

OPEN ACCESS: REDEFINING IMPACT.



Repositorios digitales para la investigación: La experiencia de CETA-CIEMAT

José Miguel Franco Valiente
josemiguel.franco@ciemat.es



CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

TRUJILLO. Una ciudad histórica

trujillo

DEJA QUE COMIENCE LA MAGIA



- Parque Nacional Monfragüe
- Turgalium, Taryala, Trujillo
- Pizarro, Orellana
- 10k habitantes
- 800k visitas anuales



CETA-Ciemat. Una apuesta estratégica

- **Centro Extremeño de Tecnologías Avanzadas**

- MICIN, CIEMAT,
Junta de Extremadura
Ayuntamiento Trujillo



Ayuntamiento Trujillo

Ciemat



GOBIERNO DE EXTREMADURA



- Establecer puentes, *“unir está en nuestras manos”*

- Computación y Ciencia (E-Ciencia)
- Ciencia y Sociedad
- América Latina \leftrightarrow Extremadura \leftrightarrow Ciemat



- **Fondo Europeo de Desarrollo Regional (FEDER)**

- Obra civil
- Infraestructura tecnológica
- Colaboraciones y proyectos I+D



UNIÓN EUROPEA
“Una manera de hacer Europa”



CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Ciemat. El respaldo de un gran centro

- **Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas**

- **Organismo Público de Investigación**
- **Ministerio de Economía y Competitividad**
- **Secretaría de Estado de I+D+i**

- **I+D+i en...**

- **Energía**
- **Medioambiente**
- **Tecnología**
- **Investigación básica**



- **Algunos datos...**

- **Recursos humanos: 1326 personas**
- **Presupuesto (2013): 98,3M €**
- **Ingresos actividades I+D+i (2013): 44,21M €**





CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Ciemat. Centros Territoriales



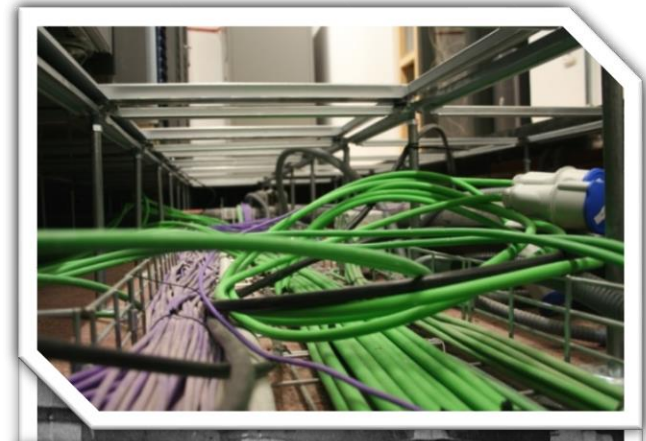


CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

CETA-Ciemat. Instalaciones e infraestructuras

- **Conventual de San Francisco (siglo XVI)**
 - 900m2, oficinas, despachos, salas
 - Vigilancia 24x7
- **Centro de Procesado de Datos**
 - > 1.000 Terabytes (+60 Gbps)
 - > 140 Teraflops (HPC, GPGPU, MPI, NUMA)
 - 3.000 cores CPU; 125.000 cores GPU
- **Conexión a Redes de fibra**
 - Regional (RCTex) y nacional (RedIRIS-NOVA)
 - 2x 10 Gbps
- **Referente nacional e internacional**
 - IBERGRID, España y Portugal
 - SCALAC, América Latina y Caribe
 - Iniciativas Grid y Cloud Europea (EGI, EFC)





CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

CETA-Ciemat y los repositorios digitales: Los comienzos

- **La GRID**
 - Compartición de recursos
 - Acceso seguro
 - Estándares abiertos
 - Organizaciones virtuales (VO)
- **Gran capacidad de cómputo y almacenamiento**
- **Explotación de almacenamiento para alojar repositorios digitales**





CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Digital Repository Infrastructure: El proyecto

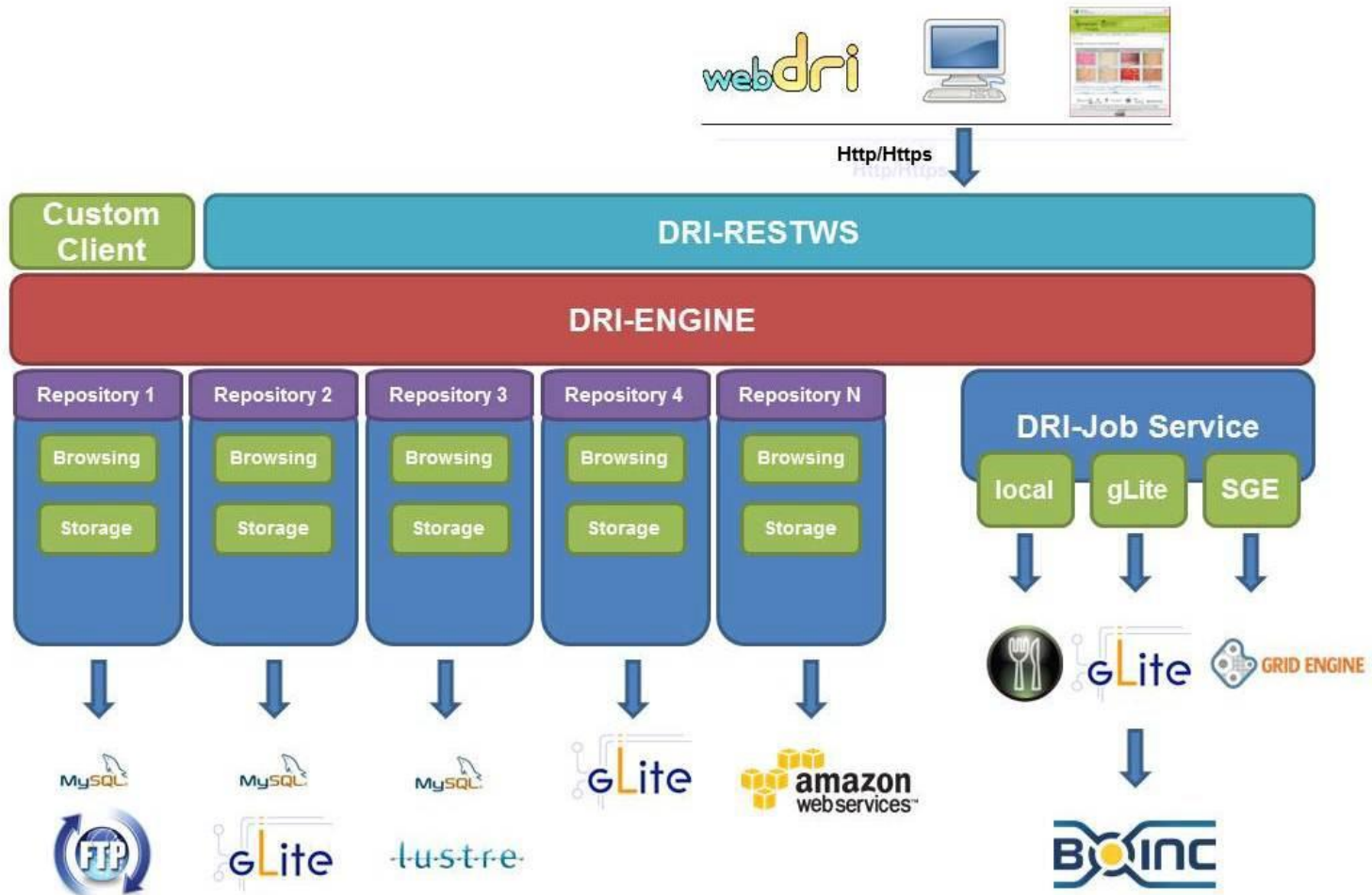
- **Motivación**
 - Alojar repositorios digitales en infraestructuras avanzadas de almacenamiento
 - Proporcionar herramientas para su generación y explotación por los usuarios
- **Colaboradores**
 - INFN-Catania (Italia), INEGI-FEUP (Portugal) y Fundecyt (España)
- **Resultado:** Plataforma DRI
 - Abstrae a los proveedores de los repositorios del almacenamiento subyacente
 - Abstrae a los desarrolladores de los servicios de la infraestructura
 - Proporciona aplicaciones para la navegación y gestión del contenido
 - Inspirado en la especificación JSR-170 (*Content Repository API for Java*)



CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Digital Repository Infrastructure: Arquitectura



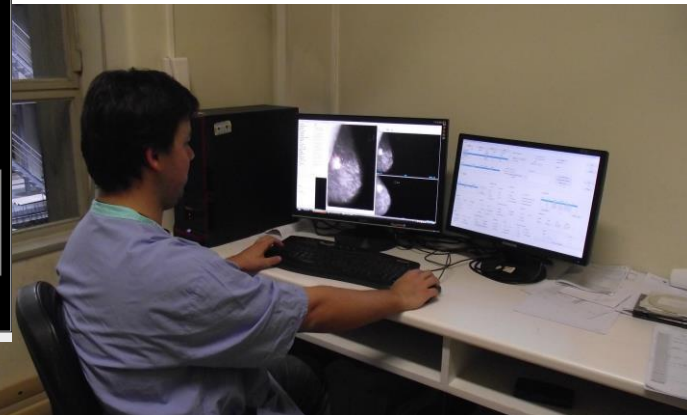
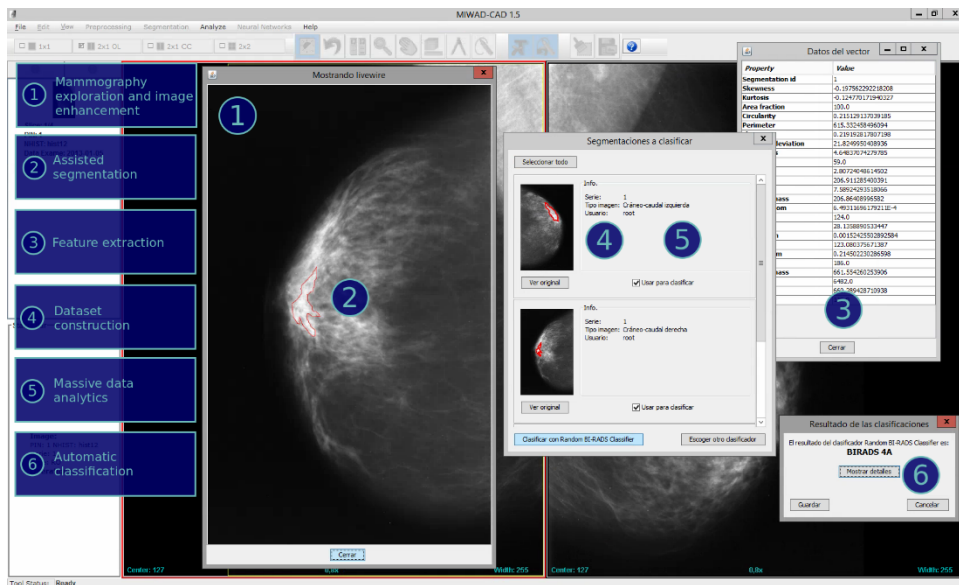


CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Casos de uso: Breast Cancer Digital Repository

- **Primer repositorio abierto** de cáncer de mama de la Península Ibérica
- **Estación CADx** para el Diagnóstico Asistido del Cáncer de Mama
- En colaboración con instituciones **portuguesas**
- Reconocido por FECYT como **buena práctica** de co-financiación FEDER



Acceso: <http://www.bcdr.eu>

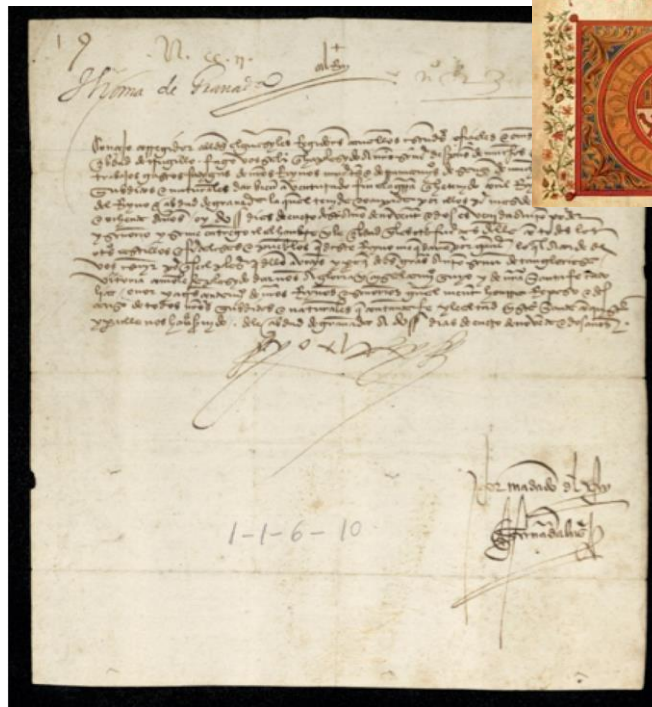


CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Casos de uso: Archivo Municipal de Trujillo

- Digitalización de fondo documental hasta el siglo XVI
- En colaboración con el Ayuntamiento de Trujillo
- Publicado en: <http://www.trujillo.cc/archivo/index.html>





CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Digital Repository Infrastructure: Estado actual

- **Desarrollo discontinuado**
 - Corrección de errores detectados por los casos de uso
- **Causa**
 - **Ocaso** de la computación GRID → **Computación en la nube**
 - Falta de **financiación**
 - **Rendimiento** pobre de la plataforma
- **Siguientes pasos:**
 - Migración a la **WWW**
 - **Adopción de tecnologías** avaladas por la **industria (BigData y NoSQL)**
 - **Evolución** de los casos de uso
 - Desarrollo y alojamiento de **portales de acceso abierto** a datos de investigación



CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Proyectos actuales: ALOE

- Plataforma de **e-Learning** para **formación** en diagnóstico del **cáncer de mama**
- Uso de **repositorios digitales** para **entrenamiento** de especialistas
- **Evolución** del caso de uso de imagen médica (BCDR)

Acceso: <http://www.aloe-health.org>

aloe alpha
Database About Sign In

A platform for radiologist training

ENTER

BREAST CANCER DIGITAL REPOSITORY
A repository of annotated breast cancer cases, reviewed by specialists and biopsy proven. It can be used to train machine learning classifiers using its API or as a reference.
DETAILS

AUTOEVALUATION FOR STUDENTS
Aloe includes a e-Learning platform which allow the auto-evaluation of the radiologist skills. Radiology students can read mammographies based in real cases anywhere.
DETAILS

MACHINE LEARNING CLASSIFIERS
Machine Learning Classifiers are integrated in the platform to tests their performance.
DETAILS

PARTNERS

CETA Ciemat
CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS

ineqi
Instituto de Engenharia, Medicina e Gestão Industrial

universidade de aveiro

SÃO JOÃO

Database Practice About Sign In

Lesion database RESET

GENDER
 Male Female

IS SEGMENTED?
 N/A No Yes

LOCATION
 Right Left

BIOPSY RESULT
 N/A Insufficient
 Benign Malign
 Suspect

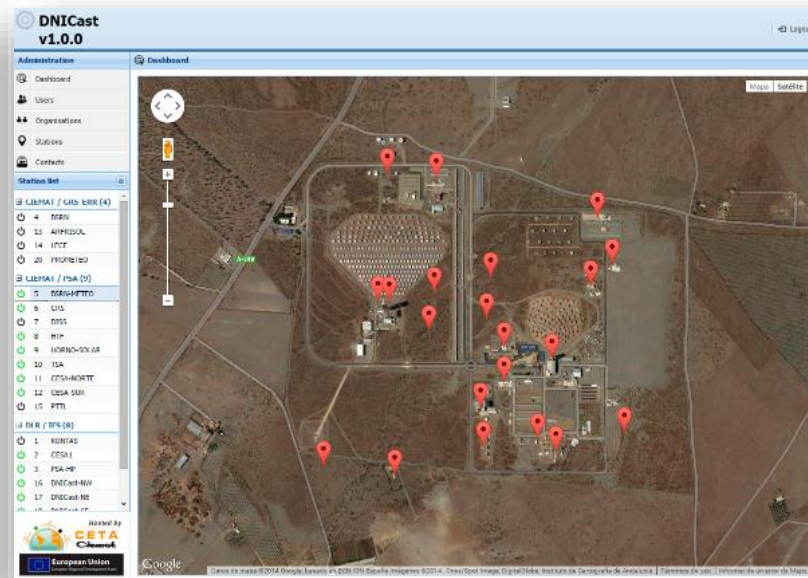
BI-RADS CLASSIFICATION

Patient	BIRADS
Age: 71, Gender: Female Lesion: Density < 25%, Location Right, Mammography Anomaly	BIRADS 6
Age: 53, Gender: Female Lesion: Density < 25%, Location Right, Mammography Anomaly	BIRADS 2
Age: 64, Gender: Female Lesion: Density 51% to 75%, Location Right, Mammography Anomaly	BIRADS 6
Age: 62, Gender: Female	BIRADS 2
Age: 65, Gender: Female	BIRADS 4A
Age: 45, Gender: Female	BIRADS 6

Proyectos actuales: **Sensoscopio**

- **Plataforma on-line para captación, normalización y visualización de datos de sensores**
- **Proyecto FP7-DNICast: Predicción de radiación directa en plantas termosolares**
- En colaboración con el **Grupo de Radiación Solar (GRS-Ciemat)**

Acceso: <http://sensoscopio.eu>



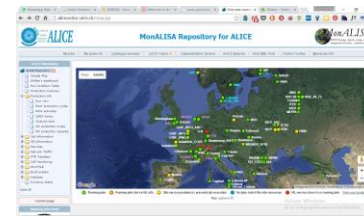


CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Otros proyectos soportados

- **ADRASE:** Acceso a los datos de radiación solar de España
 - Acceso: <http://adrase.ceta-ciemat.es>
- **ALICE:** Acceso a los datos de la colaboración ALICE (CERN)
 - Acceso: <http://alimonitor.cern.ch/map.jsp>
- **BIMED:** Base de datos de imágenes médicas (histología)
 - Acceso: <http://www.informed.unal.edu.co:8084/BiMed/>





CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Resumen

- Desde su creación en 2006, CETA-Ciemat trabaja en el **desarrollo y la explotación de repositorios digitales**
- **Digital Repository Infrastructure:** Plataforma para el desarrollo y gestión de repositorios digitales en infraestructuras avanzadas de almacenamiento
 - Breast Cancer Digital Repository
 - Archivo Municipal de Trujillo
- **Portales de acceso on-line a datos de investigación**
 - ALOE: Plataforma de formación de especialistas en diagnóstico del cáncer de mama
 - Sensoscopio: Plataforma para la visualización y gestión de datos de sensores



CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat

Agradecimientos



CETA-Ciemat agradece el apoyo recibido del **Fondo Europeo de Desarrollo Regional** (FEDER) a través de su programa operativo “Economía basada en el Conocimiento”

GRACIAS POR SU ATENCIÓN

José Miguel Franco Valiente

josemiguel.franco@ciemat.es



GOBIERNO
DE ESPAÑA

MINISTERIO
DE CIENCIA
E INNOVACIÓN



CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat



FEDER

Fondo Europeo de Desarrollo Regional

Una manera de hacer Europa