



Preliminary OSDC PIRE 2014 Amsterdam Workshop Program

Introduction & Background

The OSDC PIRE program participants who are destined for sites other than Amsterdam will come to this workshop. We will also be inviting local researchers and distinguished external speakers who have kindred interests. The workshop will stimulate research, educate participants and initiate international research relationships.

The workshop is sponsored by the NSF-funded OSDC PIRE project. It will also be supported from various resources: University of Amsterdam (<http://www.uva.nl/en/home>) It follows on from a successful OSDC PIRE 2013 Workshop, held in the Informatics Forum, University of Edinburgh.

All participants should register for the event at the following web site so that we can plan computing facilities, catering and meeting room space: <https://osdc-pire-2014-amsterdam.eventbrite.com>

Overview & Structure

Day	Topics	Social events
Monday	Welcome. Introduction to OSDC software and methods. PICO presentations ¹ : self-introductions.	Refreshments
Tuesday	Dealing with Big Data.	Dinner
Wednesday	The research challenge/bazaar	Canal cruise Dinner
Thursday	Projects developments. Technology topics of general interest with hands-on opportunities.	Free time
Friday	Conclusion challenge Wrap-up	(Optional) Museum Trip

¹ A PICO presentation allows a presenter to make a quick pitch for their topic, two slides and two minutes. They then can present their topic in more detail by a) having a full set of slides on the workshop web site, and b) talking further to interested parties beside a poster, by showing the slides or by showing a demo, in a specified location.

DAY 1 (Mon, 16-Jun 2014)

<i>Time</i>	<i>Description</i>
08:45-09:30	Registration
09:30-11:00	Session 1: Welcome & Introduction Chair: Paola Grosso & Cees De Laat
09:30-10:00	Welcome and "Why are we here?"
10:00-11:00	Title: Using the OSDC for Data-Intensive Research, Speaker: Robert Grossman, <i>U of Chicago</i> We give an introduction to the Open Science Data Cloud (OSDC), a science cloud that is operated by the not-for-profit Open Cloud Consortium. The OSDC is a persistent cloud-based infrastructure that allows scientists to manage, analyze, integrate and share medium to large size scientific datasets. The OSDC contains data from a variety of scientific disciplines, from earth science to biology. The OSDC is currently one of the largest cloud-based infrastructures devoted to scientific data.
11:00-11:30	<i>Coffee break</i>
11:30-13:00	Session 2: OSDC PIRE participant self-introduction session Chair: Heidi Alvarez, <i>FIU</i>
	<p>Participants in the workshop, including organizers and speakers, will introduce themselves and provide a record about themselves for consultation during and after the workshop. You were invited to provide two slides. You can also provide a title slide with your name and home institution on it; we'll make this title slide if you didn't. The other two slides should contain what you think is important that people at the OSDC workshop know about you. The three slides will be spliced into an auto-advancing presentation: 1 minute for the title slide, during which time the previous speaker sits down and the next takes their place on the podium. Your next two slides will auto advance after a minute (no animations!).</p> <p>You can speak to them or say related things. This is called a PICO presentation; see below. BE YE WARNED: if you don't provide slides (PowerPoint, Keynote or PDF – we'll do the conversion) we will provide two blank slides for your two minutes. The order will be random but pre-published.</p> <p>You have an important role to play during the first day of the Amsterdam Workshop. In order to get to know each other we would like you to prepare a PICO presentation as an interesting way to present yourself and the area(s) of your research. Below are the links with details for PICO presentations also known as introductory "2-minute's-madness".</p> <p>Presenting Interactive COntent, or PICO, sessions highlight the essence of a particular research area – just enough to get excited about a topic without being overloaded with information. PICO sessions combine the best of oral and poster presentations. Every PICO author presents 2 slides in a "2 minute's madness" and afterwards, all attendees have enough time to watch the presentation again (we will post on the OSDC PIRE Facebook Fan Page & Workshop website) and hold discussions with both the author and other attendees. See this short video. https://www.youtube.com/watch?feature=player_embedded&v=A36FZZiZYVE</p> <p>Find more details here http://www.egu2013.eu/guidelines/author_guidelines_pico.html</p> <p>Additionally, we would like to post your picture and a short Bio (one paragraph) on the OSDC PIRE Facebook Fan Page.</p> <p>After the PICO presentations, we will have a general discussion about the workshop and what you want to get from it.</p>
13:00-14:00	<i>Lunch</i>
14:00-17:15	Session 3: Enabling data-intensive discovery with the Open Science Data Cloud Chair: Maria Patterson, <i>U of Chicago</i>
14:00-14:30	Title: The Namibia Flood Dashboard Speaker: Jill Hardy and Race Clark, <i>U of Oklahoma</i> A report from PIRE fellows who have completed their experiences abroad working on flood and drought risk management in Namibia.



14:30-14:50	<p>Title: Combining Latent Topics with Document Attributes in Text Analysis Speaker: Nelson Auner, <i>U of Chicago</i> A report on statistical text analysis at the University of Chicago Booth School of Business</p>
14:50-15:40	<p>Title: Tutorial and Hands-on OSDC exercises Speaker: Maria Patterson, <i>U of Chicago</i> This is an introduction to using the OSDC. Please make sure to bring your laptops so you can follow along. Also please register to use the OSDC at least a week before the start of the workshop.</p>
15:40-16:00	<i>Coffee break</i>
16:00-17:00	<p>Title: Reproducible Research and Collaborative Tools Speaker: Maria Patterson, <i>U of Chicago</i> Motivation and tips for ensuring reproducibility in data analysis and enabling easier scientific collaboration using literate statistical programming. You may find these tools useful for the OSDC PIRE Challenge!</p>
17:00-17:15	<p>Title: Wrap-up, Review, and PIRE Challenge Kick-off Facilitators: Heidi Alvarez, Robert Grossman, Paola Grosso & Cees De Laat</p>
17:30- 19:00	<i>Refreshments at Polder</i>



DAY 2 (Tues, 17-Jun 2014)

09:00-10:40	Session 4: Applications & Infrastructure 10 -20 Minute Talks Chair: Robert Grossman, U of Chicago
9:00-9:20	Title: Project Matsu Speaker: Maria Patterson <i>University of Chicago</i>
9:20-9:40	Title: Yates and Tukey Speaker: Rafael Suarez <i>University of Chicago</i>
9:40-10:00	Title: Bionimbus Speaker: Robert Grossman <i>University of Chicago</i>
10:00-10:30	Title: Big Data & Scientific Remote Collaboration Projects at LARC-USP Speaker: Fernando Redigolo, USP LARC
10:30-11:00	Title: Scrying the next generation of data-intensive research infrastructure: Research at the Data-Intensive Research Group in Edinburgh Speaker: Paul Martin <i>Edinburgh</i>
11:00-11:30	<i>Coffee break</i>
11:30-15:00	Session 5a: Amsterdam Big Data Research Chair: Paola Grosso & Cees de Laat, U of Amsterdam
11:30-12:00	Title: GBIF data Speaker: Cees Hof, <i>University of Amsterdam</i>
12:00-12:30	Title: The Amsterdam DSRC Speaker: Marcel Worring <i>University of Amsterdam</i>
12:30-13:00	Title: ENVRI and Big Data Infrastructure Speaker: Cees de Laat <i>University of Amsterdam</i>
13:00-14:00	<i>Lunch</i>
14:00-15:00	Title: The Challenges of Digital Video as Big Data Speaker: Jeff Weekley, <i>Naval Postgraduate School Monterey</i>
15:00-15:30	<i>Early Coffee break (Location :B1.23 (Science Park 904))</i>
15:30-17:30	Session 5b: Using ENVRI Chair: Massimo Argenti, ESA (Location :B1.23 (Science Park 904))
15:30-17:30	Title: Tutorial and Hands-on on the ENVRI data (Location :B1.23 (Science Park 904))
17:30-19:30	Dinner <i>Greek Restaurant on the Amstel</i>



DAY 3 (Wed, 18-Jun 2014)

09:00-13:00	Session 6 The research challenge/bazaar Chair: Cees de Laat/Paola Grosso
09:00-10:30	<p>The goal is to stimulate wide-ranging research discussion around topics of interest to the OSDC PIRE Fellows. Those running the workshop or giving talks longer than 20 minutes already, cannot initiate a discussion – they have already had their chance. Every OSDC PIRE Fellow and their host counterparts are strongly encouraged to make a pitch for research they have done, research they are doing or research they would like to do. This may be research in their own institution or in the host institution they are about to visit. They should make a strong pitch to get others interested in follow up discussions.</p> <p>The session will be in two parts: Part 1: PICO pitches to raise interest – fast and furious to leave plenty of time for Part 2: Parallel and self-organized discussions with gentle shepherding.</p> <p><i>You should send us your PICO slides for this session by mid-day (BST) Monday 16 June</i>, so there is time to splice them together and to put them on the workshop web site. You should have a title slide that includes your name and home institution and two following slides without animation, for your PICO talk. After these you can have other slides which will go on the web site, but which you will not be able to use during your PICO pitch. Again the slides will auto advance. The order will correspond to when we receive your slides; the early birds get the first talks.</p> <p>When these have been presented, each PICO presenter should be prepared to lead a follow up discussion (possibly more than once) using one of the following three methods:</p> <ol style="list-style-type: none"> 1. Showing your full set of slides and talking about them, 2. Standing by a poster and talking about it (we'll try to get these posters on the workshop web site), or 3. Showing a demonstration of your research results.
10:30-12:00	Group brainstorming
12:00-14:00	<i>Lunch and move to location of the excursion</i>
14:00-18:00	Excursion Boat trip along the canals
18:20-21:30	<i>Dinner Indonesian Rijstafel</i>



DAY 4 (Thurs, 19-Jun 2014)

09:00-10:30	Session 7: Title: Tools for Development and Support of Large-Scale Data-Intensive Applications, Speaker: Shantenu Jha, Rutgers University USA (Location :B1.23 (Science Park 904))
10:30-11:00	Coffee break
11:00-11:40	Session 8: PIRE partners contributions Chair: Paola Grosso (Location :B1.23 (Science Park 904))
11:00-11:20	Title: Data-Intensive Research at AIST, Speaker: dr. Isao Kojima – ITRI- AIST
11:20-11:40	Title: Research at UWI Speaker: Brigitte Collins UWI
11:40-16:00	Session 9: Parallel Sessions on Project Development
11:40-14:00	Session 9A: Project Development in AIST Chair: Isao Kojima (Location :TBD (Science Park 904))
11:40-12:30	Orientation to research at AIST.
12:30-14:00	<i>Extended lunch with informal discussions</i>
11:40-16:00	Session 9B: Project Development in Sao Paolo Chairs: Fernando Redigolo (Location :TBD (Science Park 904))
11:30-12:30	
12:30-14:00	<i>Extended lunch with informal discussions</i>
14:00-16:00	
11:40-16:00	Session 9C: Project Development in Edinburgh Chairs: Paul Martin (Location :TBD (Science Park 904))
11:30-12:30	
12:30-14:00	<i>Extended lunch with informal discussions</i>



14:00-16:00	
11:40-16:00	Session 9D: Project Development in Amsterdam Chairs: Paola Grosso (Location :TBD (Science Park 904))
11:30-12:00	Introduction to the projects in Amsterdam: <ul style="list-style-type: none"> ○ Ana Oprescu on OSDC to OSDC ○ Zhiming Zhao on OSDC to SDN ○ Miroslav Zivkovic on OSDC and Big Data issues
12:00-13:00	<i>Lunch</i>
13:00-15:00	ExoGENI tutorial
15:00-16:00	Demo and discussions

DAY 5 (Fri, 20-Jun 2014)

Each site that will host PIRE visiting researchers should have a parallel track from morning coffee to afternoon coffee to run a detailed orientation program matching their visitors and projects. Sites may choose to use part of their time, and participants can then join in other parallel sessions.

09:00-10:30	Session 9: Challenge presentations and prizes Chair: Cees de Laat/Paola Grosso
09:00-10:30	<i>A student from each group will present their work and the jury will decide on the winner.</i>
10:30-11:00	<i>Coffee break</i>
11:00- 12:30	Session: 10 What have we learn & where next? Chair: Heidi Alvarez, Bob Grossman, Paola Grosso & Cees De Laat
11:00-12:00	What have we understood this week about data and how to make the best use of the data bonanza? What do we need to do & understand to make data-use easy and effective? How should we do that?
12:00-12:30	Wrap-up & valediction
12:30-14:00	<i>Lunch</i>
14:00-19:00	<i>Excursion (optional)</i> <i>Museumplein – the museum square with the Rijksmuseum, Van Gogh Museum and Stedelijk Museum.</i>

