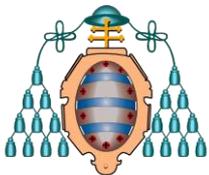


Alonso Velázquez, Pedro  
Cortina Parajón, Raquel  
Díaz Rodríguez, Susana Irene  
Fernández-Combarro Álvarez, Elías  
Ranilla Pastor, José

# GRUPO DE RECUPERACIÓN DE INFORMACIÓN Y COMPUTACIÓN PARALELA

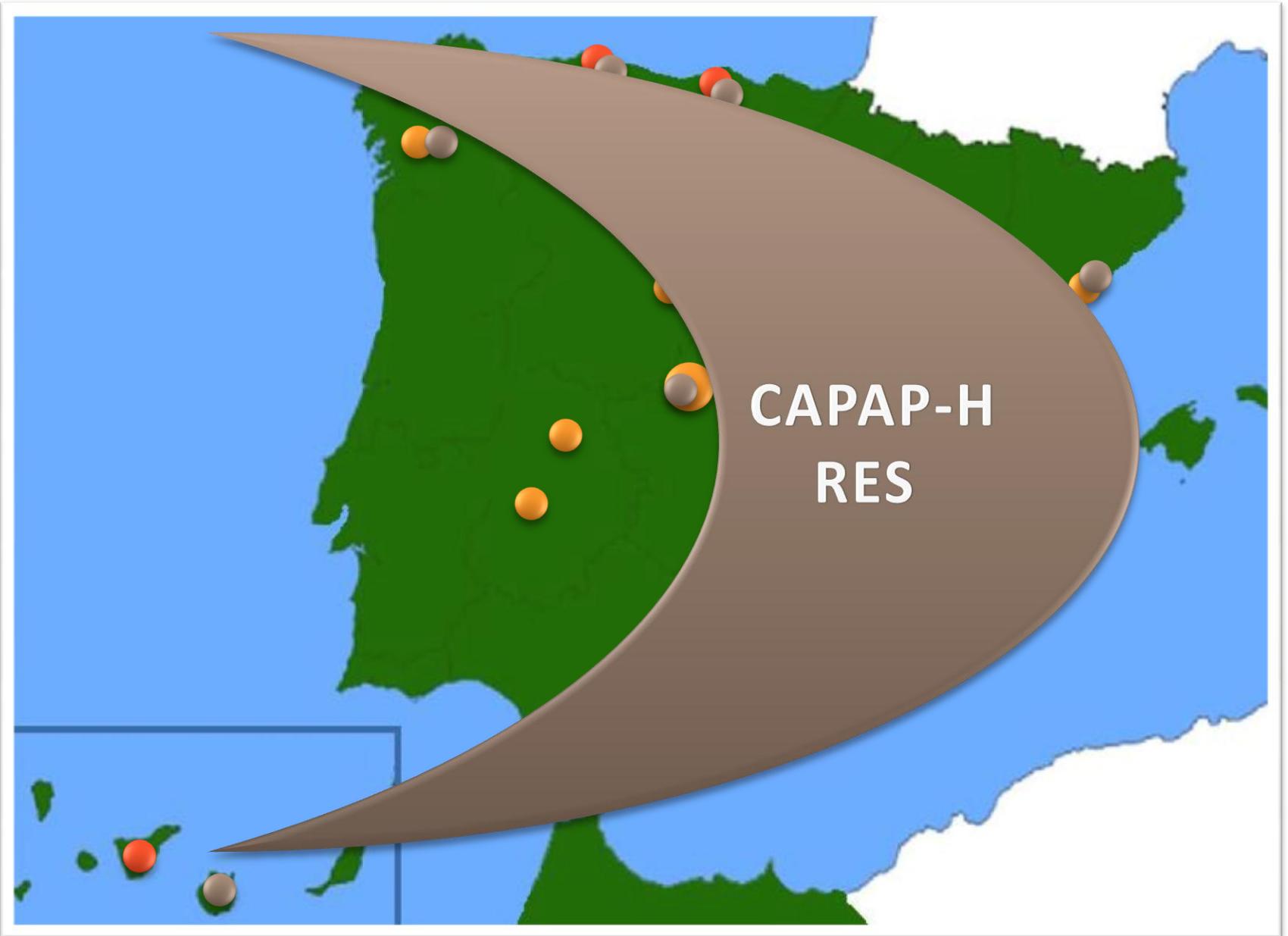


UNIVERSIDAD DE OVIEDO



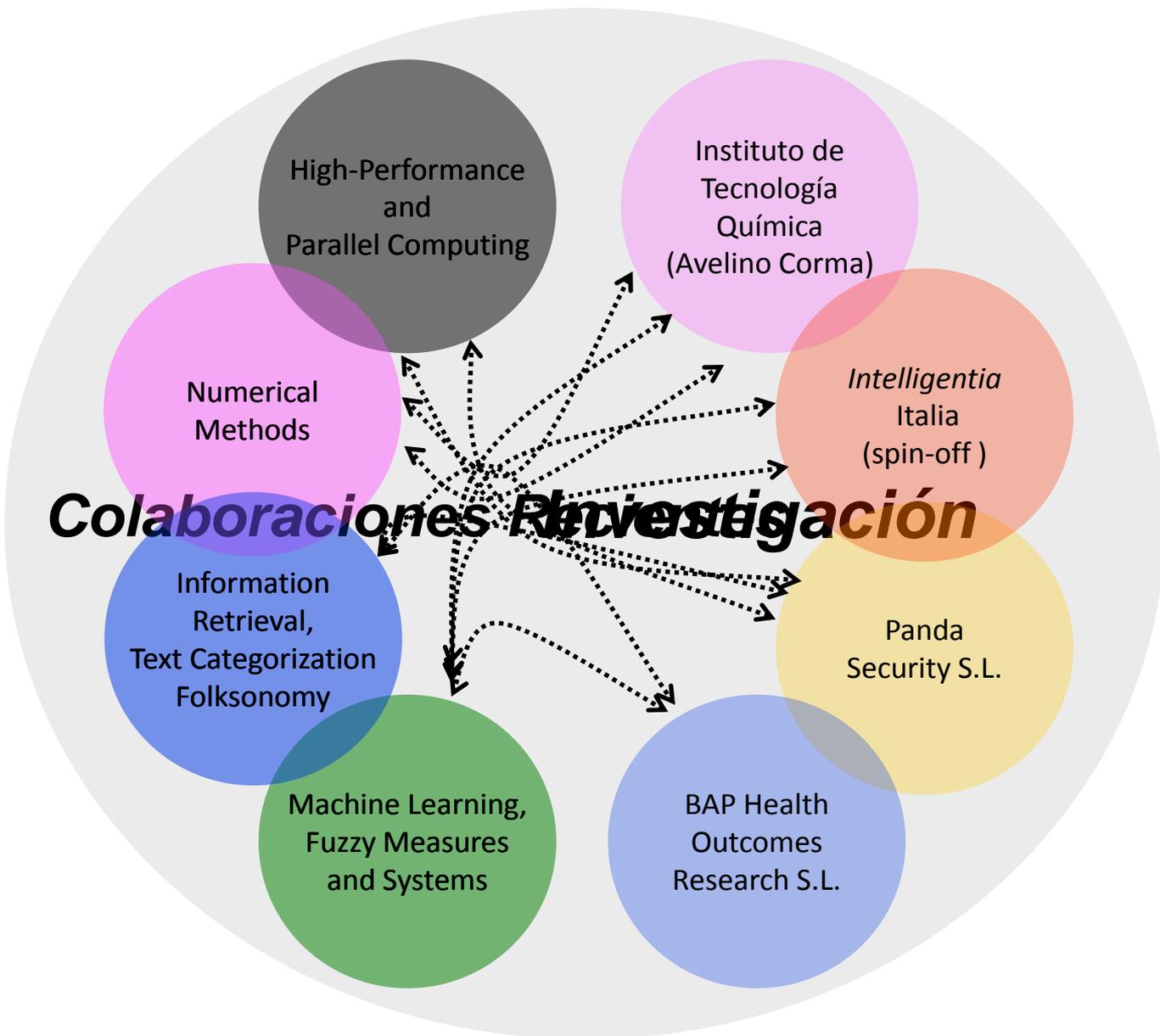
ASTURIAS  
CAMPUS DE EXCELENCIA  
INTERNACIONAL  
| AD FUTURUM |





# PROBLEMAS DE ALTA DIMENSIONALIDAD

¿Qué Hacemos?



## Definición - Aplicaciones

- Generalización del concepto de cuerpo, no necesariamente asociativos.
- Teoría de la codificación, teoría de grafos, combinatoria, etc.

## Historia

- 1960 Kleinfeld clasificó los *semicuerpos* de orden  $2^4$  (16 elementos)
- 60s Knuth y Walker clasificó los *semicuerpos* de orden  $2^5$  (32 elementos)
- 2008 Dempwolff / Nosotros clasificamos los de orden  $3^4$  (81 elementos)
- 2009 Nuestro grupo clasificó los de orden  $2^6$  (64 elementos)

## Casos Actuales

- Orden  $3^5$  (243 elementos). Más de 80 años en un procesador/core actual.
- Orden  $2^7$  (128 elementos). Más de 200 años en un procesador/core actual.



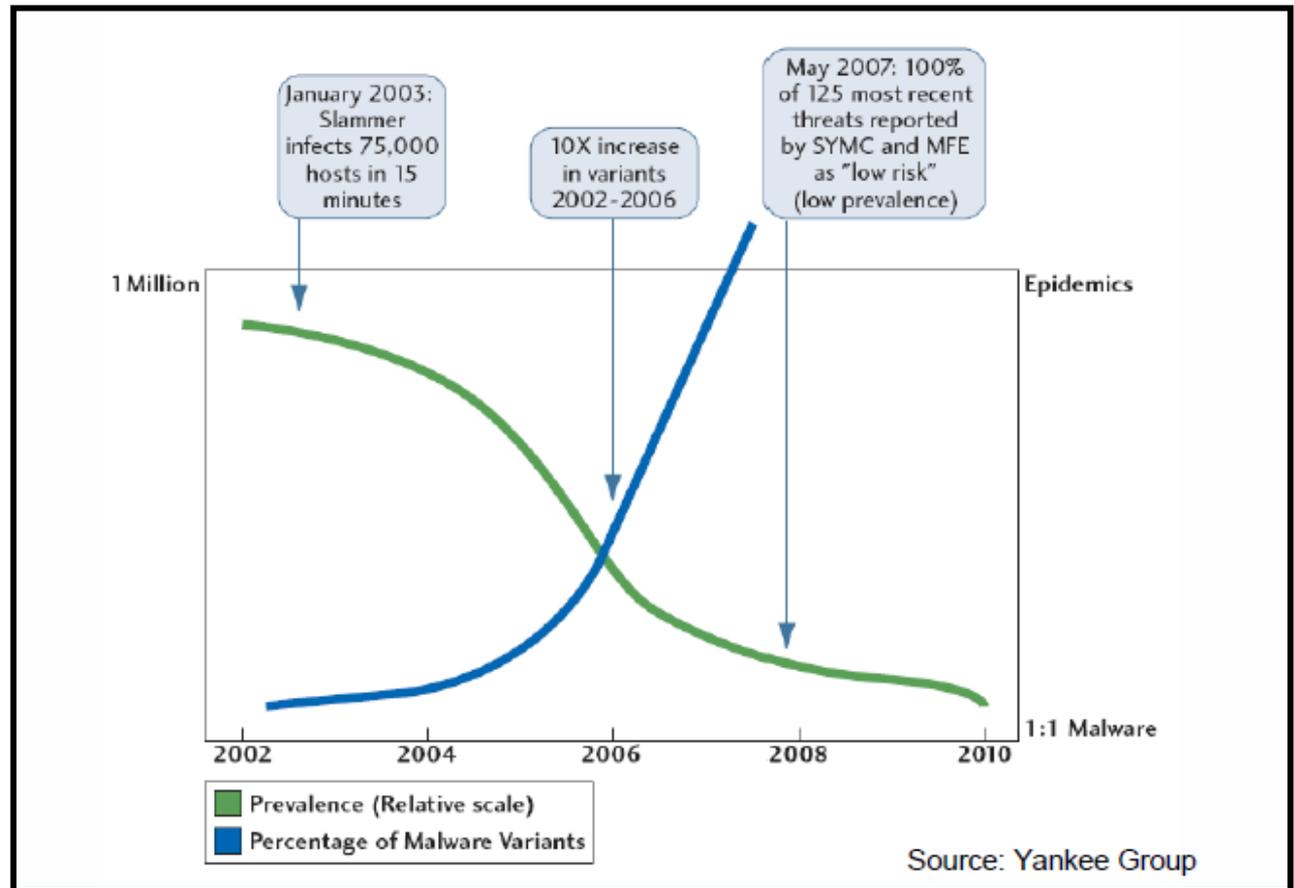
# Transferencia Tecnológica: 2008 – 2010 (sujeto NDA)

March 2, 2010

“Spanish investigators, working with private computer-security firms, have arrested the three alleged ringleaders of the so-called **Mariposa botnet** ... researchers from the **Georgia Tech Information Security Center** joined with software firm **Panda Security**...”

Panda Security

Funded by:



## Instance Selection

Millions of samples stored in DBs

Learning process unfeasible

Some Instance Selection Techniques

- Clusterization, SOM, Genetic, etc.
- Edition techniques (Wilson...)
- ...

## Feature Selection

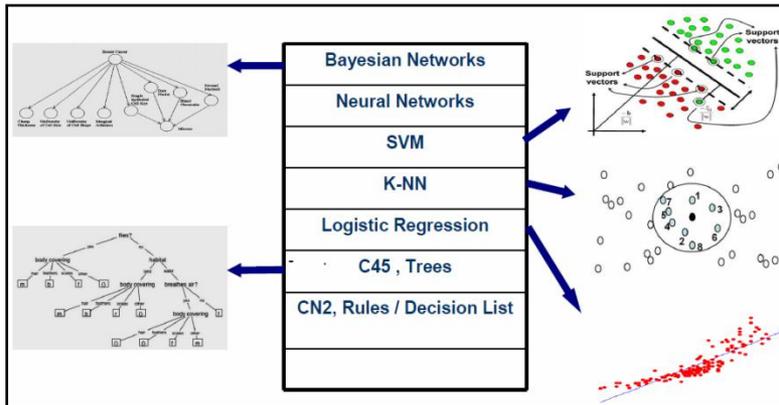
Guarantee maximum relevancy

Non-redundant evidences

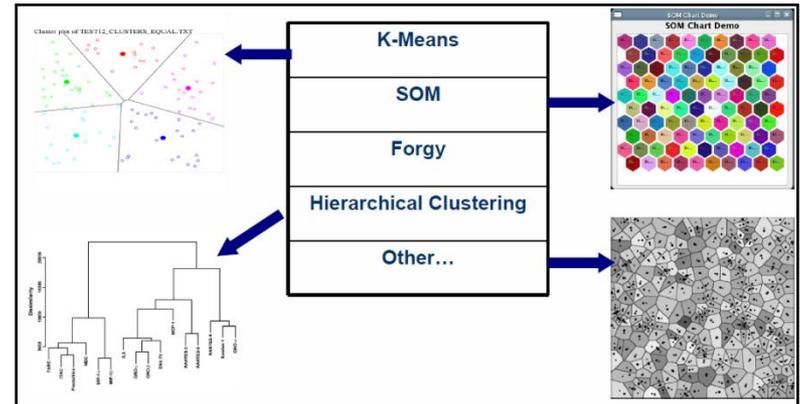
Solution

- Statistical measures
- Filter and Wrapper Approach
- ...

## ML: Supervised Methods



## ML: Unsupervised Methods



High-Performance and Parallel Computing