



OSDC PIRE 2015 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 by the University of Amsterdam (<http://www.uva.nl/en/home>). It follows from a series of successful OSDC PIRE workshops at the University of Edinburgh (2012, 2013) and at the University of Amsterdam (2014).

All participants should register for the event at the following web site so that we can plan computing facilities, catering and meeting room space at osdc-pire-2015-amsterdam.eventbrite.com

Overview & Structure

| Day | Topics | Social events |
|-----------|--|---|
| Monday | Welcome. PICO presentations ¹ : self-introductions. Introduction to big data challenges, community solutions, and the OSDC. | Refreshments Location: Polder, Science Park 205 |
| Tuesday | Dealing with Big Data around the world: on-going research at OSDC PIRE host institutions. | Dinner Restaurant Mojo, Ringdijk3,1097AA Amsterdam www.mojo-amsterdam.nl |
| Wednesday | The research challenge/bazaar | Canal cruise Dinner Kantjil & de Tijger Spuistraat 291-293, 1012 VS Amsterdam http://www.kantjil.nl/en/ |
| Thursday | Technology topics of general. Research project development breakout sessions. | Free time |
| Friday | Conclusion challenge Wrap-up | (Optional) Museum Trip Museumplein – the museum square with the Rijksmuseum, Van Gogh Museum and Stedelijk Museum |

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, 08-Jun 2015)

| <i>Time</i> | <i>Description</i> |
|-------------|--|
| 08:45-09:30 | Registration (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 09:30-11:00 | Session 1: Welcome & Introduction Chair: Paola Grosso & Cees de Laat (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 09:30-10:00 | Welcome and “Why are we here?” |
| 10:00-11:00 | Title: Clouds and Commons for the Data Intensive Science Community Speaker: Robert Grossman, University of Chicago Large scale data commons like the Open Science Data Cloud are now practical due to the maturing of data center scale computing and are needed due to the growth of big data in many science disciplines, including biology, medicine, astronomy, and earth sciences. |
| 11:00-11:30 | <i>Coffee break</i> |
| 11:30-13:00 | Session 2: OSDC PIRE participant self-introduction session Chair: Maria Patterson, University of Chicago (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| | <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.</p> <p>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 invite you 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> (Location: CWI canteen, Science Park 123) |



| | |
|--------------|---|
| 14:00-17:15 | Session 3: Big Data Challenges, Solutions, and the Open Science Data Cloud Chair: Maria Patterson, University of Chicago (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 14:00-15:00 | Title: Open Science and Research Reproducibility Speaker: Tony Hey, University of Washington |
| 15:00-15:30 | Title: A data commons for the genomics community Speaker: Allison Heath, University of Chicago |
| 15:30-16:00 | <i>Coffee break</i> |
| 16:00-16:30 | Title: Cloud-based analysis of NASA satellite data on OSDC Speaker: Maria Patterson, University of Chicago |
| 16:30-17:00 | Title: Using OSDC – demos and tutorials Speaker: Maria Patterson, University of Chicago 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. |
| 17:00-17:15 | Title: Wrap-up, Review, and PIRE Challenge Kick-off Facilitators: Robert Grossman, Paola Grosso & Cees de Laat |
| 17:30- 19:00 | <i>Refreshments at Polder</i> (Location: Polder, Science Park 205) |



DAY 2 (Tues, 09-Jun 2015)

| | |
|-------------|--|
| 09:00-10:30 | Session 4: Research at the PIRE hosts Chair: Robert Grossman, <i>University of Chicago</i> (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 09:00-09:30 | Title: Data intensive research at AIST Speaker: Jason Haga, <i>AIST</i> |
| 09:30-10:00 | Title: Research at USP LARC Speaker: Fernando Redigolo, <i>USP LARC</i> |
| 10:00-10:30 | Title: Data intensive research in Edinburgh Speaker: Malcolm Atkinson, <i>University of Edinburgh</i> |
| 10:30-11:00 | <i>Coffee break</i> |
| 11:00-12:30 | Session 5: Amsterdam Big Data Research Chair: Paola Grosso, <i>University of Amsterdam</i> (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 11:00-11:30 | Title: Global Biodiversity Information Facility (GBIF) data Speaker: Cees Hof, <i>University of Amsterdam</i> |
| 11:30-12:00 | Title: Machine Learning Speaker: Max Welling, <i>University of Amsterdam</i> |
| 12:00-12:30 | Title: ENVRI and Big Data Infrastructure Speaker: Cees de Laat, <i>University of Amsterdam</i> |
| 12:30-14:00 | <i>Lunch</i> (Location: CWI canteen, Science Park 123) |
| 14:00-17:30 | Session 6: Data Intensive Science: workflow and provenance Chair: Zhiming Zhao, <i>University of Amsterdam</i> (Location: UvA, B1.23, Science Park 904) |
| 14:00-14:45 | Title: User Centered Provenance Management for Data-Intensive Platforms Speaker: Alessandro Spinuso, <i>KNMI</i> Talk and demo |
| 14:45-15:30 | Title: dispel4py tutorial – part I Speaker: Rosa Filgueira, <i>University of Edinburgh</i> |
| 15:30-16:00 | <i>Coffee break</i> |
| 15:30-17:30 | Title: dispel4py tutorial – part II Speaker: Rosa Filgueira, <i>University of Edinburgh</i> |
| 17:30-19:30 | Dinner <i>Restaurant Mojo, Ringdijk 3, 1097 AA Amsterdam, www.mojo-amsterdam.nl</i> |



DAY 3 (Wed, 10-Jun 2015)

| | |
|-------------|--|
| 09:00-12:00 | Session 7: The research challenge/bazaar Chair: Cees de Laat/Paola Grosso (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 09:00-09:50 | Title: Towards an Open Science Commons in Europe Speaker: Tiziana Ferrari, European Grid Infrastructure (EGI) |
| 09:50-10:10 | Title: Project Matsu in Namibia Speaker: Race Clark, PIRE Fellow |
| 10:10-12:00 | <p>Group brainstorming and optional research pitches: 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 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>You should send us your PICO slides for this session by end of day (BST) Tuesday 9 June, 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. |
| 12:00-14:00 | <i>Lunch</i> (Location: CWI canteen, Science Park 123) <i>Move to location of the excursion</i> |
| 14:00-18:00 | Excursion Boat trip along the canals |
| 18:30-21:30 | Dinner <i>Restaurant Kantjil & de Tijger, Spuistraat 291-293, 1012 VS Amsterdam, www.kantjil.nl/en/</i> |



DAY 4 (Thurs, 11-Jun 2015)

| | |
|-------------|--|
| 09:00-10:00 | Session 8: Handling data Chair: Miroslav Živković (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 09:00-09:30 | Title: Anomaly Detection Speaker: Piotr Zuraniewski, TNO, The Hague |
| 09:30-10:00 | Title: We are Big Data Speaker: Sander Klous, University of Amsterdam |
| 10:00-10:30 | Coffee break |
| 10:30-11:30 | Session 9: Data Transfer and Networking Chair: Paola Grosso (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 10:30-11:00 | Title: Data transfer methods Speaker: Josh Miller, University of Chicago |
| 11:00-11:30 | Title: AmLight SDN Testbeds: The Future of Collaboration Featuring AtlanticWave-SDX Speaker: Heidi L. Morgan, Florida International University |
| 11:30-13:00 | Session 10: Parallel Sessions on Project Development <i>Each site that will host PIRE visiting researchers should have a parallel track 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.</i> |
| 11:30-13:00 | Session 10A: Project Development in AIST Chair: Jason Haga (Location: UvA, B0.207, Science Park 904) |
| 11:30-13:00 | Orientation to research at AIST |
| 11:30-13:00 | Session 10B: Project Development in Sao Paolo Chairs: Fernando Redigolo (Location: UvA, A1.06, Science Park 904) |
| 11:30-13:00 | Orientation to research at USP |
| 11:30-13:00 | Session 10C: Project Development in Edinburgh Chairs: Rosa Filgueira/Malcolm Atkinson (Location: UvA, A1.08, Science Park 904) |
| 11:30-13:00 | Orientation to research at Edinburgh |
| 11:30-13:00 | Session 10D: Project Development in Amsterdam Chairs: Zhiming Zhao (Location: UvA, A1.14, Science Park 904) |
| 11:30-13:00 | Orientation to research at UvA |
| 13:00-14:00 | Lunch (Location: CWI canteen, Science Park 123) |
| 14:00-17:00 | Session 11: Group work- Students can continue work on their challenge Available rooms: UvA, A1.04, A1.14, D1.116, D1.162, Science Park 904 |



DAY 5 (Fri, 12-Jun 2015)

| | |
|-------------|---|
| 09:00-11:00 | Session 12: Challenge presentations and prizes Chair: Cees de Laat/Paola Grosso (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 09:00-09:20 | Title: On his research at UvA Speaker: Alex Moreno, PIRE Fellow |
| 09:20-11:00 | <i>A student from each group will present their work and the jury will decide on the winner.</i> |
| 11:00-11:30 | <i>Coffee break</i> |
| 11:30-13:00 | Session 13: What have we learned & what next? Chair: Heidi Morgan, Paola Grosso & Cees de Laat (Location: CWI, Eulerzaal (Z009), Science Park 123) |
| 11:30-12:30 | 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? |
| 12:30-13:00 | Wrap-up & valediction |
| 13:00-14:00 | <i>Lunch</i> (Location: CWI canteen, Science Park 123) |
| 14:00-19:00 | <i>Excursion (optional)</i> <i>Museumplein – the museum square with the Rijksmuseum, Van Gogh Museum and Stedelijk Museum.</i> |