

## Minutes of WP2 Intro Meeting along the INDIGO kick-off meeting at Bologna

23<sup>rd</sup> April, 10h-13h, Bologna

### Participants:

	Present	Apologies
P0 IFCA-CSIC Santander	Jesus Marco, Fernando Aguilar	
P1 UPV-Valencia		Ignacio Blanquer, Erik Torres
P2 CIRMMP-Uni Firenze	Antonio Rosato	Andrea Giachetti, Enrico Morelli
P3 INAF	Angelo Antonelli	Riccardo Smareglia Cristina Knapic
P4 U.Utrecht	Alexandre Bonvin	
P5 CMCC Lecce	Sandro Fiore	Giovanni Aloisio
P6 ICCU Rome	Sara di Giorgio	
P7 EGI.eu Amsterdam	Peter Solagna	Małgorzata Krakowian, Yin Chen, Sara Coelho
P8 CNR Milan, Rome, Bari	Federico Zambelli, Giuseppe Profitti	
P9 INGV Rome	Paolo Favali, Massimiliano Rossi, Manuela Sbarra	Lucio Badiali
P10 RBI Zagreb	Karolji Skala	

See <https://grid.ifca.es/wiki/INDIGO/WP2/Meetings> for additional documentation

### Introduction

The meeting started with the welcome of Jesús Marco, from CSIC, the institution that coordinates this WP2. Introductory slides were circulated in advance by mail; as they include contact details and photos, they are posted in the wiki now without this info.

Notice that relevant slides about User Communities in INDIGO were presented the previous day in the afternoon session (see link: <https://agenda.infn.it/conferenceTimeTable.py?confId=9460#20150422>)

INDIGO User Community involvement (by Jesus)

EGI-Engage and the EGI ecosystem (by Peter)

INAF use cases: from Astronomical Archives and analysis to big projects like SKA and CTA

Astrophysics (by Cristina and Angelo)

INDIGO: an opportunity for ELIXIR-ITALY to develop services on a solid and open cloud software platform (by Giuseppe)

The proposed agenda was adopted:

0-Introduction of partners and details on their participation in WP2

1-T2.1: Start discussion on requirements

2-T2.2: Research Data Management and Requirements

3-T2.3: Test Use Cases

4-T2.4: Dissemination plan

5-Plans next steps (today and next two months) in particular: D2.1, D2.2, including actions, and agree on next meetings (bi-weekly teleconf and short meeting in Lisbon at EGI.eu conference)

**Notice that we have included some extra info at the end of the minutes.**

## **0-Introduction of partners and details on their participation in WP2**

General information including partners' profiles and members can be found in the WP2 intro slides circulated previously. A brief summary of the confirmation of the participation of the partners was done. We have used a simple counting to identify partners (P0-P10 as there are 11 partners participating in WP2). Notice that all percentages of participation are indicative, but useful for the management and actions in the WP.

### P0- IFCA-CSIC

Jesús Marco (50% time in INDIGO) and Fernando Aguilar (100% in INDIGO). Both only in WP2. Related projects: EGI-Engage Lifewatch Competence Center, COOPEUS.

### P1- UPV

Ignacio Blanquer (contact) and Erik Torres could not attend to this meeting. They have experience in gathering requirements from previous and ongoing projects (OpenCloud Brazil, VENUS-C)

### P2-CIRMMP

Antonio Rosato (contact) (~25% time). His team will include Andrea Giachetti, Enrico Morelli. They have experience in BioMedBridges, WeNMR. Currently involved in MoBRAIN CC.

### P3-INAF

**Rome group:** Angelo Antonelli, working towards CTA, now in ASTRI prototype. They will try to understand if the prototype works within the INDIGO project. His team includes around 10 people.  
**Trieste group:** Riccardo Smareglia will be the contact person, Cristina to be contacted to understand the participation

### P4-U. Utrecht

Alexandre Bonvin will be the contact person. He is **leading the EGI MoBrain CC** also, and will not declare effort in INDIGO. A post-doc (100% in INDIGO for 2 years) will be hired by september.

### P5-CMCC

Sandro Fiore was in WP6, he presented later in the session (see below).

### P6-ICCU

Sara Di Giorgio will be the contact, she is involved in other projects (in particular PARTHENOS). She will help in DMP and preservation issues. Her team is interested in the Cultural Gateway, developed in collaboration with Catania (another INDIGO partner). She is interested also in the "cost" of digital preservation platforms. Davide Madonna will also collaborate.

### P7-EGI.eu

Peter Solagna was in WP3, he joined later the session (see below). Małgorzata Krakowian, Yin Chen (previously working in ENVRI) and Sara Coelho (Communications Manager) will join the team. **See additional info in slides of session.**

### P8-CNR

Federico Zambelli will be the contact, he is working in Elixir, and his expertise includes bioinformatics and big data. Giuseppe Profitti is also in CNR team. CNR considers as a key application the Galaxy workflow, and had an agreement with INFN Bari (another INDIGO partner) for this work.

### P9-INGV

Paolo Favali, leader of EMSO, Massimiliano Rossi, who will be the contact for INDIGO, and

Manuela Sbarra (strong IT background), together with Lucio Badiali (apologies, he joined on Friday) . **See additional info in slides of session, at the end of the INTRO slides.**

#### P10-RBI

Karolj Skala will be the contact person. His team (6 people) has experience in dissemination and is working on the deployment of Galaxy in the Cloud (project with John Hopkins), and leading the EGI DARIAH Competence Center. **See additional info in slides of session.**

### **Brief overview and tasks**

Before starting with the discussion task by task, few points were discussed. We will have an agile methodology approach during the project (5 slides to be sent by Jesus about this).

The first two deliverables (D2.1, D2.2) are due in June, and D2.3 in December, where the test use cases need to be in place. See info in intro slides.

#### **T2.1: Start discussion on requirements**

Some initial comments were made: we need to define our scope, and have in mind the different e-infrastructures (EGI.eu, EUDAT, PRACE) and the other related projects in calls like E-INFRADEV-4, INFRADEV-3 and the recent call on VRE.

Sandro showed us the Data Analytics Modelling language that they have developed. It is a collection of icons, symbols and diagrams that describe graphically a workflow or dataflow (see the attached slides). These diagrams are very helpful for ICT people, but the intermediate step to connect with final users is not direct. See the slides.

Peter gave us an overview of EGI and their experience with communities (in FedCloud there are currently 26 communities). For outreach they have two main channels: direct contact with the communities (requirements gathering, services needed, look for a EGI partner, etc.) and participation in conferences, workshops, etc (see wiki for slides).

FedCloud which is developed on standards, has many successful use cases like Chipster, Happi, INERTIA, etc. Within INDIGO they are ready to support use cases.

Some comments:

- We should profit of the activities in EGI-Engage Competence Centers
- Galaxy is an example of common interest for three of the partners.
- We will set up a wiki for collecting all the information and a potential scheme for requirements
- We need to contact the other WPs

ACTIONS:

*A2.1.1: Develop a questionnaire/template for requirements from user communities (assigned to Peter, Karolj, Fernando). Start from a guide closer to researchers. Initial target: before Lisbon meeting (18<sup>th</sup> may).*

#### **T2.2: Research Data Management and Requirements**

Massimiliano (Max) presented the ideas from INGV (see slides at the end of INTRO).

We may need a survey on the research communities to collect and analyze the individual Data Management Plans (DMP) and data-life-cycle documentation, as well as the use of a Reference Model. We need to define clearly what is under an integrity test is not so obvious, how it related to data ingestion, and the role of validation tests, etc. This is key for preservation, a field where ICCU (Sara) has experience. Indigo guarantee access and data preservation but does not provide storage

itself. From the point of view of INAF (Angelo) there are three important issues: standardization (data format), metadata (well organized, type independent access) and interoperability. Sara also remarked on the importance of semantic layers (some ESFRIs use them) and knowledge bases, and of the integration between contents and repositories.

*ACTION (a posteriori):*

*A2.2.1: Develop a discussion scheme for these issues in the Open Data session at EGI Conference (LW, EMSO and CTA invited to join the meeting and present) (Action on Jesus/Fernando in contact with Paolo/Max in EMSO and Angelo in CTA and confirm them with Lukasz) target; next week*

### **T2.3: Test Use Cases**

Alexandre commented on the definition of the **Test** use cases. All the communities should define clearly their potential test use cases in terms of software needed, data used, computing resources, etc. As D2.3 is for M9, this first step must take place well **before the end of June**.

The complete list of applications included in INDIGO proposal for Test use cases was reviewed and found in principle ok.

An interesting example was pointed out: will it be possible to deploy a complete virtualized HADDOCK portal, supported by a cluster in FedCloud?

*ACTION (a posteriori):*

*A2.3.1: Start compiling the info required for the TEST use cases. Action on Fernando (to trigger the discussion and circulate the initial template provided (on Friday) by Alexandre, see also wiki).*

*Target: before EGI conference*

### **T2.4: Dissemination Plan**

Karolj presented (see slides in wiki) the initial proposal and ideas. He described the PRIDE event, the possibility of activities in training (on Galaxy), etc. The need for a dissemination plan including all partners was indicated.

*ACTION:*

*A2.4.1: Start compiling a list of events of interest for the different partners in relation to INDIGO.*

*Assigned to all partners but see below as this action will be more general within INDIGO.*

Finally regarding T2.5 on sustainability, a potential action A2.5.1 focused on compiling a “state of the art and ongoing initiatives map: ESFRIs, e-infrastructs, etc” was proposed. We could use this map to check where all the use cases are located, where we can focus the dissemination, etc. (Assigned initially to Jesús, but see below on dissemination/exploitation)

**5-Plans next steps (today and next two months) in particular: WP2 management, D2.1, D2.2, including actions, and agree on next meetings (bi-weekly teleconf and short meeting in Lisbon at EGI.eu conference)**

It was proposed and accepted to formally nominate Jesus Marco from CSIC as WP2 coordinator and Peter Solagna from EGI.eu as deputy.

We have a mail list ([indigo-wp2@lists.cnaa.infn.it](mailto:indigo-wp2@lists.cnaa.infn.it)) (**Fernando to check everybody is included**) and an Adobe Connect room (<http://connect.ct.infn.it/data-cloud-wp2/>) that will be used for the bi-weekly online meetings. Meetings will follow the SAMOSAS rules.

The next meeting of WP2 will take place on **May 7th at 11:00**.

We need to deploy a wiki page where all the information can be uploaded

(see <http://grid.ifca.es/wiki/INDIGO/WP2> as a temporal solution)

## **POST-MEETING INFORMATION**

Jesus and Peter were confirmed as WP managers in the INDIGO Collaboration Board meeting on Friday morning.

EGI Meeting in Lisbon (18-22): please confirm to Fernando who will be around and when to define a potential short f2f meeting. Notice that EMSO and LifeWatch will have a slot in the Open Data session on Thursday, CTA is also invited to join (tbc).

Alexandre has sent a first draft of formulary for requirements (see wiki under T2.3).

INDIGO dissemination plans have to be coordinated among T1.4 (leading), T2.4, T2.5 and T3.4. A meeting with INDIGO management in WP1 session with Karolj, Peter, Jesus, agreed to prepare an initial scheme.

The INDIGO deliverable template will be prepared soon under WP1 (Quality assurance task)

**INDIGO industrial partners would like to get involved in WP2, suggestion is to invite them to task 2.5 (Jesus to confirm after checking with INDIGO management)**

Tentative INDIGO calendar (FYI): 2<sup>nd</sup> week of Nov, EGI.eu meeting in Bari

There could be an intermediate F2F meeting by end of June if required.