



GOBIERNO
DE ESPAÑA

MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



Open Science?

Jesus Marco

(IFCA, CSIC-UC, Santander, Spain)

“Las Mañanas IFCA”

Santander 16 Dec 2016

Three Questions

*Make Science more accessible,
efficient and **transparent***

- 1. Why do I talk about Open Science ?*
- 2. Is Science already Open ?*
- 3. What happens now ?*

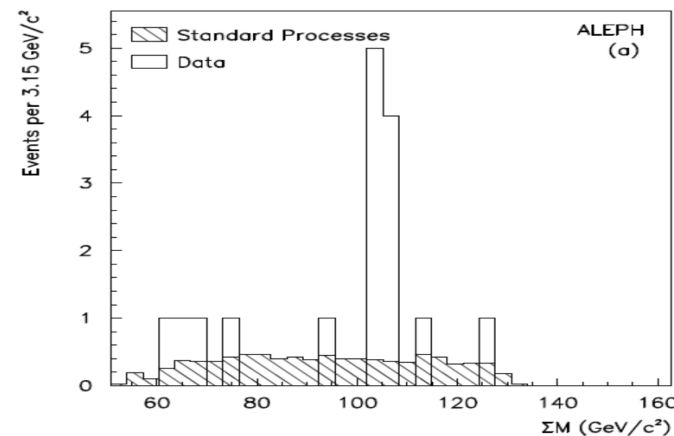
Why do I write these slides in English?... DISCLAIMER

Why do I talk about Open Science?

✱ *We were (are?) so excellent and **open**...*

Story of a plot...

(thanks Celso for reminding me about the author!)



- ✱ *What we did next*
- ✱ *A lesson not learnt*
- ✱ *New hopes thanks to Open Data?*
- ✱ *Yet working on it... 20 years later*

But the “real” world is out there...

Why do I talk about Open Science?

- ✚ *We were (are?) so excellent and open...*
 - ✚ *Physics, Astro **open** repositories*
 - ✚ *...in contrast with other areas (see later)*
 - ✚ *HEP not so open to share data, even after embargo period... nor methods*
- ✚ *LEP experience*
 - ✚ *The four-jet saga (the in-famous plot)*
 - ✚ *We learned to “share” information, but even so...*
 - ✚ *...we had an extended period of data collection*
 - ✚ *What about preserving the data as it was?*
 - ✚ *And the methods?*
 - ✚ *RE-USE (at LHC era) + REPRODUCE (if needed)*

Transparency...versus credit and fame

Why do I talk about Open Science?

LEP experience

-  *Sharing “events”*

-  *Preserving the “hard way” (ask Rafa, Iban, Miguel Angel)*

- *Tapes (data) + Machines (software)*



CERN had no funds in 2001 to preserve 100 TB of “unique” data

Why do I talk about Open Science?

✚ *A new era: LHC*


✚ *CMS open data and long term preservation*

- *Finally, a policy! (thanks Teresa!)*
- *The implementation*
 - *Difficulties: who needs it? who will pay?*
 - *Real life: the Higgs Boson discovery analysis reproducibility*

✚ *From data + software (DPHEP) to **knowledge** (2014)*

- *Not completed (yet! but thanks Kati, Alicia, Ana, Luca, Barbara)*
 - *Validation*
 - *All info: github, wikies, cadl, internal reviews...*
 - *Analysis description*
 - *Ontological approach*

We were ambitious but right: CMS Open Data portal @ IFCA




[ABOUT](#)
[SEARCH](#)
[EDUCATION](#)
[RESEARCH](#)

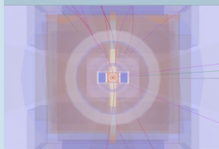
[Home](#) > [Research](#) > [CMS](#)

CMS Open Data are available in the same format as used in analysis by CMS physicists. A CMS-specific analysis framework is needed, and it is provided as a Virtual Machine image with the CMS analysis environment. The data can be accessed directly through the VM image. Basic information of the data contents is provided in [About CMS](#) and in [About CMS Physics Objects](#). The original data are in primary datasets, i.e. no selection nor identification criteria have been applied (apart from the trigger decision), and these have to be applied in the subsequent analysis step. The 2011 data release includes simulated Monte Carlo datasets, but no simulated datasets are provided for the 2010 release.


VMs



Getting started!



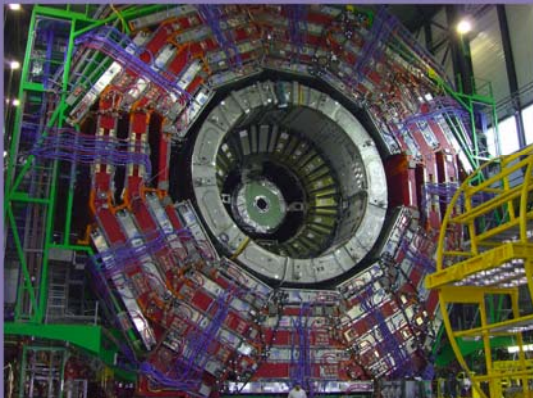
Software and tools



CMS Primary Datasets

CMS primary datasets are AOD (Analysis Object Data) files, which contain the information that is needed for analysis

CMS OPEN DATA @ IFCA.ES BETA



START YOUR ANALYSIS

ABOUT

★

Look to the LHC CMS detector from inside, start analyzing its data.

Instituto de Física de Cantabria provides you with a virtual environment for CMS Open Data analysis for educational use, developed in collaboration with [aeonium](#).

7

Why do I talk about Open Science?

HEP and Open Access

CERN repositories

SCOAP



SCOAP³ – Sponsoring Consortium for Open Access Publishing in Particle Physics

home What is SCOAP³ SCOAP³ Partners SCOAP³ Journals SCOAP³ Repository FAQs Resources Contact Search

SCOAP³ is a global partnership of 3,000 libraries, funding agencies and research institutions from 47 countries and Intergovernmental Organizations.



INSPIRE

Welcome to INSPIRE, the High Energy Physics information system. Please direct questions, comments or concerns to feedback@inspirehep.net

HEP HEPNAMES INSTITUTIONS CONFERENCES JOBS EXPERIMENTS JOURNALS AYUDA

Información Referencias (16) Citas (36) Posters Gráficos

Four jet final state production in e^+e^- collisions at center-of-mass energies of 130-GeV and 136-GeV

ALEPH Collaboration (D. Buskulic et al.) [Mostrar todos los 392 autores](#)

Apr 1996 - 31 pages

Z.Phys. C71 (1996) 179-198
DOI: [10.1007/s002880050163](https://doi.org/10.1007/s002880050163)
CERN-PPE-96-052, CERN-PPE-96-52
Experiment: [CERN-LEP-ALEPH](#)

INSPIRE: [electron positron colliding beams](#) | [electron positron annihilation](#) | [jet multiple production](#) | [final state n multiple production](#) | [mass spectrum \(2jet\)](#) | [jet charge](#) | [angular distribution](#) | [new particle search for new production](#) | [Higgs particle](#) | [sparticle](#) | [quark excited state](#) | [event shape analysis](#) | [ALEPH](#) | [experimental results](#) | [for](#)

015-03-31

Springer Link

Page not found

Sorry, the page you requested is unavailable. The link you requested might be broken, or no longer exist.

SpringerLink is providing researchers with access to millions of scientific documents from Journals, Books, Protocols and Reference works.

Why not start at our homepage?

Log in to check your access to this article

Buy (PDF) EUR 42,29

- Unlimited access to full article
- Instant download (PDF)
- Price includes local sales tax if applicable

[Find out about institutional subscriptions](#)

Export citation

Share article

Zeitschrift für Physik C Particles and Fields

June 1996, Volume 71, Issue 2, pp 179–197

Four-jet final state production in e^+e^- collisions at centre-of-mass energies of 130 and 136 GeV

Authors [Authors and affiliations](#)

ALEPH Collaboration, D. Buskulic, I. De Bonis, D. Decamp, P. Ghez, C. Goy, J. P. Lees, A. Lucotte, M. N. Minard, P. Odier, B. Pietrzyk, M. P. Casado, M. Chmeissani, J. M. Crespo, M. Delfino, [show 378 more](#)

Article

First Online: 01 June 1996
DOI: [10.1007/BF02906976](https://doi.org/10.1007/BF02906976)

Cite this article as:
ALEPH Collaboration, Buskulic, D., De Bonis, I. et al. Z. Phys. C - Particles and Fields (1996) 71: 179. doi:10.1007/BF02906976

1 Citations 77 Downloads

Abstract

The four-jet final state production in e^+e^- collisions at center-of-mass energies of 130 and 136 GeV is studied with the ALEPH detector at LEP. The results are compared with the predictions of the Standard Model and with the predictions of the $q\bar{q}g$ model.

We use cookies to improve your experience with our site. [More information](#) [Accept](#)

20 YEARS LATER...

Is Science already Open?

✚ *A real world is outside!*

✚ *Remote monitoring platform at CdP*

(thanks Maria, Nacho, Jose and ECOHYDROS team!)

- *Can **our** Data be OPEN?*
- *Model of the complete system,*
 - *Meteo data, bathimetry, ...*
- *Software and ... Reference Papers (validation!)*

(thanks Fernando, Dani)

✚ *LifeWatch: an ESFRI for Biodiversity & Ecosystems*

- *COOPEUS, where is all the OPEN DATA???*
- *Legal questions...*

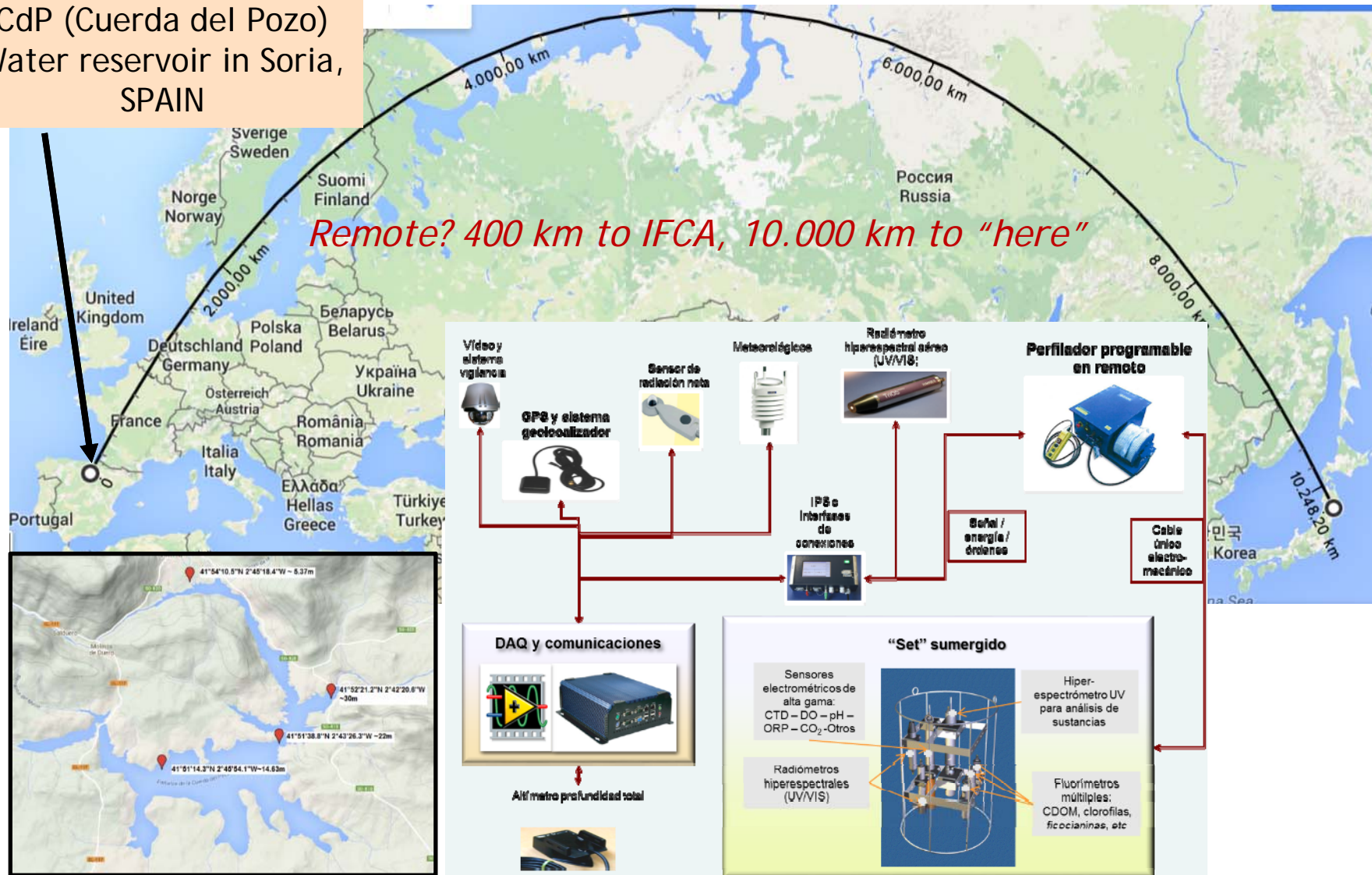
✚ *INDIGO-DataCloud: 12 different communities...*

✚ *Research Data Alliance*

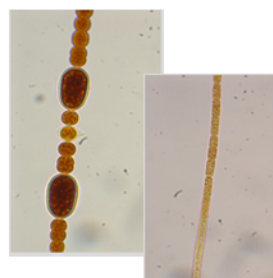
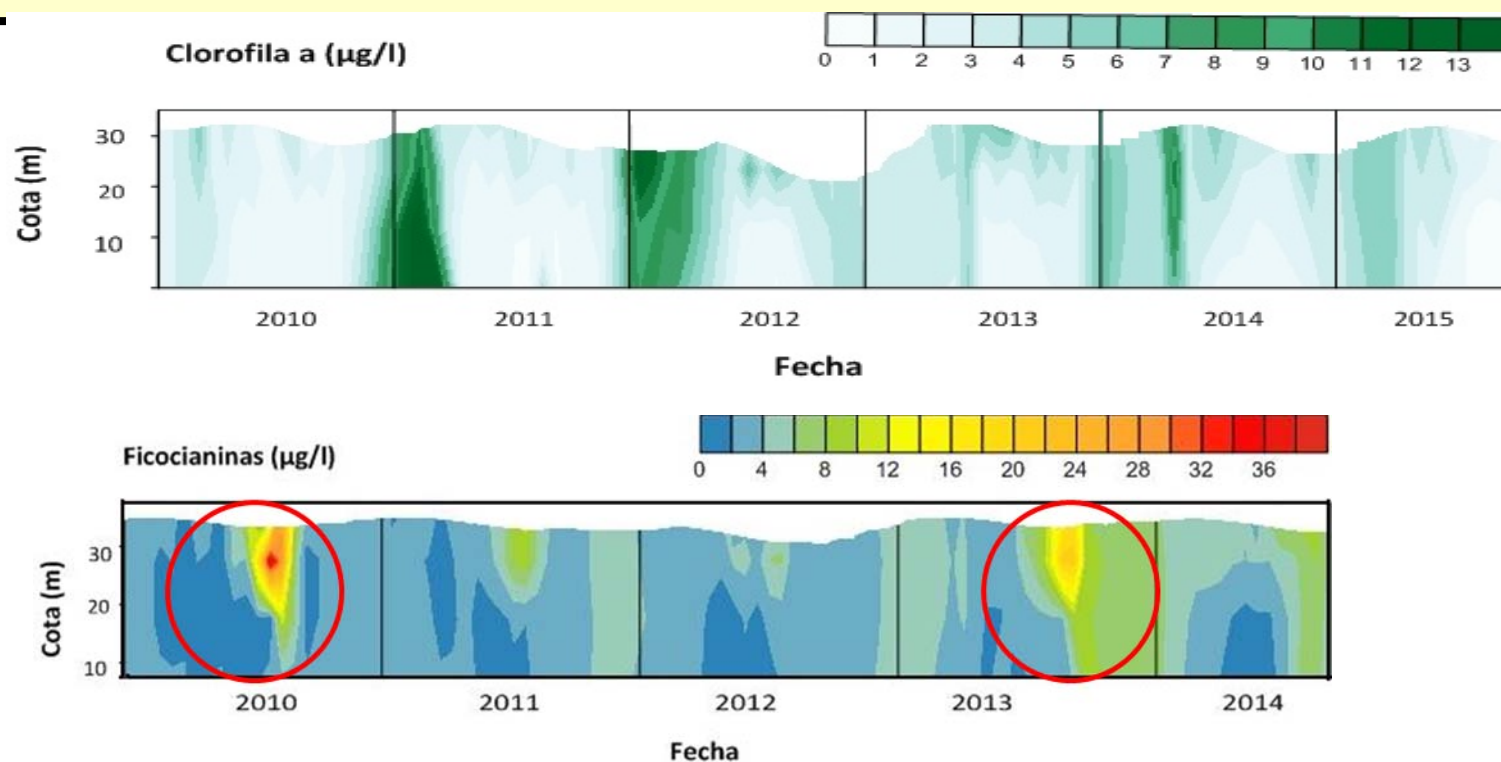
So we decided to implement THE Open Science Framework!

Monitoring a Remote Water Reservoir

CdP (Cuerda del Pozo)
Water reservoir in Soria,
SPAIN

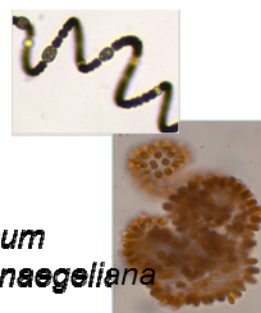


From monitoring to a CHAB warning system



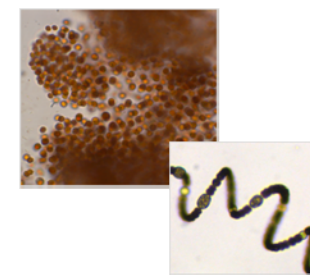
2010 - 2011

Dolichospermum planctonicum
Aphanizomenon flos-aquae



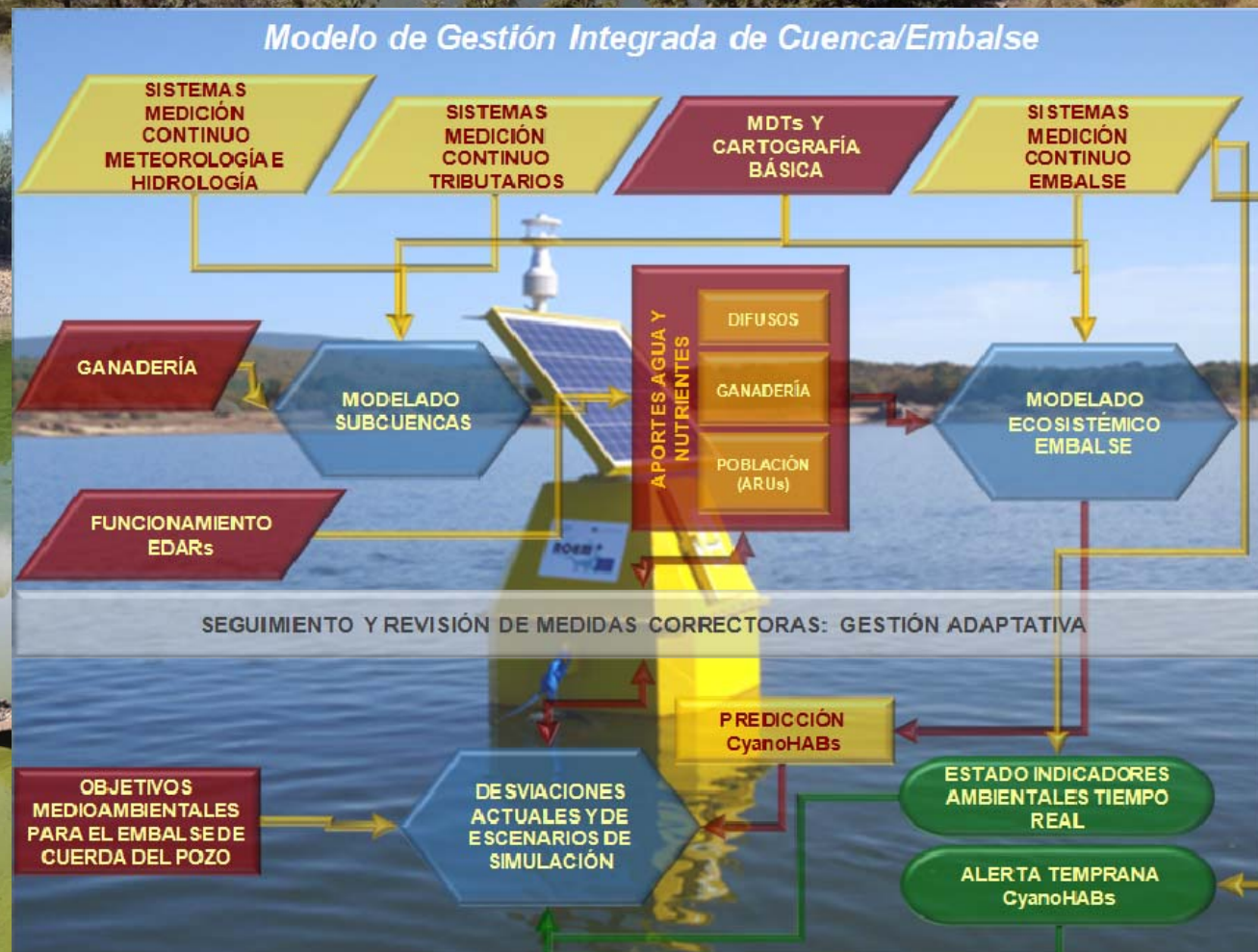
2013

Dolichospermum crassum
Colonias Woronichinia naegeliana



2014

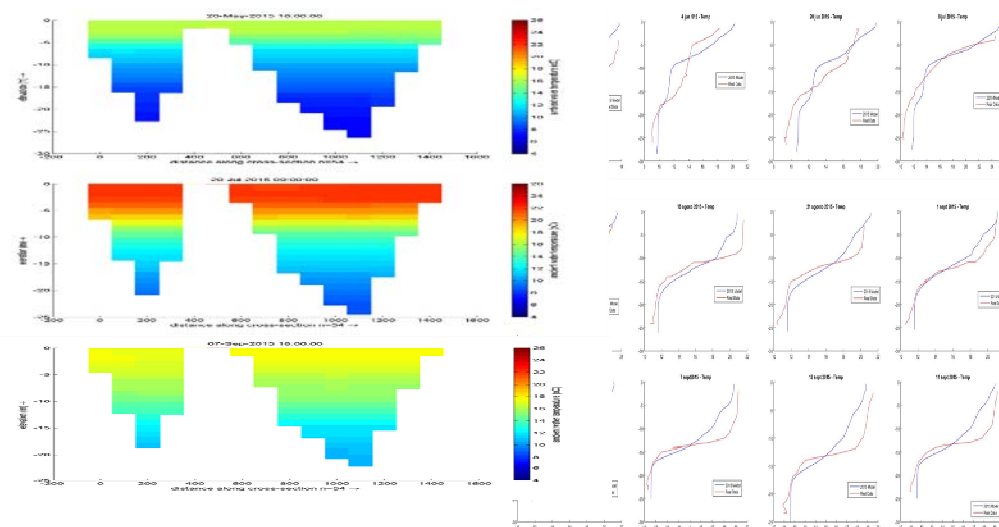
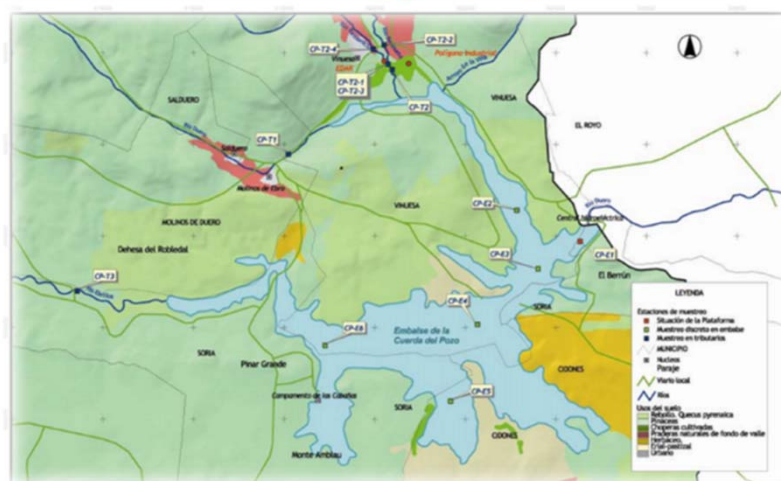
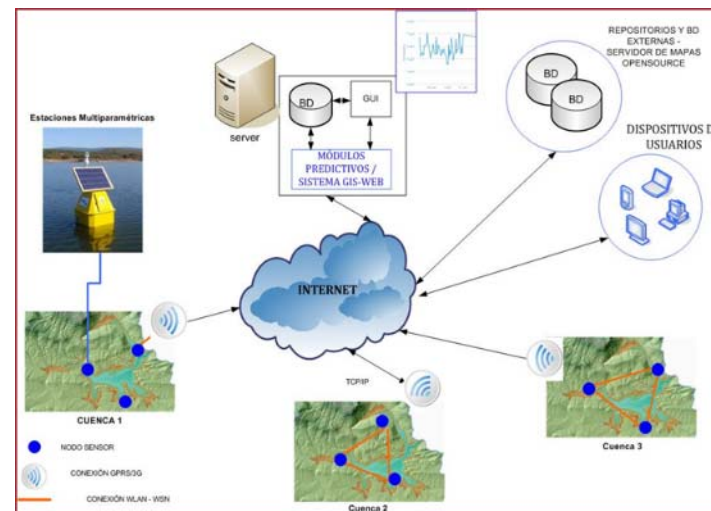
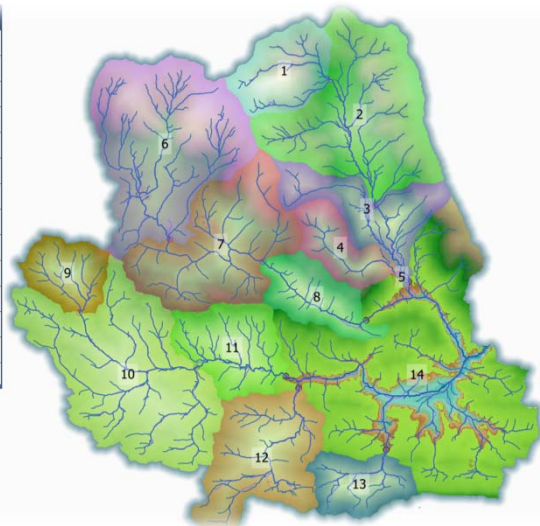
Colonias Microcystis novacekii
Dolichospermum crassum



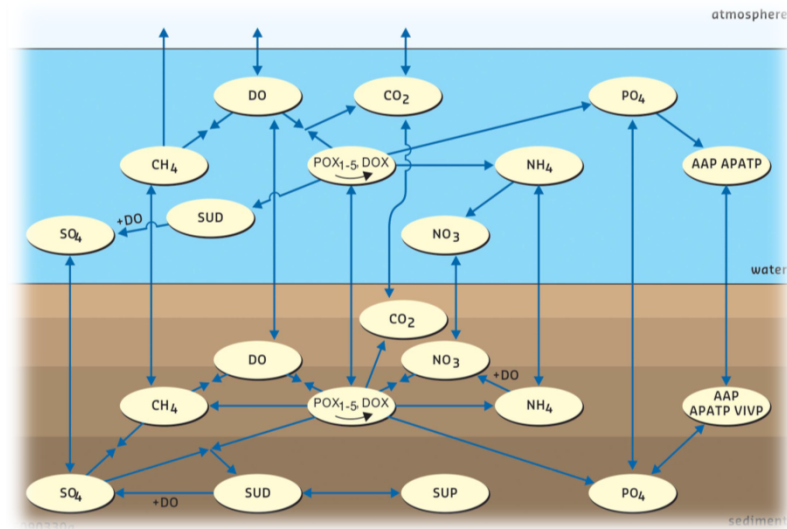
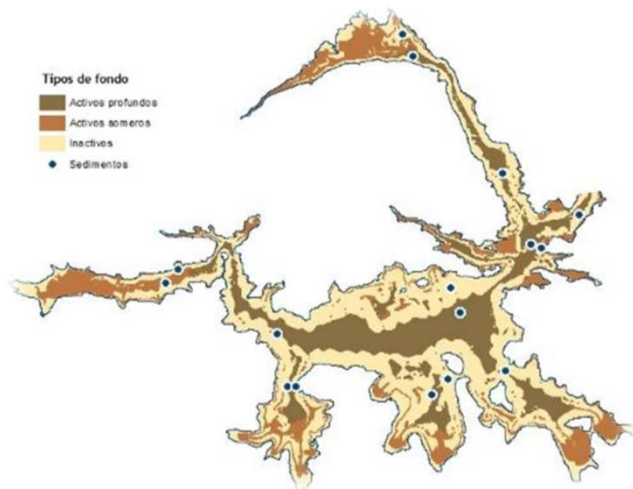
FROM A.MONTEOLIVA PhD THESIS, 2016

(Validated) Hydrological Model: Delft3D

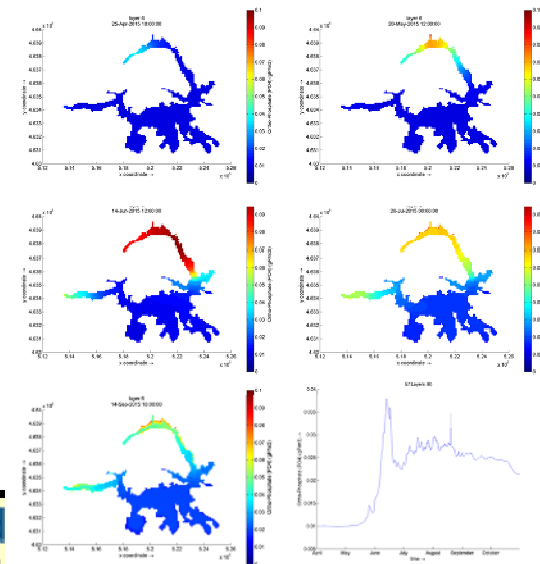
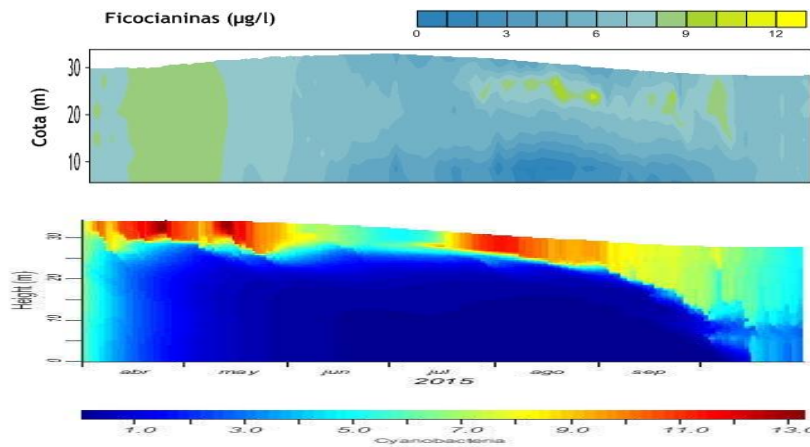
ID	Nombre	AREA (Ha)
1	Revinaesa	2.142
2	Revinaesa	6.003
3	Revinaesa	2.530
4	Remoncio	1.397
5	Remoncio	13
6	Duero	6.720
7	Duero	4.878
8	Duero	1.810
9	Ebrillos	1.569
10	Ebrillos	7.249
11	Ebrillos	2.495
12	Dehesa	3.812
13	Bajero	1.615
14	CED	12.390



BioGeoChemical Models: large number of processes and parameters !!!



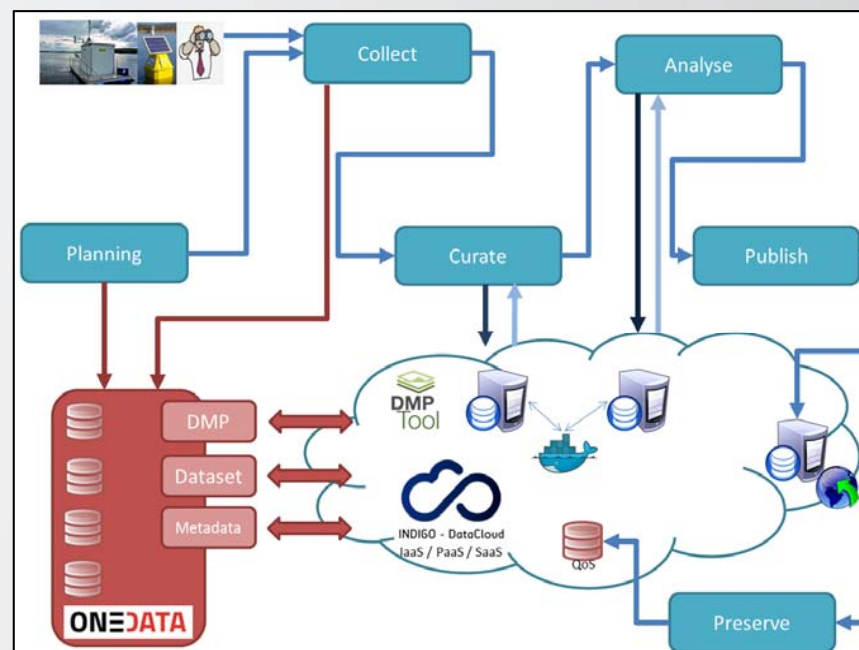
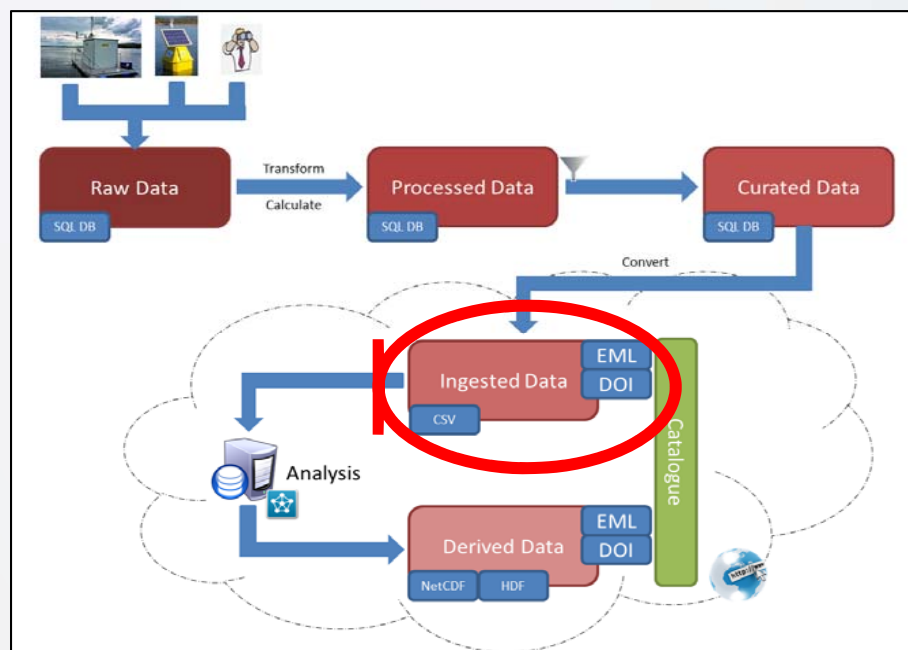
1-D models are not enough, **we need to reproduce the evolution in 3D**



Ingested Data in the Life Cycle scheme



INDIGO - DataCloud



AND the FAIR + R Rules!

Projects follow the Data Life-Cycle: Publish

Plan

Describe and Collect

Curate

Integrate

Analyze

Preserve

Publish

DATA LIFE CYCLE

Publish

Make records available publicly for all users of the framework. Once you select publish your record it will become available for curation in the default community LifeWatch Open Science Framework and in any other you have chosen. If the record is accepted, it will be shortly discoverable for the general public.

Project records

Data Management Plan for Cuerda del Pozo 2015-11-05 Data Management Plan for Cuerda del Pozo

PID lifewatch.openscience/1 | open | dmp | archived

↗ Publish

Processed Physicochemical Water Parameters 2013 2015-11-05 Processed physicochemical water data taken at Cuerda del Pozo Reservoir during 2013. Files of the dataset include depth and temperature for the thermocline analysis

PID lifewatch.openscience/3 | DOI 10.5281/1wdaap.3 | open | dataset | public | archived

↗ Public

Thermocline Analysis Software 2015-11-05 Thermocline parameters calculation. It expects processed data with depth and temperature parameters.

PID lifewatch.openscience/4 | open | software

↗ Publish

Thermocline Demo Analysis 2015-11-05 This is a demo analysis using the processed dataset of Cuerda del Pozo.

PID lifewatch.openscience/5 | open | analysis | public | archived

↗ Public

A framework oriented to Projects

The screenshot shows the 'Projects' section of the LifeWatch Open Science Framework. The header includes the LifeWatch logo and navigation links: SEARCH, COMMUNITIES, PROJECTS (highlighted), DEPOSIT, ANALYZE, ADMIN, and a user icon. A breadcrumb trail shows 'Home > Projects'. The main heading is 'Projects created by LifeWatch Open Science Framework users', accompanied by a 'New Project' button. A search bar with a 'Search' button is present. Two project cards are displayed: 'Cuerda del Pozo' (described as covering reservoir data and analysis, created by admin) and 'Test' (created by admin). The footer contains contact information, funding acknowledgments from the European Commission and Spanish government, logos for the Spanish government and European Union, and mentions of funding from the European Union and support from INVENJO and AEONIU.

LifeWatch Open Science Framework

SEARCH COMMUNITIES **PROJECTS** DEPOSIT ANALYZE ADMIN

Home > Projects

Projects created by LifeWatch Open Science Framework users [New Project](#)

[Search](#)

Cuerda del Pozo [View](#)

This project covers Cuerda del Pozo reservoir's data and analysis.

Created by: admin

Test [View](#)

Created by: admin

LifeWatch Open Science Framework
[About](#) | [API](#) | [Contact](#) | [Privacy Policy](#) | [Terms of Services](#) | [Use Guide](#)

This project has been funded with support from the European Commission


GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

UNIÓN EUROPEA
Fondo Europeo de Desarrollo Regional
Una manera de hacer Europa

Powered by **INVENJO**
Developed by **aeonium.eu**

Thanks Ana, Enol, Aida, David, Fernando, Daniel, and also to
AEONIUM and VIAVANSIS

Key question: Incorporate Digital Knowledge

- ✚ Software (VM) + Data preservation is not enough
 - Ideas explored under CMS preservation
 - Validation
 - Analysis Description
 - CHEP 2013 discussion  Knowledge Preservation
- ✚ Incorporate Digital Knowledge from start
 - Understand the use of “ontologies” / “semantics” (thanks Pablo, Guadalupe, Esther)
 - Ontologies are not taxonomies
 - Ontologies are not metadata
 - Ontologies are not (restricted) vocabularies
- ✚ *Under analysis for Fresh Water VRE:*
 - *Ontological Framework explored: SWEET*
 - *Ontologies: EML and WaterEML*
 - *Consider report at RDA 2015 (Paris) devoted session*
 - *What about INSPIRE directive?*
- ✚ *First try:*
 - *Build on SWEET*
 - *Start from vocabularies used in CdP*
- ✚ Yet, how to integrate into DMP? Ideal (unify) is obvious...

SOLUTIONS EXPLORED

See OPEN DATA Commons session @ RDA Paris

- Support external resources (data, tools): **VRE**
- Enable a “/lifewatch/home” for each researcher/each community, accessible with ID via a **preservation portal**
- Users will define the “openness” of their
 - DATA (private/**embargo**/open/published-DOI)
 - ANALYSIS (R/python, via github)
 - WORKFLOWS at SaaS level (R,python)
- Support it with a global (federated) distributed storage
 - OneData (Data Commons basic component)
- Integrated also with FedCloud computing resources
 - We will rely on INDIGO project developments to optimize!
- Enforce DMP (Data Management Plan)



*If it needs to be preserved => **DMP** & **OPEN** (after embargo)*

Is Science already Open?

- ✚ *Do you know “Science Europe” ?*
 - ✚ *CSIC participation*
 - ✚ *Research Data group*
 - ✚ *Text Mining: A “horrible story” (you don’t want to know)*
 - *Legal questions again...*
 - *Lobbies and Lobbies and more Lobbies*
 - ✚ *Open Access policies*
 - ✚ *Open Repositories need data as well!*
- ✚ *CSIC gets a “policy” proposal!*
 - ✚ *On Open Science!*
- ✚ *But... how to implement it?*

The best business in the world!

Is Science already Open?

✦ *Open Science and Open Access*

Open Access, as defined in the Berlin Declaration,¹ means unrestricted, online access to peer-reviewed, scholarly research papers for reading and productive re-use, not impeded by any financial, organisational, legal or technical barriers. Ideally, the only restriction on use is an obligation to attribute the work to the author.



✦ *Visit OpenAire, meet library experts...*



Launched at the 'Berlin 12' conference in December 2015, the OA2020 initiative aims to accelerate the transition by transforming subscription-based scientific journals to OA business models. OA2020 is based on a financial analysis published by the Max Planck Digital Library. According to the analysis, there should be enough money in the system to allow for a transition to OA at potentially neutral cost.

The OA2020 initiative is outlined in an Expression of Interest statement that was endorsed by 53 parties at the end of June 2016, including SE Mos (SNSF, CSIC, NWO, MPG, Leibniz Association, FCT, DFG and FWF).



[PARTICIPATE](#)
[SEARCH](#)
[MONITOR](#)
[SUPPORT](#)
[OPEN ACCESS](#)

Search in 17,577,022 publications 31,747 datasets from 5,789 repositories and OA journals

EU and Latin America
working together
towards a common
Open Access
implementation



RESEARCHERS

Why Open Access. How to comply. What

DATA PROVIDERS

How to make your content more visible. What

Bienvenidos a **DIGITAL.CSIC**, el repositorio institucional del Consejo Superior de Investigaciones Científicas.

DIGITAL.CSIC organiza, preserva y difunde en acceso abierto los resultados de investigación del CSIC.

Memorias DIGITAL.CSIC



DIGITAL.CSIC
CIENCIA EN ABIERTO

 Envíanos tus trabajos



Noticias destacadas

DIGITAL.CSIC en el Bootcamp de THOR [21/11/2016]

El **proyecto internacional THOR** para la promoción de identificadores persistentes en la comunicación científica organizó el Bootcamp "Tecnología y Servicios para Datos de Investigación" el 17 de noviembre pasado. DIGITAL.CSIC participó en el programa con una **presentación sobre sus servicios a la comunidad científica CSIC productora de datos de investigación**.

DIGITAL.CSIC celebra la Semana Internacional del Acceso Abierto 2016 [24/10/2016]

Este año la Semana Internacional del Acceso Abierto (Octubre 24-30), bajo el lema "Open in Action", está dedicada a medidas concretas que promueven el acceso abierto a los resultados de investigación. **DIGITAL.CSIC ha incorporado nuevos servicios** para que la comunidad científica CSIC pueda poner en práctica el acceso abierto más fácilmente, tanto a sus publicaciones como a los datos de investigación generados durante sus proyectos.

Material del curso de DIGITAL.CSIC sobre datos abiertos [21/10/2016]

Is Science already Open?

Open Science and Research Data Management (RDM)



Funding research data management and related infrastructures

Knowledge Exchange and
Science Europe briefing paper

May 2016



Given the diversity in Europe, a common vision, strategy and funding practice is not easy to accomplish. The increasing shift to an Open Science approach offers a good starting point for the layout of a layered, component-based RDI with complementary RDM support functions at various levels: international/ national/local and mono/inter/multi-disciplinary, offering various types of RDI services (computing, storage, network, data, research support, training and education).

I don't have WhatsApp but... Jesus leaves the group!

🌀 *Science Europe TDM Workshop*

Is Europe Falling Behind in Data Mining? Copyright's Impact on Data Mining in Academic Research

Lucie Guibault, professor of information law at the University of Amsterdam, pointed out that TDM is not directly mentioned in the Directive on Copyright in the Information Society (Infosoc Directive). Nevertheless, the technology is hindered by the directive because TDM regularly involves making copies of the works to be mined. This infringes the reproduction right that is broadly protected by Article 2 of the Infosoc Directive. Copying protected works either needs an appropriate license or an exception within the law.

Professor Guibault focused particularly on Article 5 of the Infosoc Directive, which lists exceptions to the reproduction right and the right of communication to the public. Article 5(1) of the Infosoc Directive allows for broad “transient and incidental reproductions” of copyrighted works. However, transient copies for TDM purposes are only allowed if they are an integral and essential part of a technological process whose sole purpose is to enable (a) a transmission in a network between third parties by an intermediary, or (b) a lawful use. However, a use can be ‘lawful’ if it is authorised by either the rights owner or by law. Since there is no specific provision in the Directive authorising TDM, this means that the article does not provide a guarantee of the right to carry out TDM without the consent of rights holders.

Protecting a very large business... Health, Economy...

What happens now?

CSIC Vision 2020: "A Digital Knowledge Strategy to support Open Science"

"Open Science is a broad term, covering the many exciting developments in how science is becoming more open, accessible, efficient, democratic, and transparent. This Open Science revolution is being driven by new, digital tools for scientific collaboration, experiments and analysis and which make scientific knowledge more easily accessible by professionals and the general public, anywhere, at any time..."

The experience from different research projects and activities lead by CSIC (Spanish National Research Council), indicates that supporting this concept of Open Science is a key first step towards a new way of discovering, sharing and preserving knowledge.

Three key components must be considered under this approach:

A) Open Access to research publications, enabling direct access, without any kind of restriction, registration or subscription.

B) Enhanced Research Data Management, covering the full data cycle, from planning, acquisition and curation to publication, integration in analysis and preservation.

C) Advanced e-Infrastructures enabling the process of large datasets, the mining of scientific databases and literature, as well as the distributed collaboration among researchers at all levels, including the contribution from citizen science.

What happens now?

🌀 Under the vision proposed for 2020:

- Researchers are able to directly explore, access and use research data and publications of different areas when preparing new interdisciplinary studies, employing a well defined framework. They will be able to integrate and analyze the data, using the required computing infrastructure, and also to store and publish the new results including a description of the analysis under a semantic framework so they can be further shared and preserved.*
- Relevant data and analysis results published will be further explored, re-used and referenced by the research community, and proper recognition to their quality and impact attributed to the authors*
- Adequate technical and financial support is provided to these Open Science pillars, including the formation of new specialists and the dissemination of the techniques and results.***
- Citizens are engaged in the support of science, being able to directly explore new results and contribute, when possible, in different ways, from data provision to crowd sourced tasks.*
- The research initiatives launched to target the integration of the semantic framework in the Open Science context provide successful examples of interdisciplinary achievements.*

What happens now?

A) Open Access (OA)

1) Research publications are one of the main results of research process. Both Research Performing and Research Funding institutions share the vision of increasing the impact and reducing the costs of research publications by moving to a system of Open Access

- How to assure that research publications are either published in an Open Access journal or deposited, as soon as possible, in a repository?
- It is crucial to support any valid approach to achieve Open Access goals (green-gold), recognising repositories as a key strategic infrastructure.
- **The hybrid publication model as currently defined and implemented by publishers, is not a working and viable pathway to Open Access. The “double dipping” must be prevented and publishers cost transparency improved.**

2) Open Access is not only about the right of access, but also about the use and re-use information, subject to proper attributions.

- The final goal is to shift to a research publication system in which free access to research publications is guaranteed. This involves a move towards Open Access, replacing the present subscriptions system with other publications models redirecting and reorganising the current resources accordingly.

What happens now?

B) Enhanced Research Data Management (RDM)

1) There is a clear need for a (common EU) policy for RDM activities

- The framework must establish what structure must be used to assure an effective organization of RDM activities, which responsibilities in the data cycle should be defined and how to assure that curation activities are close to the required expertise.

*2) How to convince all actors (RFO, RPO, research teams) of the importance of RDM activities and **open data reuse and exploitation**?*

- By defining institutional policies, and enforcing them, for example new indicators for assessment exercises.
- Data licensing issues should be carefully considered to guarantee proper attribution (following for instance open source software licensing experience)

3) Definition of the scope of the RDM activities, and in particular long term preservation, for the datasets collected/produced in a given project

- In that sense it should be established how to define the interest of the datasets, and associated software and recipes; how to promote that researchers use correctly data embargo and also that open data and metadata formats employed are useful for reuse, and finally how to balance the investment required with this interest on reuse.

What happens now?

C) e-Infrastructures (e-INFRA)

1) e-Infrastructures must be offered and accessed in a unified way as services supporting Open Science

- Assuring a coordination of the different Data, HPC and Distributed Computing/Cloud Computing resources
- Providing Single Sign On mechanisms and tools for management of Virtual Organizations.

2) Virtual Research Environments must be productive for researchers

- Enabling new capacities/capabilities (access to new algorithms, to new resources, simplifying the deployment of new applications)
- Providing a transparent way to share data, analysis and discussions.

What happens now?

Additionally, two clear findings are transversal to these pillars:

INTEGRATION: How to support RDM and OA activities using services on top of e-infrastructures

- Guaranteeing the closeness to institution/experts and guarantee national involvement
- Exploiting an adequate scale factor

FUNDING: How to assure a baseline funding for RDM, OA and e-INFRA

- Considering a formal overhead (3-10%) for any project, depending on the weight of these activities
- By transforming a punctual funding into a long-term budget within the institution

✚ *European
Open Science Cloud*

*New H2020 projects,
~50M€ in 2017-2018*

Complex!

*Research +
e-Infrastructures*

We will be there

(thanks Alvaro, Pablo, Fernando)

Quien le pone el cascabel al gato...



Some “collateral” impacts

✿ *Research Information and Open Science!*



Research Performing Organisations (RPOs) and Research Funding Organisations (RFOs) collect and use data about their own activities from various and heterogeneous sources. This kind of data - data about research activities rather than research data generated by researchers - is stored in research information systems. RPOs and RFOs use research information systems for a variety of different purposes, such as monitoring and evaluating research activities and outputs, allocating funding, supporting decision making on their policies and strategies, tracking researchers' careers, and describing their systemic role to policy-makers, stakeholders and the public. As a result, decision makers and research organisation managers alike increasingly depend on indicators, reports and studies that draw data from research information systems.

Be FAIR! get an ORCID now!

Some “collateral” impacts

- ✪ *Open Data and Open Access open a lot of possibilities...if you have resources*
- ✪ *Example: Deep Learning applied to image recognition (ex: how to become a world expert on plants identification! Ask Ignacio!)*
- ✪ *Is it true that Google and US-GS store whole (all time) Sentinel ESA data (and EU is not able to do it? Ask Javier)*
- ✪ *Who can track all drugs and clinical essays in the market?*

So what can be done???

What happens now?

UIMP-University of Cantabria MASTER IN DATA SCIENCE (2017-2018)

1.- FUNDAMENTALS MODULE (OCTOBER 2017-JANUARY 2018)

This is a compulsory module, and includes five subjects grouped under three topics:

Data Science Panorama	Data Science Methods	Data Management
Introduction to Big Data and Open Science	Statistics for Data Science	Data models and Information systems
	Data Mining	Data Life-cycle: from acquisition to presentation.

2.- SPECIALIZATION MODULE (JANUARY-APRIL 2018)

The student must choose one of the following areas of specialization:

Data Science Analytics	Data Science Engineering	Open Data Management
Machine Learning I	Computer Systems for Big Data	Data Access Services and Portals
Machine Learning II	Cloud for Data Science	Data Preservation

What happens now?

3.- PROFESSIONAL MODULE (APRIL-JUNE 2018)

This module includes the following subjects (compulsory for all students):

Security, Privacy and Legal Aspects
--

New developments in Data Science (based on seminars)

4.- PROFESSIONAL ORIENTATION MODULE (MAY-SEPTEMBER 2018)

The student, according to qualifications and future interest, can opt for external practices and/or “Data Labs” on different areas:

Data Labs		
------------------	--	--

Biomedicine	Environment and Meteorology	Physics and Astronomy
-------------	-----------------------------	-----------------------

Economics and Finance	Internet of Things	Social Sciences
-----------------------	--------------------	-----------------

External practices at selected companies or research groups
--

5.- MASTER THESIS (started in MAY, to be presented by SEPTEMBER 2018)

An advanced work carried out autonomously by the student under the supervision of a professor of the Master. The subject and orientation of this work will depend on the chosen specialty. It will assume a work of initiation to the professional context that will allow you to join a company or a research group.

It may be developed under a three months external remunerated internship in one of the collaborating companies or research groups.