are also looking into such workflow systems to offer services for scientists. It is not yet fully clear which environment will be offered in these cases or exactly what types of services data infrastructures should offer. EUDAT will continue to work with community experts to test service concepts that allow users to execute workflows on data stored in the EUDAT data domain.

AGENDA

Workflows Chairs: Morris Riedel, Christian Page
Morris Riedel: Workshop Results
Christian Page: Climate Use Case
Erhard Hinrichs: Linguistics Use Case
Other contributions and General discussion on Workflow Issues in EUDAT
Update New Services
Final Round on New Services

CHAIR: LEIF LAAKSONEN, COLLABORATION DIRECTOR, CSC – IT CENTRE FOR SCIENCE & RDA EUROPE COORDINATOR

DATE & TIME: WEDNESDAY 30TH OCTOBER – 11:00 -13:00





PLENARY ROOM:RAFFAELLO + TINTORETTO + BOTTICELLI


OVERVIEW:

Leif Laaksonen, CSC – IT Centre for Science & Coordinator of the RDA Europe initiative – the European driver of RDA chairs the third plenary session looking at **Global Data Infrastructure Components**. This session provides an insight to the Research Data Alliance, launched in March 2013, and an update on the progress to date by John Wood. The session includes updates from some RDA groups - **Data Foundation and Terminology** – Raphael Ritz, **Metadata** – Bill Michener, **PID Info Type** - Tobias Weigel, **Data Type Registry** - Daan Broeder & **Practical Policy**- Rainer Stotzka - comprising of stakeholders who voluntarily come together to exchange knowledge, share discoveries, discuss barriers and potential solutions, explore and define policies and test as well as harmonise standards to enhance and facilitate global data sharing. RDA promotes and encourages both a bottom-up and interdisciplinary approach to solving global challenges, inviting researchers, scientists, data practitioners and other interested stakeholders from around the world to work together to achieve the RDA vision - An open, seamless, self-regulatory global digital data infrastructure that is the foundation for discovery and progress. Leif Laaksonen

AGENDA:

11:00-11:20	The Research Data Alliance: Status Update, John Wood, RDA Europe Council Chair & Secretary General Association Commonwealth Universities
	<p>Professor John Wood CBE, FEng is the Secretary-General of the Association of Commonwealth Universities and High Level Expert Group on Scientific Data Information Chair & European Research Area Board Chair. He graduated from Sheffield University in metallurgy and went to Cambridge University for his Ph.D. (Darwin College) where he subsequently stayed on as Goldsmith’s Research Fellow at Churchill College.</p> <p>He was a founder member of the European Strategy Forum for Research Infrastructures and became chair in 2004 where he was responsible for producing the first European Roadmap. He became the first chair of the European Research Area Board in 2008 responsible for high level advice to the European Commission and in 2009 produced a long term strategic vision entitled “Preparing Europe for a New Renaissance.”</p> <p>He was elected as a fellow of the Royal Academy of Engineering in 1999 and is currently a member of their Council and International Committee. He was made a Commander of the British Empire in 2007 for "services to science," and in 2010 was made an “Officer of the Order of Merit of the Federal Republic of Germany”.</p>

11:20-11:30	Data Foundation and Terminology - Raphael Ritz, Head of Data Science Group at RZG, Max Planck Society, Germany
	Raphael Ritz leads the Data Science and Services group within the computing center of the Max Planck Society. In addition, he represents the MPG in international contexts focusing on data handling.
11:30-11:40	Metadata - William (Bill) Michener, Professor and Director of e-Science Initiatives for University Libraries, University of New Mexico & DataONE Principal Investigator
	<p>Bill Michener is Professor and Director of e-Science Initiatives at the University of New Mexico's University Libraries. He serves as Project Director for two large National Science Foundation supported projects: (1) Data Observation Network for Earth (DataONE)—a large DataNet project that supports cyberinfrastructure development and community engagement for the biological, environmental, and Earth sciences; and (2) the New Mexico Experimental Program to Stimulate Competitive Research. He is actively involved in research related to creating information technologies supporting data-intensive science, development of federated data systems, and community engagement and education. He has a PhD in Biological Oceanography from the University of South Carolina and has published extensively in marine science, as well as the ecological and information sciences.</p> <p>Bill has authored five books and more than 100 journal articles and book chapters. He has a strong background and training in organizational sustainability and governance, project management, and meeting facilitation. Presently, he serves on the Board of Directors (or Administrative Board) for Dryad, Inc., the Organization for Tropical Studies, and the Cornell Lab of Ornithology, as well as the Governance Committee for an emerging international organization that seeks to nurture a network of organizations that are involved in Public Participation in Scientific Research (i.e., citizen science). He serves as editor of the Ecological Society of America's Ecological Archives, Associate Editor for Ecological Informatics, and as a member of the Ecology Editorial Board, and was recently appointed to the Technical Advisory Board for the Research Data Alliance</p>
11:40-11:50	PID Information Types - Tobias Weigel German Climate Computing Center (DKRZ) / University of Hamburg
	Tobias Weigel holds a Diploma degree in Geoinformatics from the University of Münster, Germany. Since 2010, he has been working at the German Climate Computing Center (DKRZ) as a software engineer and architect for various e-science infrastructure and project activities, including IS-ENES and EUDAT, with emphasis on lightweight web services and reusable components. Current activities concern usage scenarios and tool support for PIDs across these areas, which also relates to the work done in RDA. He is also a PhD student at the University of Hamburg, working on a PID-centric topic with particular focus on massive PID usage at data centers and infrastructures.
11:50-12:00	Data Type Registry – Daan Broeder, Max Planck Institute For Psycholinguistics, Nijmegen
	<p>Daan Broeder, for many years senior developer for archive and infrastructure solutions and now deputy director of TLA unit of the MPI for Psycholinguistics, is technologist (electronics and IT) and has a long record in leading development tasks in international projects. Currently he is a member of the executive board of the Dutch CLARIN project and participating several EU projects concerned with research infrastructure development as DASISH and EUDAT.</p> <p>In addition he is convener of new ISO standards in the linguistic domain (TC37/SC4: Component Metadata Infrastructure "CMD" and persistent identification "PISA").</p>

12:00-12:10	RDA Working Group Practical Policy – Rainer Stotzka, Karlsruhe Institute of Technology (KIT), Institute for Data Processing & Electronics, Software Methods Group Head
	<p>Dr. Rainer Stotzka studied technical cybernetics in Stuttgart and received his Ph.D. degree in computer science from the University of Mannheim, Germany, in 1995. At Karlsruhe Institute of Technology (KIT), Institute for Data Processing and Electronics, he leads the department “Software Methods” and a software development group “Big Data”. Within the Helmholtz activity “Large Scale Data Management and Analysis” he is responsible for the “Data Life Cycle Lab Key Technologies” inducing novel research data management methods for many scientific communities. He is involved in various national and international research projects and initiatives, e.g. DARIAH, eCodicology, EUDAT, RDA and the German Alliance for Research Data Management. Within the last five years he and his research groups published more than 40 papers in peer-reviewed conferences and journals.</p>
12:10-13:00	Panel discussion



PLENARY SESSION: WRAP-UP AND CONCLUSIONS

CHAIR: DAMIEN LECARPENTIER, EUDAT PROJECT MANAGER, CSC – IT CENTRE FOR SCIENCE

DATE & TIME: WEDNESDAY 30TH OCTOBER – 13:00 -13:30

PLENARY ROOM:RAFFAELLO + TINTORETTO + BOTTICELLI

OVERVIEW:

Damien Lecarpentier – EUDAT Project Manager chairs the wrap-up & conclusion plenary session which summarises the conference proceedings and the outcome of discussions during the parallel tracks. The winner of the EUDAT New Services Survey will be announced at the end of this session.



Damien Lecarpentier has an MA in Political Science and PhD in Social Sciences from the Advanced School of Social Sciences in Paris. He joined CSC in 2009 and now works as Project Director within the Research Infrastructures unit. Damien has been involved as work package and task leader in policy work and outreach activities in several FP7-funded projects in the areas of grids (EGI_DS), HPC (DEISA2, PRACE, EESI) and e-Infrastructure policy (e-IRGSP2, e-Infranet). He is currently managing the EUDAT project (www.eudat.eu) which aims to build a sustainable cross-disciplinary and cross-national data infrastructure providing a set of shared services to access and preserve research data.

WEDNESDAY 30 OCTOBER 2013 - MEETINGS & ASSOCIATED WORKSHOPS

ROOM	PUCCINI	TIZIANO	VERDI
13:30 - 14:30		<i>Networking Lunch</i>	
14:30 - 16:00	EPIC User Meeting (Closed)	PID Information Types WG Meeting	DIGITAL PRESERVATION OF CULTURAL DATA workshop
16:00 - 16:30		<i>Coffee</i>	
16:30 - 18:00	EPIC User Meeting (Closed)	PID Information Types WG Meeting	DIGITAL PRESERVATION OF CULTURAL DATA workshop

WORKSHOP: PID INFORMATION TYPES WG MEETING

DATE & TIME: WEDNESDAY 30TH OCT – 14:30 – 18:00

ROOM: TIZIANO

OVERVIEW:

The WG meeting will take place just a couple of weeks after the 2nd RDA Plenary. We will review the outcomes of the plenary meeting and continue potential discussion topics, review the use case documents, discuss possible PID types, mix and synchronize with other WGs, get newcomers up to speed and have a glance at the next steps. The meeting will be interactive and creative, we will try to have a minimum set of presentations and hopefully a maximum of ideas and pragmatism. If you are from the EU and did not have a chance to attend the 2nd RDA Plenary, you should join us in Rome as this is a good way for you to keep track of WG activities.

The meeting is open to anyone interested in WG activities, newcomers and known faces alike - if you are not a member of the WG yet, feel free to use the meeting to become involved!

CONTACT PERSON:

Tobias Weigel - Department of Data Management - Deutsches Klimarechenzentrum GmbH (German Climate Computing Center)

DATE & TIME: WEDNESDAY 30TH OCT – 14:30 – 18:00

ROOM: VERDI

OVERVIEW:

Medicine and Natural sciences, including astronomy, biology, chemistry, earth sciences and physics, already make use of standardized formats and e-Infrastructures services to generate, curate, share and analyse research data.

The need for novel more efficient and affordable solutions for digital preservation is now increasing also in the Social Science and Humanities, in particular the Digital Cultural Heritage (DCH) sector is producing a large volume of digital content that needs to be safely stored and curated, permanently accessed, and easily shared and re-used by researchers.

Each digitisation programme is currently addressing the issue of preservation in a separate manner, a shared implementation of common e-Infrastructure layers could be beneficial and cost effective.

Moreover, preservation models are often inspired by the ISO OAIS standard, where transfers and preservation are built on information packages containing both data and metadata. Even if the transferred files are in standard formats, the implementation of standards cannot be guaranteed and it is not in control neither by the institutions that produces the software for implementing them, nor by the memory institutions.

E-Infrastructures and DCH communities entered a dialogue in the last years and several data-infrastructure projects exist and look how to set up data infrastructures, including DCH use cases:

- DCH-RP: Digital Cultural Heritage Roadmap for Preservation
- SCIDIPES: SCience Data Infrastructure for Preservation - Earth Science
- APARSEN: Alliance Permanent Access to the Record of Science in Europe network
- EUDAT: Towards a European Collaborative Data Infrastructure
- CHAIN REDS: Coordination and Harmonization of Advanced e-Infrastructures for Research and Education Data Sharing
- DARIAH: Digital Research Infrastructure for Arts and Humanities
- DASISH: Data Service Infrastructure for the Social Science and Humanities
- CLARIN: Common Language Resources and Technology Infrastructure
- SCAPE: SCALable Preservation Environments

At the same time, new projects are about to start, such as a joint Pre-Commercial Procurement project – which is now under negotiation – whose main objective is the development of an open source software licensed reference implementation for different format standards as a tool to be used by memory institutions to check conformance with standard specifications.

Aim of the workshop is to bring together such kind of projects and initiatives working world-wide in the domain of DCH, e-infrastructures and digital preservation to share and present the advancements in the state of the art, find synergies and discuss opportunities for cooperation, starting from concrete use cases.

Target Users:

- Researchers in the humanities
- Teaching and learning actors (schools, training centers, university courses)
- Cultural and creative industry for the creative use and re-use of the digital cultural content
- Content providers (e.g. cultural managers of national institutions and libraries, small institutions, private and public publishers, etc.).
- Policy makers and programme owners
- E-infrastructure providers, technology providers and R&D institutions
- R&D projects and initiatives focusing on digital preservation

AGENDA:	
14:30 – 14:45	Welcome and introduction - Antonella Fresa, Promoter Srl
<i>First Part: DCH and the e-infrastructures</i>	
14:45 – 15:10:	Using EUDAT services to replicate, store, share, and find cultural heritage data in Poznań Supercomputing and Networking Center - Maciej Brzeźniak, Poznan Supercomputing and Networking Center – Damien Lecarpentier, CSC – IT Center for Science
15:10 – 15:35	Authentication and Authorisation in the Cultural Heritage community - Roberto Barbera, INFN
15:35 – 16:00	Scalability in preservation of cultural heritage data - Simon Lambert, Scientific Computing Department – STFC
16:00 – 16:30	Coffee Break
<i>Second Part: OAIS model, standards, provenance and authenticity</i>	
16:30 – 16:50	Standard models and formats for digital preservation - Börje Justrell, Swedish National Archives
16:50 – 17:15	Implementation of authenticity evidence record model for supporting preservation scenarios - Luigi Briguglio, Engineering R&D Lab
17:15 - 17:30	Coordination of digitisation, digital access and digital preservation in Sweden - Sanja Halling, Digisam
17:30	Conclusions - Antonella Fresa, Promoter Srl

CONTACT PERSON:

For more information please contact Claudio Prandoni, Promoter Srl, prandoni@promoter.it or visit the workshop web page at <http://www.digitalmeetsculture.net/article/digital-preservation-of-cultural-data/>

PARTICIPANTS LIST

This list includes all registered participants up to & including 24th October 2013 and ONLY those that agreed to publish their contact details.

Last Name	First Name	Organisation	Country
Ansari	Agnès	CNRS/IDRIS	France
Bailo	Daniele	EPOS -INGV	Italy
Baxter	Rob	EPCC, University of Edinburgh	United Kingdom
Berni	Alessandro	STO-CMRE	Italy
Bruch	Christoph	Helmholtz Open Access Coordination Office	Germany
Bunakov	Vasily	STFC	United Kingdom
Cacciari	Claudio	CINECA	Italy
Carenton-Madiec	Nicolas	CNRS-IPSL	France
Carli	Anna Caterina	Ministero dello Sviluppo Economico	Italy
Cimino	Giampaolo	NATO CMRE	Italy
Clausen	Nanna Floor	Danish Data Archive	Denmark
Coro	Giampaolo	CNR - ISTI	Italy
De La Mar	Jurry	T-Systems International	Germany
De Smedt	Koenraad	University of Bergen	Norway
Denvil	Sebastien	CNRS/IPSL	France
Dillo	Ingrid	Data Archiving and Networked Services (DANS)	Netherlands
Dooley	Rion	Texas Advanced Computing Center/UT Austin	USA
Doorn	Peter	Data Archiving and Networked Services (DANS)	The Netherlands
Dovey	Matthew	Jisc	UK
Eide	Hans	University of Oslo	Norway

Last Name	First Name	Organisation	Country
Fares	Massimo	INGV	Italy
Fiameni	Giuseppe	CINECA	Italy
Fotland	Margaret Louise	University of Oslo	Norway
Frackowiak	Richard	CHUV University Hospital	Switzerland
Friis-Christensen	Anders	European Commission - Joint Research Centre	Italy
Gagliardi	Fabrizio	Barcelona Supercomputing Center	Spain
Gjesdal	Anje Müller	UNIVERSITY OF BERGEN	Norway
Håkansson	Carl Johan	Royal Institute of Technology Center for High Performance Computing	Sweden
Hanahoe	Hilary	Trust - It Services LTD	United Kingdom
Hanke	Thomas	Universität Hamburg	Germany
Harmsen	Henk	Data Archiving and Networked Services (DANS)	Netherlands
Hasler	Tim	Zuse Institute Berlin	Germany
Iozzi	Maria Francesca	University of Oslo	Norway
Islam	Mohammed Shahidul	AL-ARAFAH ISLAMI BANK LIMITED	Bangladesh
Jamin	Bert	Utrecht University Netherlands	The Netherlands
Jaunsen	Andreas	UNINETT Sigma	Norway
Jones	Bob	CERN	Switzerland
Kavoussanakis	Kostas	EPCC, University of Edinburgh	United Kingdom
Kemppainen	Teemu	CSC-IT Center for Science	Finland
Koski	Kimmo	CSC-IT Center for Science	Finland
Kupiainen	Irina	CSC - IT Center for Science Ltd.	Finland
Kvale	Live	University of Oslo Library	Norway
Laaksonen	Leif	CSC - IT Center for Science Ltd.	Finland

Last Name	First Name	Organisation	Country
Lannom	Larry	CNRI	USA
Laure	Erwin	KTH	Sweden
Lecarpentier	Damien	CSC - IT Center for Science	Finland
Mallmann	Daniel	Juelich Supercomputing Centre	Germany
Marker	Hans Jørgen	Swedish National Data Service	Sweden
Martin	Paul	UNIVERSITY OF EDINBURGH	UK
Mayo-Garcia	Rafael	CIEMAT	Spain
Michelini	Alberto	INGV	Italy
Michener	William	University of New Mexico	USA
Nadiradze	Kakha	Association for Farmers Rights Defense, AFRD	Georgia
Nardi	Silvia	Istituto Nazionale di Geofisica e Vulcanologia	Italy
Ostasz	Marcin	CO-CHAIR Barcelona Supercomputing Center Eudat WP2	Spain
Purcell	Andrew	International Science Grid This Week	Switzerland
Pursula	Antti	CSC - IT CENTERE FOR SCIENCE	Finland
Rauste	Päivi	CSC-IT Center for Science	Finland
Rice	Robin	University of Edinburgh	United Kingdom
Rombouts	Jeroen	3TU.Datacentrum	The Netherlands
Rosén	Victoria	University of Bergen	Norway
Ruggieri	Federico	GARR/INFN	Italy
Rybicki	Jedrzey	Forschungszentrum Juelich	Germany
Sabatino	Roberto	Dante	UK
Salmivalli	Riina	CSC - IT Center for Science Ltd.	Finland
Schwardmann	Ulrich	GWDG	Germany
Smeele	Ton	Utrecht University	The Netherlands

Last Name	First Name	Organisation	Country
Stehouwer	Herman	MPG/RDA- Europe	Netherlands
Stotzka	Rainer	Karlsruhe Institute of Technology	Germany
Svensson	Gert	KTH/PDC	Sweden
Toussaint	Frank	Deutsches Klimarechenzentrum	Germany
Toutain	Thierry	USIT - University of Oslo	Norway
van Wezel	Jos	Karlsruhe Institute of Technology	Germany
Weigel	Tobias	DKRZ	Germany
Widmann	Heinrich	DKRZ	Germany
Wood	John	Association of Commonwealth Universities (ACU)	United Kingdom
Ashley	Kevin	Digital Curation Centre, University of Edinburgh	UK
Beckers	Toine	DDN	Netherlands
Berber	Fatih	Gesellschaft für Wissenschaftliche Datenverarbeitung MbH Göttingen	Germany
Bertazzo	Matteo	CINECA	Italy
Bijsterbosch	Magchiel	SURF	Netherlands
Birney	Ewan	EBI	UK
Blake	Richard	STFC Daresbury Laboratory	UK
Block	William	Cornell University	USA
Botha	Harald	Nordforsk	Norway
Brieger	Leesa	Renci, University of North Carolina - Chapel Hill	USA
Broeder	Daan	MPI for Psycholinguistics	Netherlands
Bruce	Rachel	Jisc	UK
Brzeźniak	Maciej	Institute of Bioorganic Chemistry of the Polish Academy of Sciences Poznan Supercomputing and Networking Center	Poland
Budich	Reinhard	MPI für Meteorologie	Germany
Carter	Adam	EPCC, University of Edinburgh	UK
Castelli	Donatella	CNR - ISTI	Italy

Last Name	First Name	Organisation	Country
Castrignanò	Tiziana	Cineca	Italy
Contino	Ugo	CINECA	Italy
Cook	Chuck	EMBL - European Bioinformatics Institute	UK
Coutin	Stephane	CINES	France
Coveney	Peter	University College London	UK
D'Antonio	Mattia	Cineca	Italy
de la Torre	Victor	INB	Spain
de Witt	Shaun	STFC	UK
Di Iorio	Angela	Sapienza University of Rome	Italy
Doldirina	Catherine	Joint Research Centre, the European Commission	Italy
Elbers	Willem	Max Planck Institute for Psycholinguistics	Netherlands
Ellenbroek	Anton	FAO	Italy
Erastova	Elena	RZG	Germany
Ferrari	Tiziana	EGL.eu European Grid Infrastructure	Netherlands
Fiore	Nicola	Lifewatch	Italy
Formisano	Ciro	Engineering Ingegneria Informatica	Italy
Freire	Nuno	The European Library	Netherlands
Fuentes Brenes	Daniel	LifeWatch	Spain
Galeazzi	Fulvio	Consortium GARR	Italy
Ganguly	Raman	University of Vienna	Austria
Gebre	Binyam Gebrekidan	Max Planck Institute for Psycholinguistics	Netherlands
Girona	Sergi	Barcelona Supercomputing Center	Spain
Glaves	Helen	British Geological Survey	UK
Glinos	Kostas	European Commission	Belgium
Hajic	Jan	Charles University in Prague	Czech Republic

Last Name	First Name	Organisation	Country
Hanganu	Gabriel	Jisc	UK
Harris	Nicole	Terena	Netherlands
Hasan	Adil	UNINETT Sigma	Norway
Hellmich	Martin	CERN	Switzerland
Hodson	Simon	CODATA	France
Huber	Robert	Universität Bremen- MARUM	Germany
Hyvärinen	Pietari	CSC - IT Center for Science Ltd.	Finland
Jacquemot-Perbal	Marie-Christine	INIST-CNRS	France
Jiménez-Peris	Ricardo	Universidad Politécnica de Madrid	Spain
Jones	Sarah	Digital Curation Centre	UK
Justrell	Borje	Riksarkivet (National Archives of Sweden)	Sweden
Kakaletris	Georgios	University of Athens	Greece
Kálmán	Tibor	GWDG	Germany
Kamocki	Pawel	IDS Mannheim / University Paris V / University of Münster	Germany
Kanellopoulos	Christos	GRNET	Greece
Kennedy	Alison	EPCC, University of Edinburgh	UK
Kleinjans	Jos	Maastricht University	Netherlands
Kurzawe	Daniel	Niedersächsische Staats- und Universitätsbibliothek - SUB - Göttingen	Germany
Lambert	Simon	STFC	UK
Lampen	Lari	Max Planck Institute for Psycholinguistics	Netherlands
Lengert	Maryline	ESA	Italy
Livenson	Ilja	KTH	Sweden
Llurba	Rosend	NWO/EW	Netherlands
Lukkarinen	Ari	CSC - IT Center for Science	Finland
Maggi	Marcello	Istituto Nazionale Fisica Nucleare	Italy
Magoutis	Kostas	ICS-FORTH	Greece

Last Name	First Name	Organisation	Country
Malapela	Them bani	FAO Food and Agriculture Organisation of the United Nations	Italy
Manghi	Paolo	CNR-ISTI	Italy
Manieri	Andrea	Engineering Ingegneria Informatica Spa	Italy
Manzano	Cristina	Forschungszentrum Juelich	Germany
Manzi	Andrea	CERN	Italy
Marinucci	Monica	Oracle	Switzerland
Massol	Marion	CINES	France
Meyer	Jörg	Karlsruher Institute of Technology	Germany
Meyer	Norbert	Institute of Bioorganic Chemistry of the Polish Academy of Sciences Poznan Supercomputing and Networking Center	Poland
Mirtl	Michael	Environment Agency Austria	Austria
Morelli	Giovanni	CINECA	Italy
Morris	Huw	Rutherford Appleton Laboratory	UK
Muscella	Silvana	Trust - It Services LTD	UK
Nordling	Josefine	CSC	Finland
Öster	Per	CSC - IT Center for Science Ltd.	Finland
Pagano	Pasquale	CNR - ISTI	Italy
Page	Christian	Cerfac	France
Parodi	Antonio	Cima Research Foundation	Italia
Pereira	José	INESC TEC & U. Minho	Portugal
Pezzati	Luca	CNR INO	Italy
Pilicheva	Ekaterina	EMBL - EBI	UK
Pozzo	Riccardo	CNR - Dipartimento Scienze Umane e Sociali, Patrimonio Culturale	Italy
Ramakrishnan	Chandrasekhar	ETH Zurich	Switzerland
Reetz	Johannes	Rechenzentrum Garching / MPG	Germany
Riedel	Morris	Juelich Supercomputing Centre	Germany
Ritz	Raphael	RZG-MPG	Germany

Last Name	First Name	Organisation	Country
Saarinen	Samuli	CSC-IT Center for Science Ltd.	Finland
Salvadore	Francesco	CINECA	Italy
Schentz	Herbert	Umweltbundesamt Gmbh	Austria
Simarro	Christian	BSC	Spain
Sipos	Gergely	EGL.eu European Grid Infrastructure	Netherlands
Skow	Dane	Herculesstichting	Belgium
Straňák	Pavel	Charles University in Prague	Czech Republic
Streit	Achim	Karlsruhe Institute of Technology (KIT)	Germany
Thongtra	Patcharee	Uni Research Ltd	Norway
van de Sanden	Mark	SURFsara	Netherlands
van Dijk	Paul	Surfnet	Netherlands
Van Uytvanck	Dieter	Max Planck Institute for Psycholinguistics / CLARIN ERIC	Netherlands
Vermeulen	Ivar	VU University Amsterdam	Netherlands
Vitlacil	Dejan	PDC	Sweden
Zasada	Stefan	UCL	UK