IBERGRID EN EOSC: Despliegue y Servicios

Isabel Campos - IFCA / CSIC
Jorge Gomes - LIP & INCD

CSIC - Spanish National Research Council
LIP – Laboratório de Instrumentação e Partículas
INCD – Infraestrutura Nacional de Computação Distribuída

SOBRE IBERGRID

CAPÍTULO PRIMERO

Que trata de la condición y ejercicio del famoso y valiente hidalgo[I] don Quijote de la Mancha[1]

IBERGRID

IBERGRID es la Infraestructura Ibérica de Computación Distribuida

→ 2006 como parte de los acuerdos Ibéricos de cooperación en materia de e-infraestructuras:

→ Plan Ibérico Común: https://www.ibergrid.eu/institutional-information

Actualización del **Memorando de Entendimiento**: Cumbre Ibérica de Valladolid en **2018**

GOBIERNO DE ESPAÑA — MINISTERIO DE CIENCIA E INNOVACIÓN

FCT Fundação para a Ciência e a Tecnologia

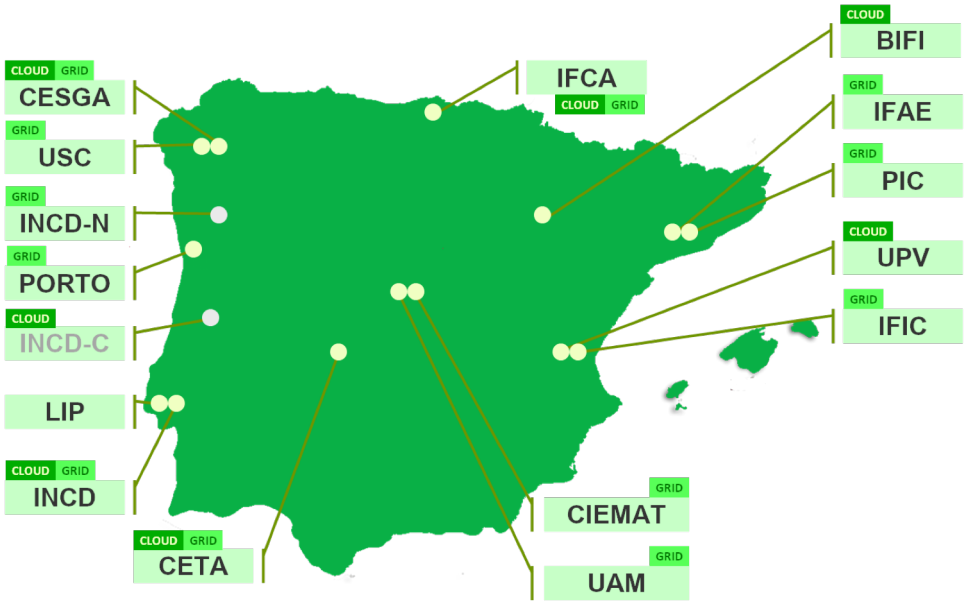
CSIC Consejo Superior de Investigaciones Científicas

INCD

Technical coordination

IFCA Instituto de Física de Cantabria

LIP Computing



Infraestructura de Computación Distribuida

a) Federación de infraestructuras operadas por centros de investigación e instituciones académicas de PT y ES enfocadas a:
   - Computación en Cloud
   - Computación en Grid
   - Procesamiento de datos distribuidos

a) Da soporte a la participación conjunta en la Iniciativa Grid Europea (EGI.eu) y otras iniciativas (EOSC, ESFRIs, LHC Computing Grid (WLCG), etc)
   - Da soporte a nivel de coordinación y de operaciones a las actividades de computación y procesamiento de datos de muchas comunidades científicas

b) Sirve como Forum para dar soporte a actividades comunes, compartir conocimientos, creación de sinergias, etc…
   - Participación conjunta en proyectos EU, incluyendo R&D y proyectos orientados a infraestructuras
   - Series de conferencias y eventos (IBERGRID 2021)

Cumbre Hispano-Portuguesa Valladolid (2018-2024)

Valladolid,
21 de Noviembre de 2018

2. Na área de Computação Distribuída, os Signatários pretendem promover a colaboração nas seguintes áreas:

a) Apoiar a realização anual da conferência IBERGRID, realizada alternadamente no território de ambos os Estados dos Signatários;

b) Reforçar a colaboração entre os Signatários, com vista a otimizar e apoiar a participação da infraestrutura Ibérica IBERGRID em infraestruturas e iniciativas internacionais de computação distribuída e repositórios de dados, entre as quais se destaca o *European Open Science Cloud* (EOSC) e o *European Grid Infrastructure* (EGI);

c) Apoiar o desenvolvimento e integração de serviços temáticos de interesse para a comunidade científica a disponibilizar através da infraestrutura Ibérica IBERGRID;

d) Apoiar e fomentar a utilização da infraestrutura Ibérica IBERGRID no apoio à participação em projetos científicos estratégicos de interesse comum tais como a participação no CERN, o suporte aos ESFRIs e o AIR Center.

Assinado em Valladolid, no dia 21 de novembro de 2018, em dois originais, nas línguas portuguesa e castelhana, sendo ambos os textos igualmente válidos.

Pelo Governo da República Portuguesa	Pelo Governo do Reino de Espanha
O Primeiro-Ministro	**O Presidente do Governo**
Antonio Costa	Pedro Sánchez Pérez-Castejón

EOSC

Servicios para las comunidades Científicas

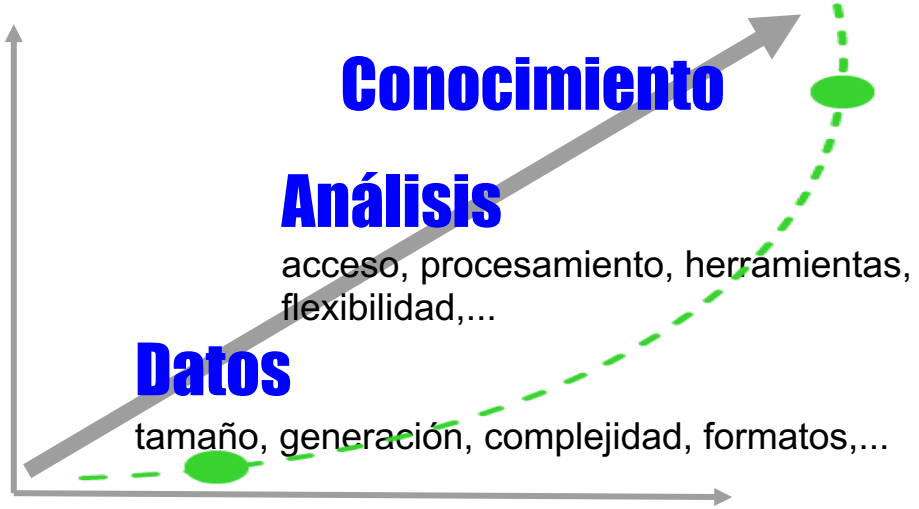
CERN, soporte a ESFRIs, AIR Center

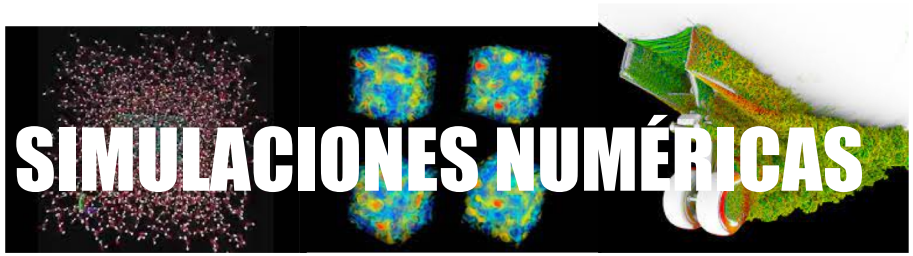
Todo comenzó con el LHC...

CAPÍTULO III

Donde se cuenta la graciosa manera que tuvo don Quijote en armarse caballero[1]

Retos científicos a gran escala
Motivación para evolucionar

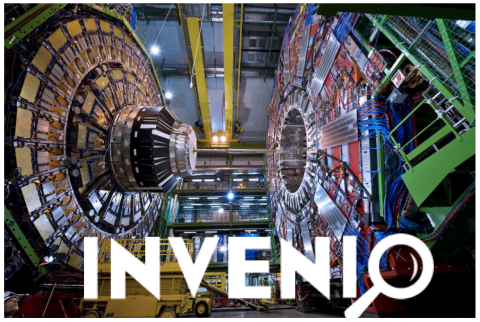




CERN LHC simulaciones y análisis de datos

Computación en GRID

Datos en Física de Altas Energías

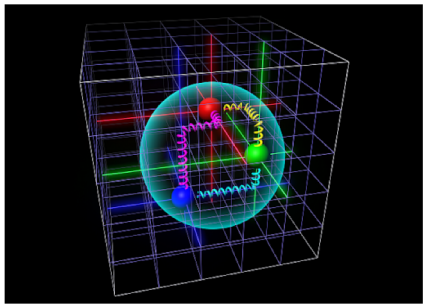


https://inveniosoftware.org

zenodo

- **Instrumentos**- Experimento CMS en el LHC@CERN
  - (Software) Orientado a Objetos (OO) (datos -> objetos, datos y métodos)
  - Software: más cantidad, y más complejo
  - Grandes volúmenes de datos: GRID/WLCG
  - **Se acordó seguir el paradigma de Datos Abiertos**
    - ¿Quién puede aprovecharlo? Datos + Software no son suficientes para extraer el conocimiento tras los datos.
    - Es necesario el soporte de aquellos que generaron los datos. Documentación, pero más.
    - Un enorme apoyo institucional (CERN)

- **Simulaciones FAE**: International Lattice QCD Data Grid
  - ILDG: un esfuerzo (> 12 años) de la comunidad científica de Lattice QCD (esquema de metadatos, formatos de datos, acceso a datos "á la Grid".



- **Tecnicamente**: necesita un esfuerzo de mantenimiento considerable por parte de los científicos y personal técnico de apoyo.
- **Humano**: ¿Por qué debería compartir esto que me ha costado tanto generar)
  Competición por puestos de trabajo, indicadores, etc.

Cooperación: hay vida más allá de nuestra propia área de investigación

CAPÍTULO XXXI

Que trata de muchas y grandes cosas

¿Por qué necesitamos a otros?

- **La cooperación es fundamental**
  - Los problemas que quedan por resolver son sumamente complejos
  - Se requiere un alto nivel de especialización para abordarlos
  - Experiencia que se obtiene sólo después de varios años de trabajo dedicado y enfocado en un problema en particular.

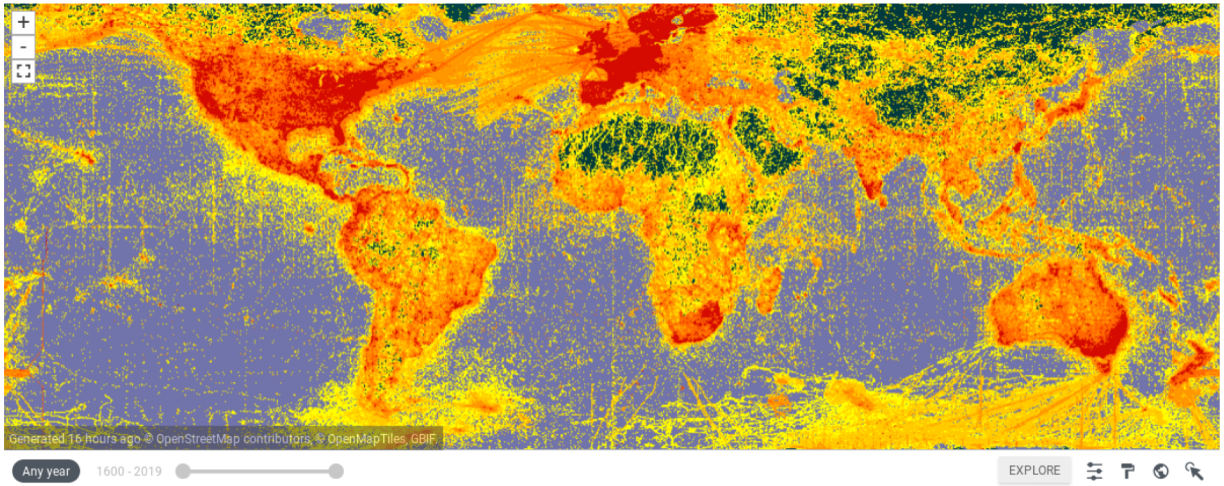
- **Desarrollo conjunto de herramientas y metodologías**
  - También ayudará a resolver desafíos específicos del dominio.
  - Si bien las discusiones técnicas son posibles, la sociología sigue siendo demasiado compleja (a menudo por razones que se reducen a la competencia por financiación).

- **Estudios interdisciplinares: en el área de datos**
  - Impacto climático y planetario, análisis de la biodiversidad,...
  - Regulaciones de privacidad de datos (medicina pero también actividades diarias)

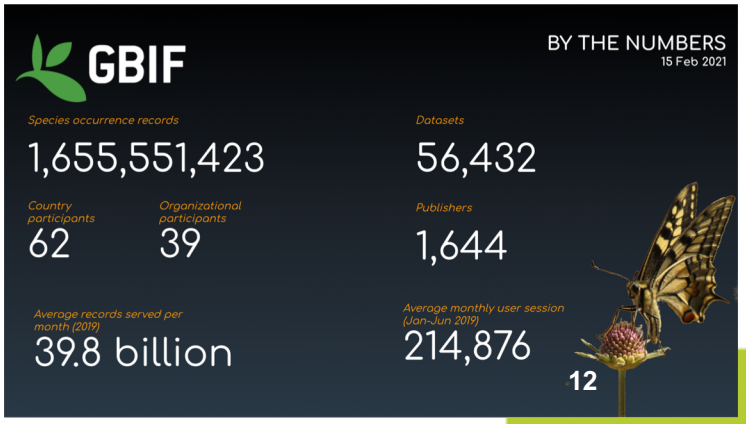
La ciencia debería explotarlos en beneficio de toda la sociedad.

Biodiversidad : GBIF

Transparencias cortesía de GBIF_PT and GBIF_ES



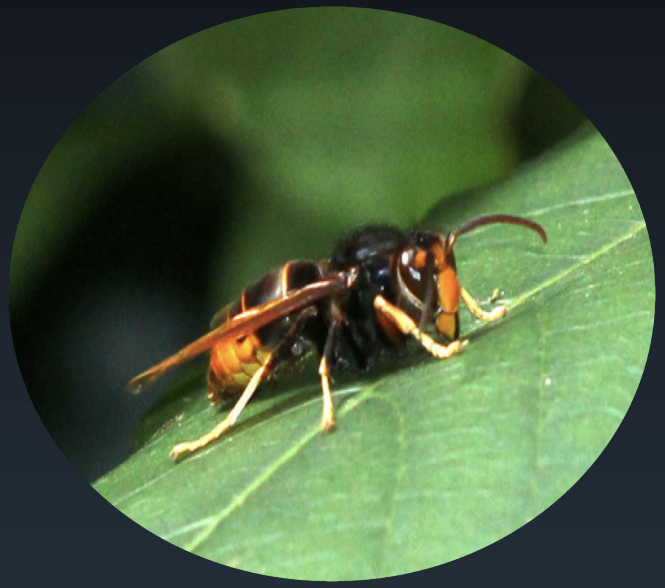
GBIF— Global Biodiversity Information Facility— Red Internacional e infraestructura de investigación financiada a nivel gubernamental, con el objetivo de proveer acceso abierto a datos sobre toda la vida en la Tierra.



Especies Invasoras

Expansión en Europa de la Avispa Asiática (*Vespa velutina*)

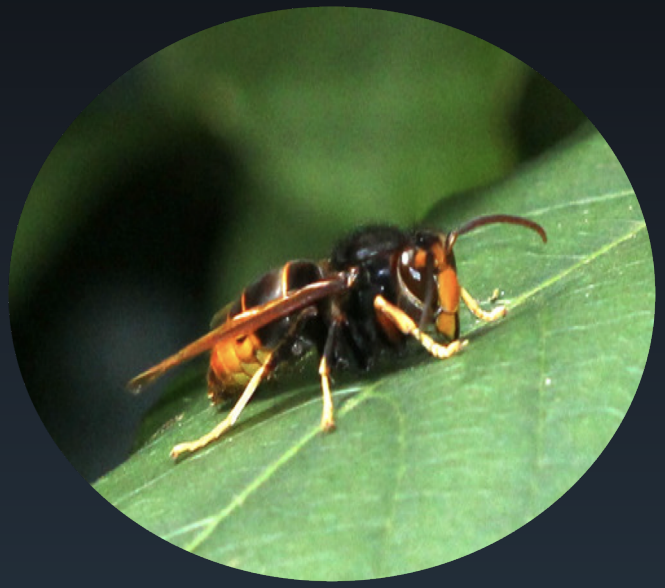
2005



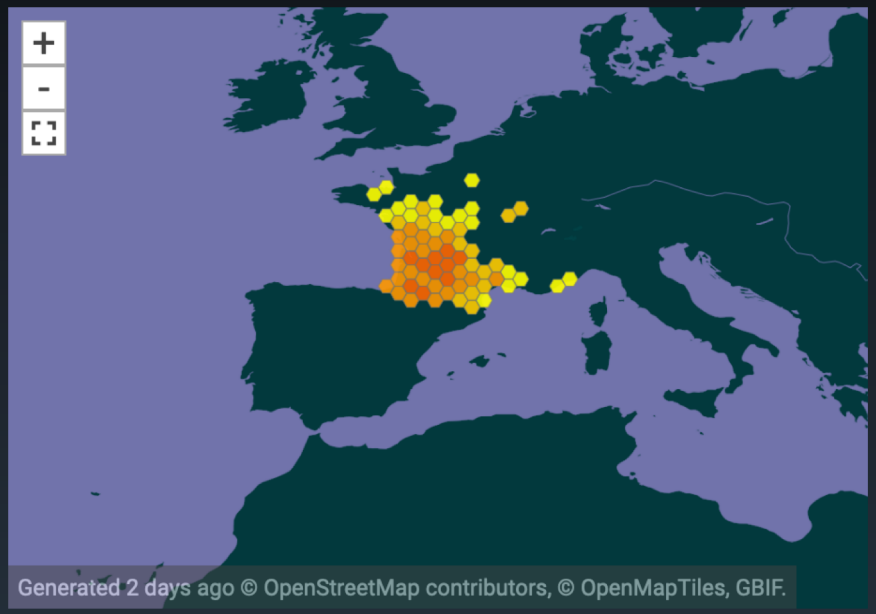
Asian hornets were first introduced to **Europe** when they **arrived** in France in 2004, thought to have been unknowingly transported in cargo. From there they rapidly spread with numerous sightings of the **hornets** across Western **Europe**. Gavin says, 'At the moment we're hoping people notice the nests early enough.

Especies Invasoras

Expansión en Europa de la Avispa Asiática (*Vespa velutina*)

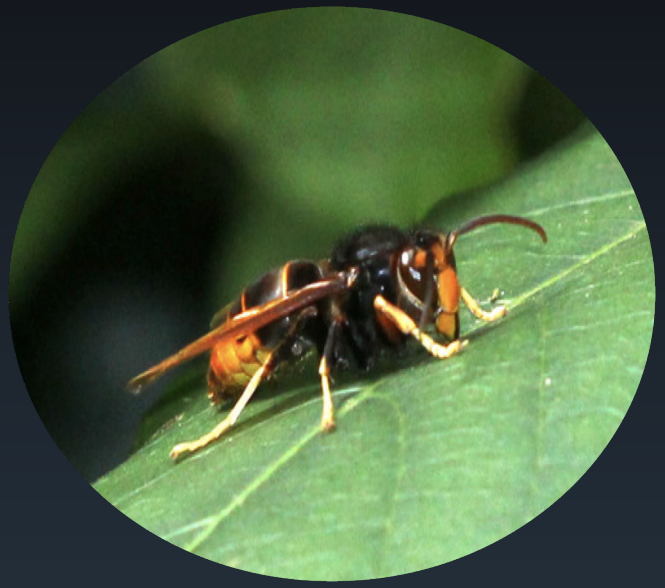


2010

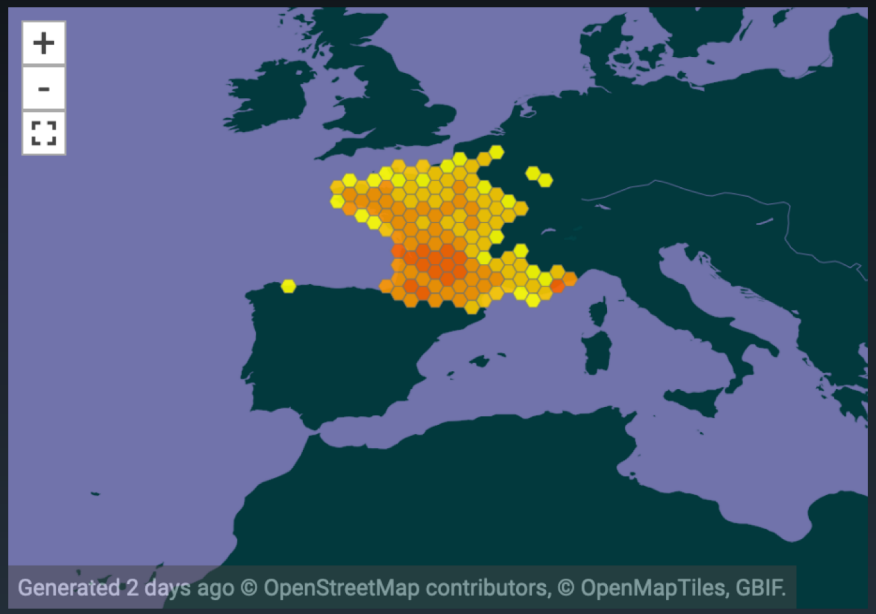


Especies Invasoras

Expansión en Europa de la Avispa Asiática (*Vespa velutina*)

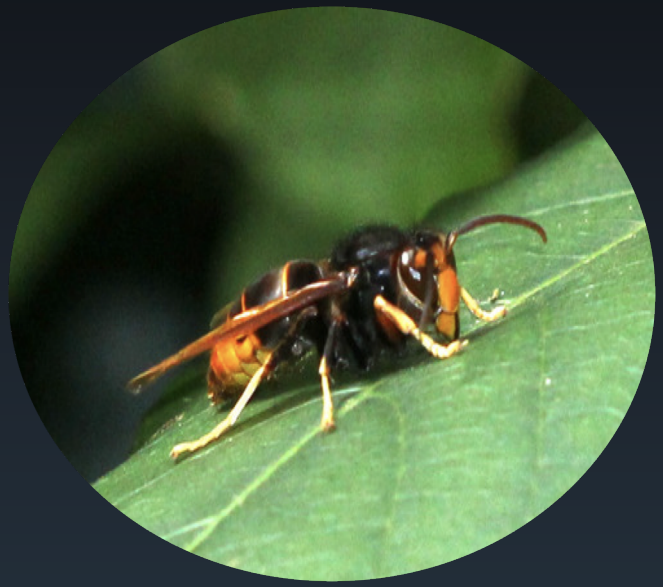


2015

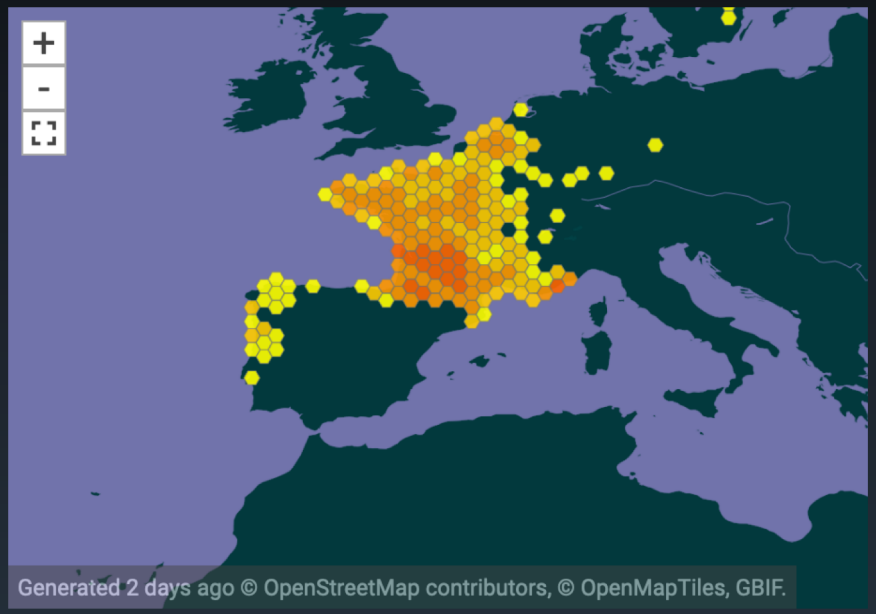


Especies Invasoras

Expansión en Europa de la Avispa Asiática (*Vespa velutina*)



2019



Una avispa velutina, comunmente conocida como avispa asiática, cerca de Santiago de Compostela. EFE

SUCESOS

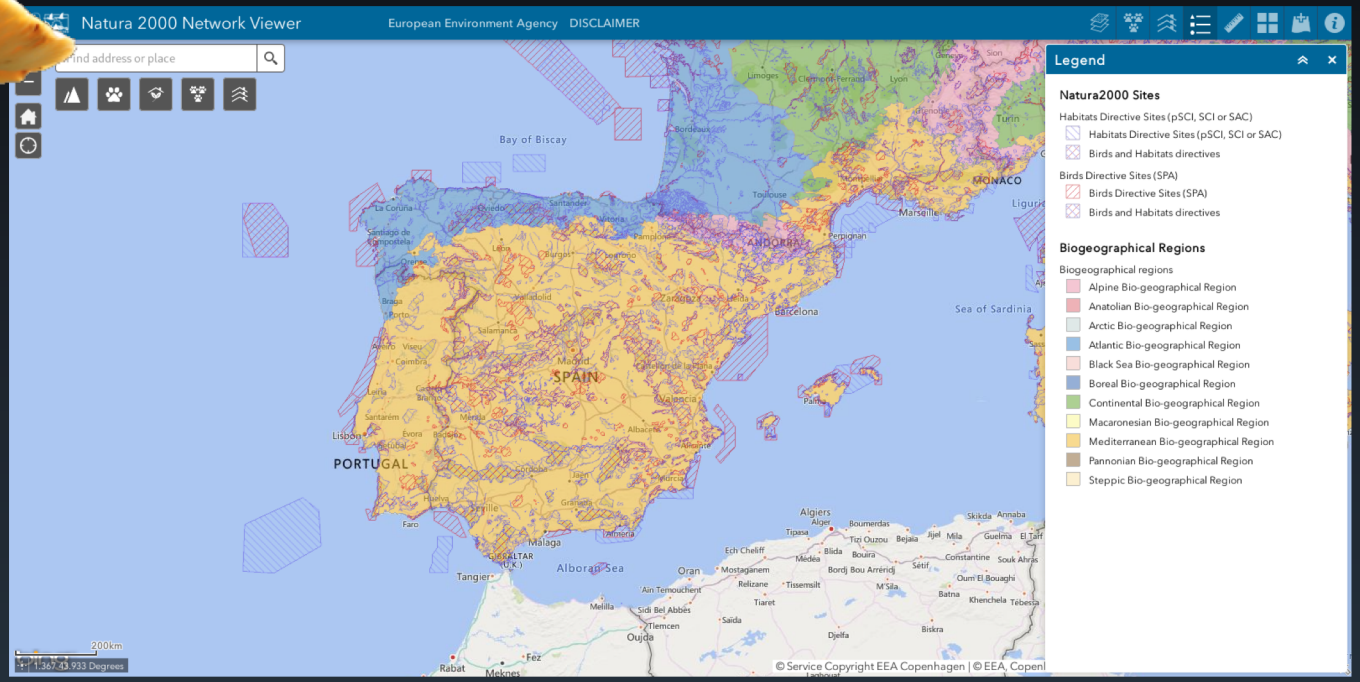
La amenaza de las avispas asiáticas en España: muere un gallego de 54 años por su picadura

El fallecido, aficionado a la apicultura, mientras exploraba sus colmenas observó un enjambre de estas avispas, que le picaron en la ceja.

11 mayo, 2020 - 09:39

Directivas Europeas de Conservación

Las directivas de Habitat y Aves, se hacen a nivel nacional, pero requieren una aproximación regional



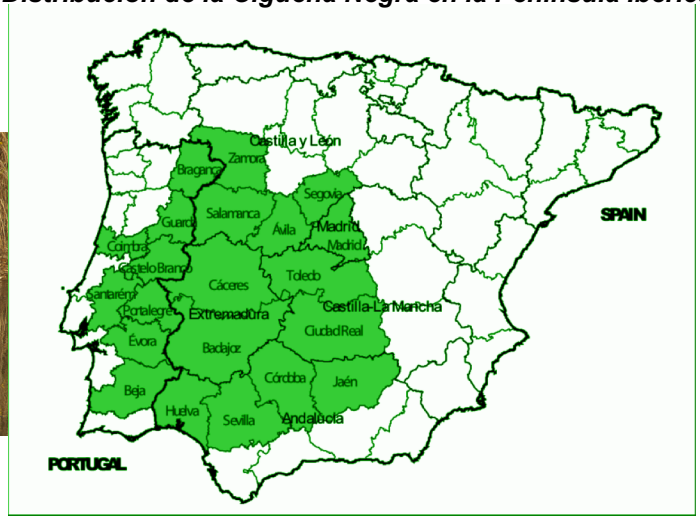
Integración de las bases de datos de GBIF a nivel Ibérico

- **IBERGRID alberga las bases de datos de España y Portugal, por separado:**

https://dados.gbif.pt & https://datos.gbif.es

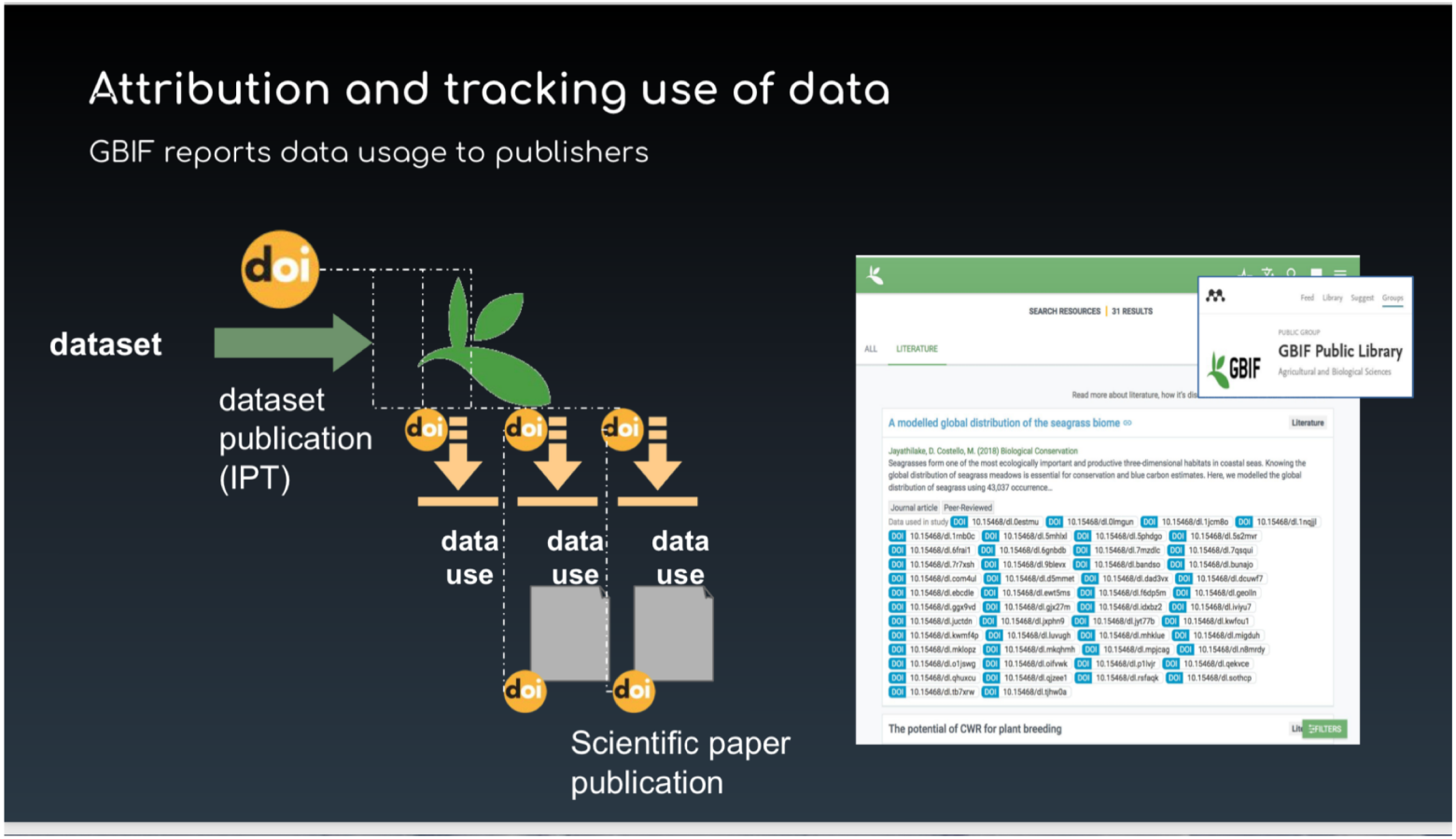
- **Creación de un Espacio de Datos común para GBIF en la península Ibérica, como piloto de la integración de entornos de biodiversidad con características comunes**

Distribución de la Cigüeña Negra en la Península Ibérica



Cigüeña negra cruzando Extremadura, en dirección Alentejo

Hay vida más allá de arXiv.org

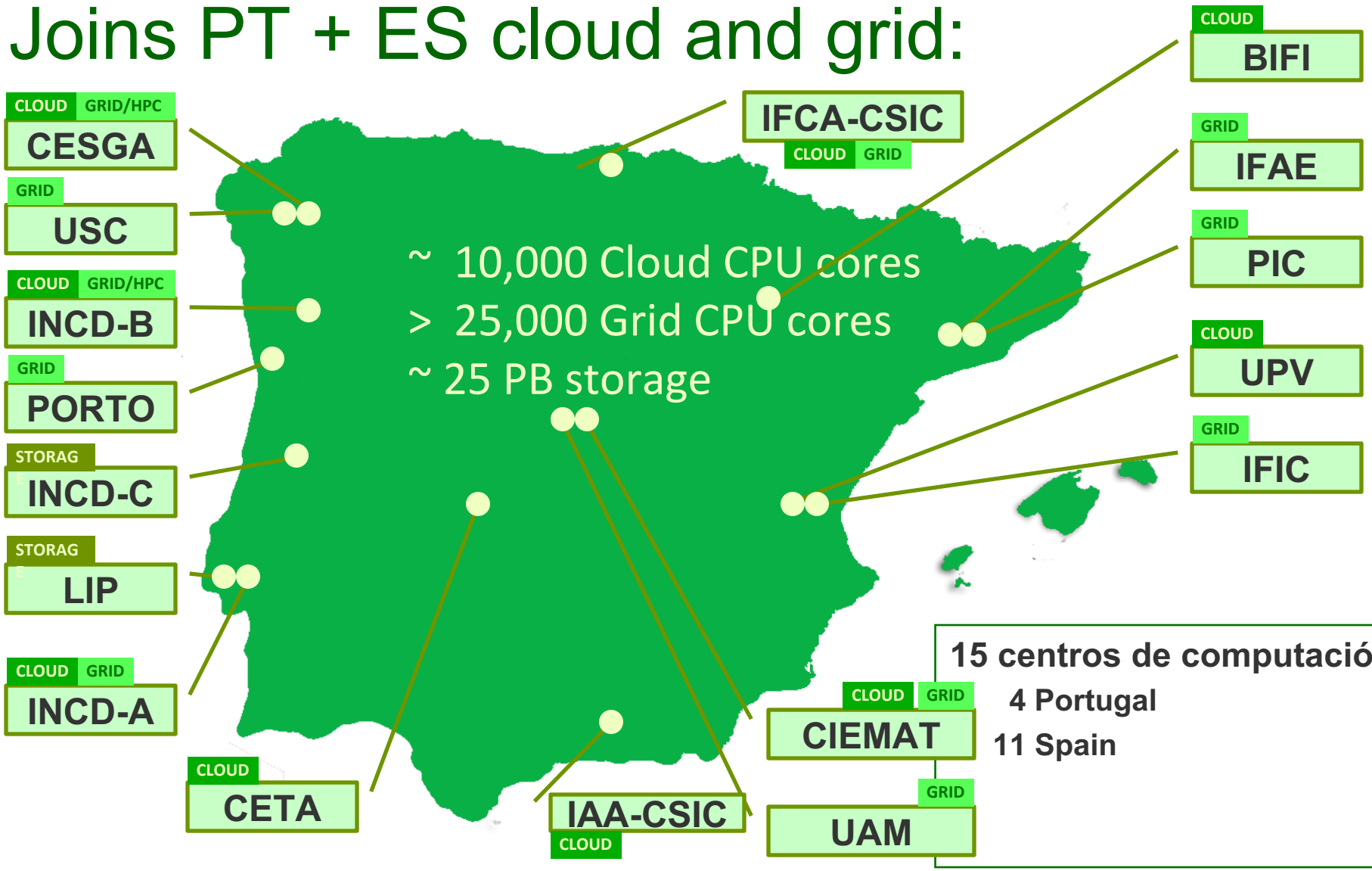


IBERGRID en EOSC: Recursos de Infraestructura

CAPÍTULO VIII

Del buen suceso[1] que el valeroso don Quijote tuvo en la espantable y jamás imaginada aventura de los molinos de viento,[2] con otros sucesos dignos de felice recordación

Joins PT + ES cloud and grid:



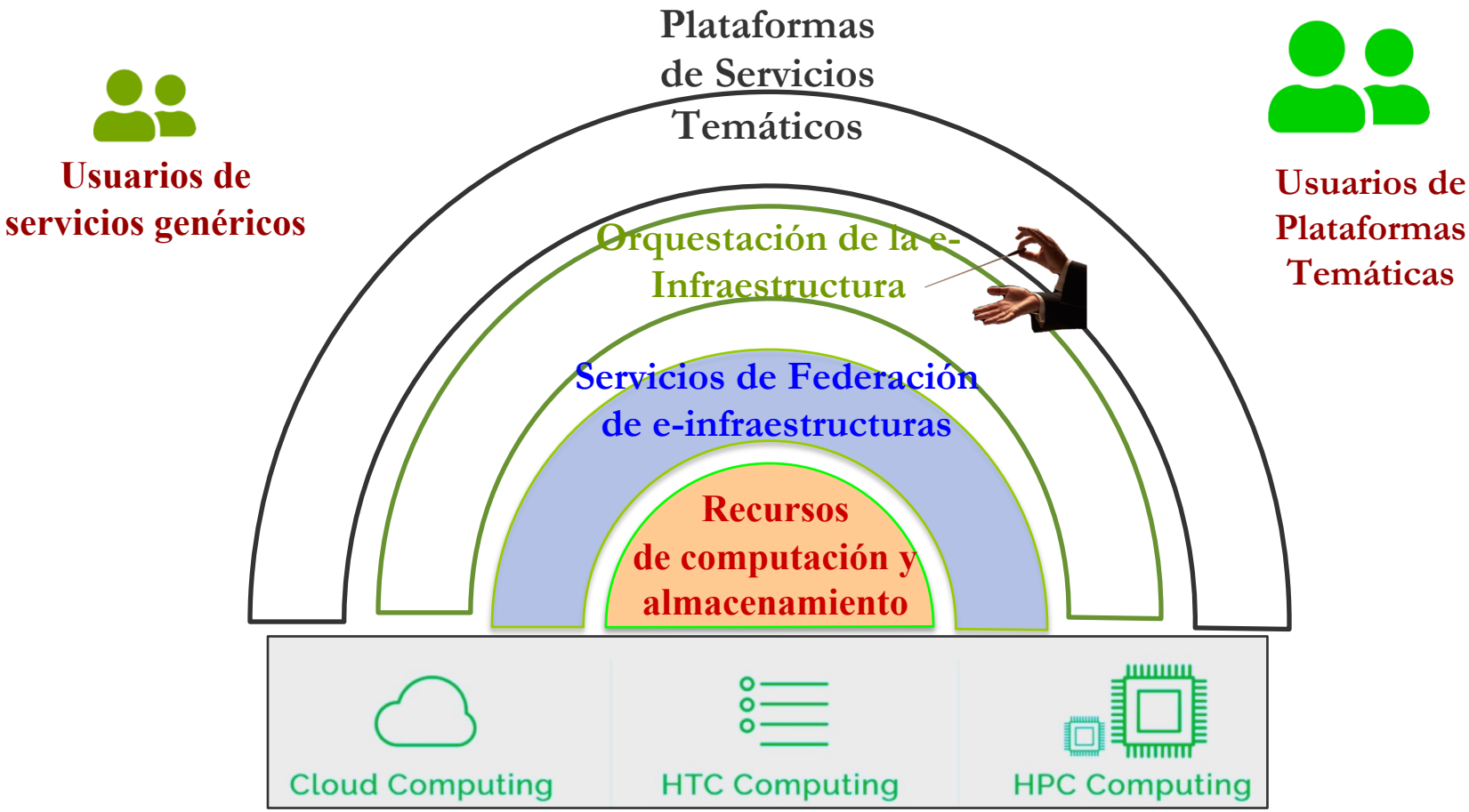
15 centros de computación:

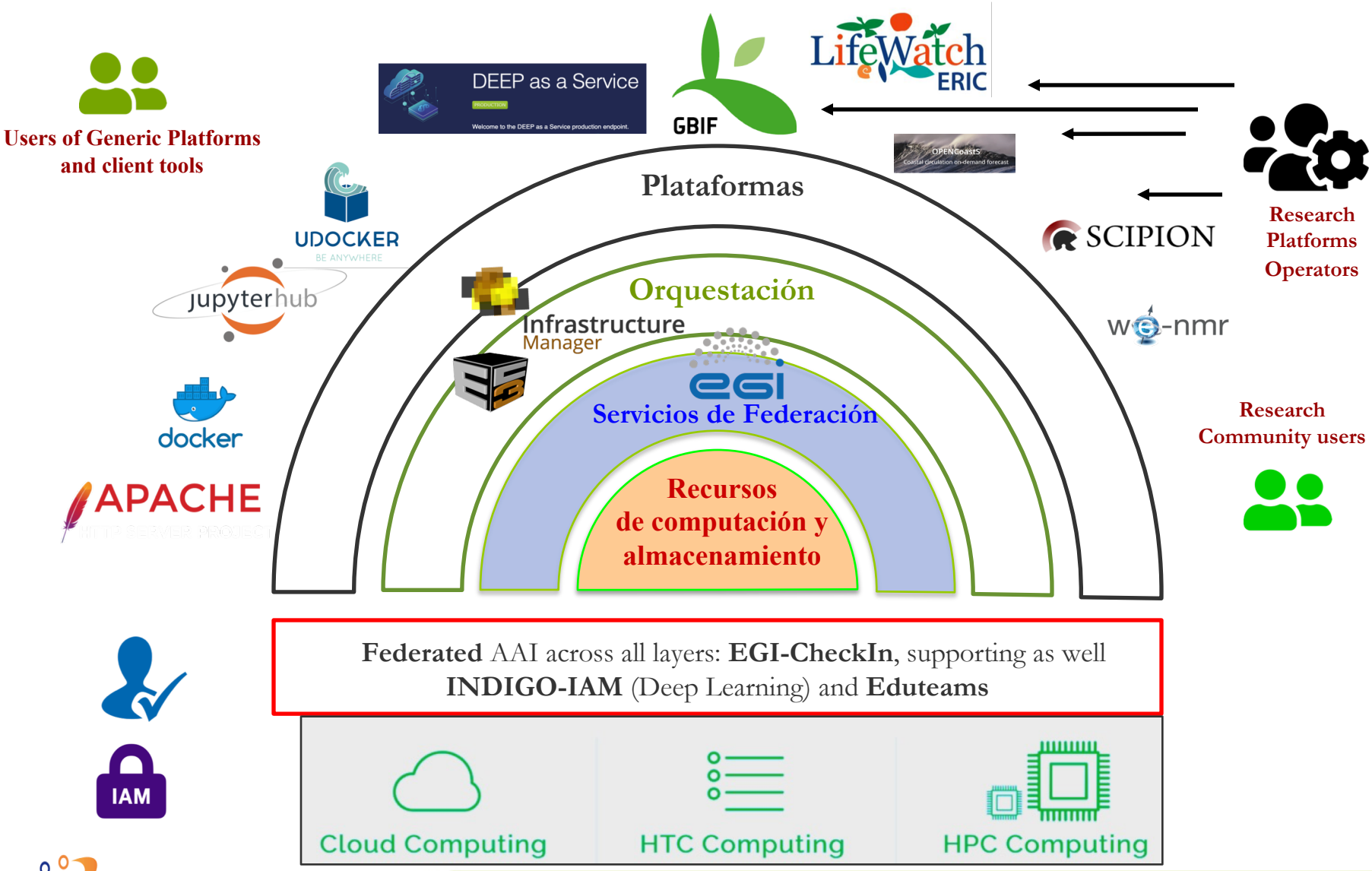
- 4 Portugal
- 11 Spain

IBERGRID en EOSC:
Integración de Servicios Generales y Temáticos

CAPÍTULO XXIIII

Donde se cuentan mil zarandajas tan impertinentes como necesarias al verdadero entendimiento desta grande historia


Usuarios de servicios genéricos
Plataformas de Servicios Temáticos
Usuarios de Plataformas Temáticas
Orquestación de la e-Infraestructura
Servicios de Federación de e-infraestructuras
Recursos de computación y almacenamiento
Cloud Computing
HTC Computing
HPC Computing


Users of Generic Platforms and client tools
DEEP as a Service
Welcome to the DEEP as a Service production endpoint.
GBIF
LifeWatch ERIC
OPENCoasts
Coastal circulation on-demand forecast
Research Platforms Operators
UDOCKER
BE ANYWHERE
jupyterhub
SCIPION
Plataformas
Orquestación
Infrastructure Manager
we-nmr
docker
EGI
Servicios de Federación
Research Community users
APACHE
Recursos de computación y almacenamiento
Federated AAI across all layers: EGI-CheckIn, supporting as well INDIGO-IAM (Deep Learning) and Eduteams
IAM
Cloud Computing
HTC Computing
HPC Computing
eduTEAMS

Servicios generales desarrollados en IBERGRID

UDOCKER

Soporte a la ejecución de contenedores

Infrastructure Manager

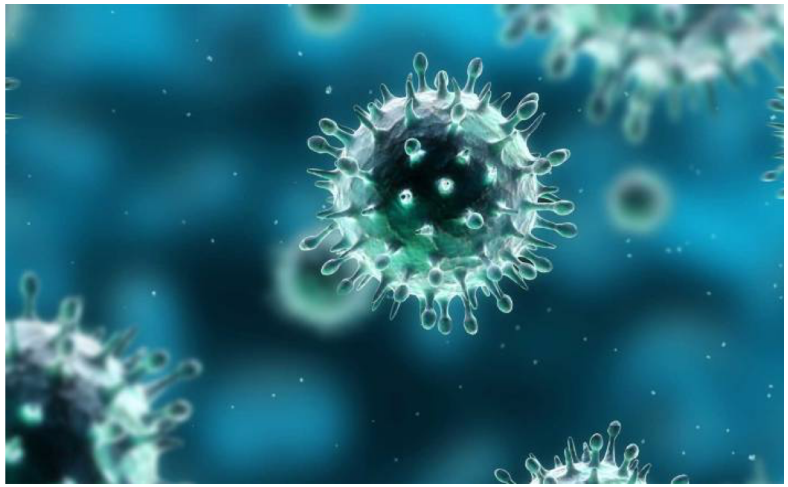
Despliegue de clusters de recursos a la carta

Despliegue de clusters “elásticos”

OsCAR: Open Source Serverless Computing

Repository	User base
github.com/indigo-dc/udocker	“user-level tool” widely expanded in BioInformatics, HEP, Fusion...
github.com/grycap/im	Application architect and application manager Cloud orchestration tool to deploy virtual infrastructures as code, integrated in the EGI Dashboard of the Federated Cloud.
github.com/grycap/ec3	User-level tool for deploying self-managed batch and Kubernetes clusters, used in different scientific disciplines.
github.com/grycap/oscar	Application developer tool for building applications based on functions activated through triggers.

WeNMR



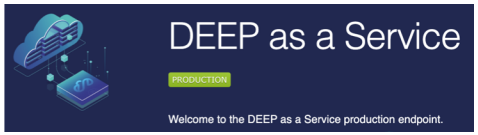
- WeNMR es una plataforma temática dedicada al estudio de la biología estructural.
- Utilizada por la comunidad de usuarios en el **ERIC INSTRUCT**

- Los retos actuales de la era pos-genómica incluyen plataformas que incluyen las herramientas para el análisis y compartición de datos, y el acceso a los recursos necesarios.
- Composiciones de Software muy complejas

965,902 horas de CPU en IBERGRID desde 2012
Cooperación en el desarrollo de herramientas

UDOCKER

Plataformas Temáticas - I



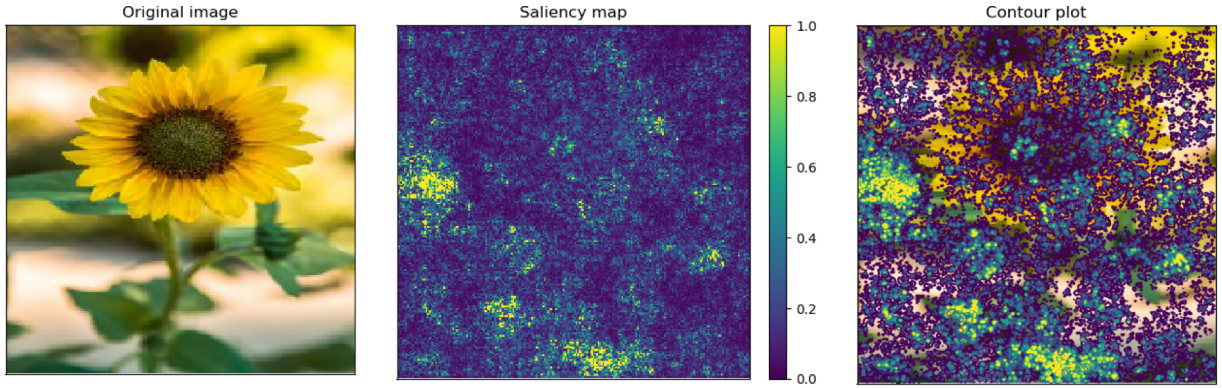
SCIPION

Info & Entry point	User base
https://deepaas.deep-hybrid-datacloud.eu/ https://marketplace.eosc-portal.eu/services/deepaas-training-facility	Deep learning tools to support training, data processing: environmental and earth observation
https://marketplace.eosc-portal.eu/services/opencoasts-portal **Generating on-demand coastal forecasts using EOSC resources: the OPENCoastS service** https://indico.egi.eu/event/5000/sessions/4490/#20201102 by M. Joao Rogeiro, 02/11, 15:45	Coast observation: environment and civil engineering
http://scipion.i2pc.es/ **Towards FAIR CryoEM workflows in EOSC** https://indico.egi.eu/event/5000/contributions/14322/ by Laura del Cano, 02/11, 13:40 - 13:55	Structural Biology: 3D electron microcospy

También sabemos clasificar plantas

- Usando técnicas de Deep Learning
- Se ejecuta en los recursos de Ibergrid, integrado en el Cloud Federado de EGI, usando GPGPUs
- Desarrollos del proyecto DEEP-Hybrid-DataCloud

http://deep.ifca.es/#plants



Helianthus annuus

5,178,562 horas de CPU en IBERGRID desde 2019

OPENCoastS

- OPENCoastS – Servicio para proveer de un sistema de previsión y vigilancia costera en la costa Atlántica Norte de Portugal/Galicia

- OPENCoastS: nivel de agua, parámetros de las olas, etc… hasta 72 horas, basado en simulaciones numéricas.

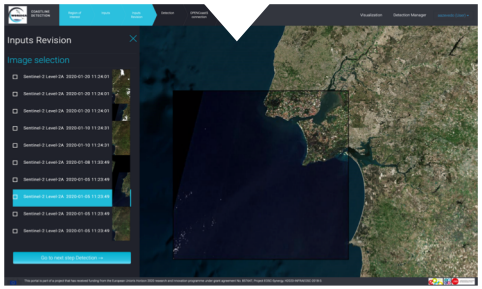
Servicio Temático en EOSC soportado por IBERGRID (2,762,426 horas) desde 2019
LIP, LNEC, INCD, UNICAN, CNRS, CSIC

OpenCoastS y WORSICA

WORSICA

Plataforma Cloud de Monitorization de Agua

Detección de niveles de inundación, fugas de agua en canales naturales y artificiales, riesgo de Tsunami, etc... detection on irrigation networks.



LIP

LNEC
LABORATÓRIO NACIONAL DE ENGENHARIA CIVIL

EOSC SYNERGY

https://www.youtube.com/watch?v=RVCxjhpsIVs

Conexión con Latinoamérica

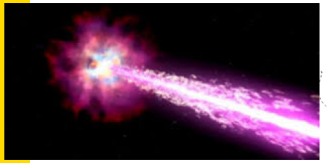
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

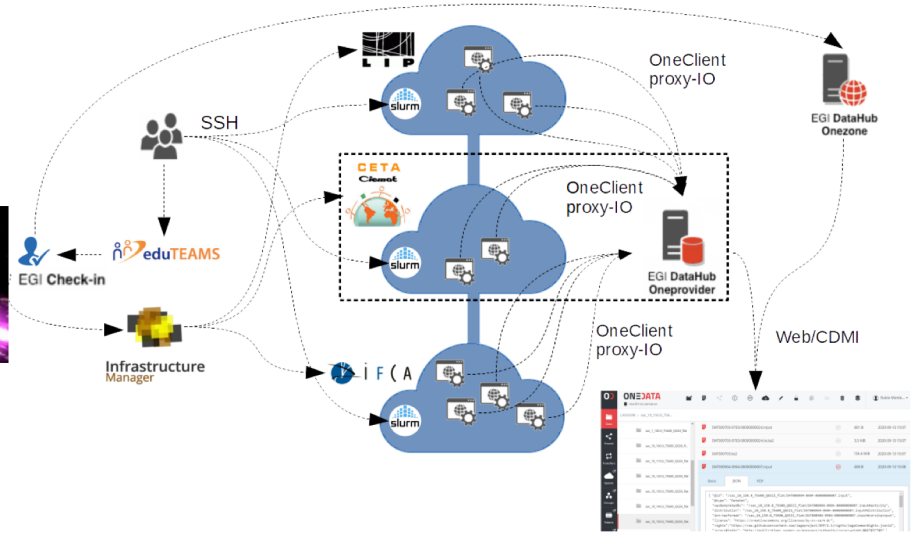
LAGO

http://lagoproject.net/

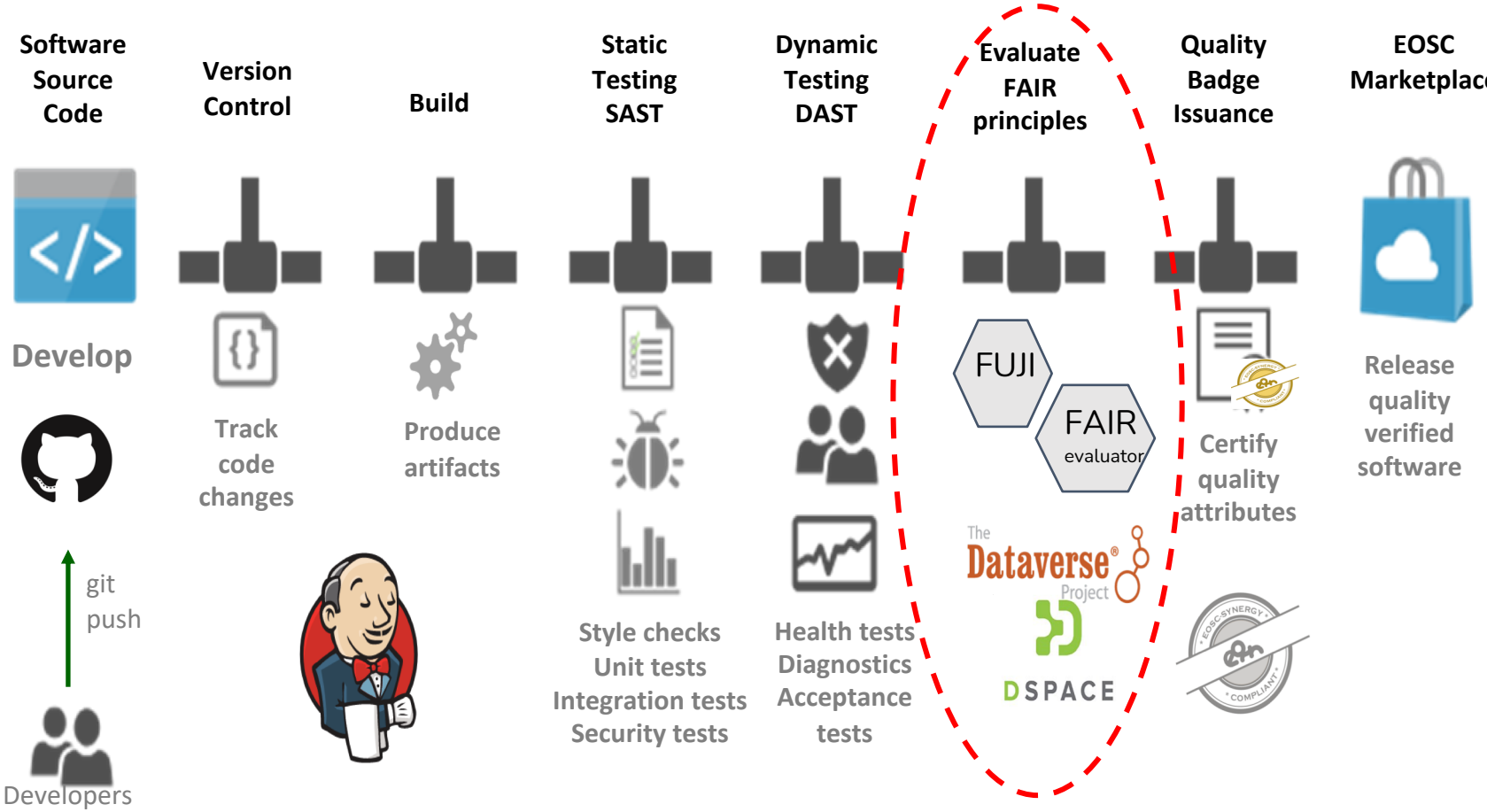
Latin American Giant cosmic ray Observatory - LAGO

LAGO es un observatorio de rayos cósmicos especializado en investigar su influencia en el diseño de materiales





¿Dónde se guardan los datos?



Observatorio AUGER

- El Observatorio Pierre Auger, en la Pampa Argentina, estudia las partículas más energéticas del Universo, en los llamados “Rayos Cósmicos”

- Los rayos cósmicos a energías extremadamente altas siguen siendo un misterio en Física de Partículas. El observatorio Pierre Auge tiene como objetivo entender los enigmas del origen del Universo.

12,714,524 horas de CPU en IBERGRID desde 2009

Zonas Semiáridas en Brasil

SAPS (SEB Automated Processing Service)

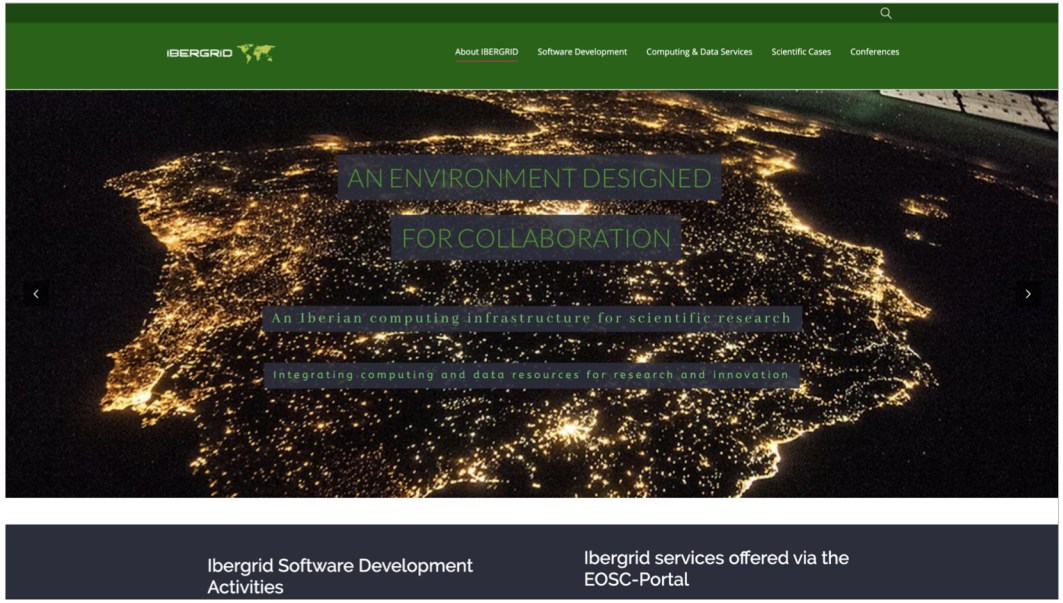
Estimación de la Evotranspiración y otros datos ambientales que se aplican para estudiar la evolución de las masas forestales, así como las condiciones en que las especies invasoras proliferan.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

UFCG

https://www.youtube.com/watch?v=RVCxjhpsIVs

https://www.youtube.com/watch?v=mM6xJJRS3Cs

http://www.ibergrid.eu/

¡ Gracias & Obrigado !