SIGMA CRIS
An implementation of VIVO
What we will see…

1. Why we need to improve the institutions visibility?
2. How to select the best tools? Analyzing the national and international scenarios
3. Why belong to an Open Source community?
4. SIGMA’s experience with VIVO
5. The CERIF2VIVO project
6. Conclusions
Introduction

Introducing us…

SIGMA is a nonprofit organization of Spanish public universities where we develop and maintain their scholarly information systems. In this way, SIGMA is owned by the universities that compose it.
Focusing on the Institutional Visibility
About the CRIS…

Focus on the institutional visibility: dissemination of research information

“A Current Research Information System commonly known as “CRIS”, is any information tool dedicated to provide access to and 

disseminate research information”
Institutional Visibility
Institutional Visibility

The main goals for the institution’s visibility are, among others:

- Transfer the Research results to the business sector and to the society in general.
- Reporting to the rankings providers, for evaluation, reuse...
- Access to the media and other colleagues to public profiles: discoverability

To do this better, the institution needs to easy the public access to their results.
One of the goals of the SIGMA strategic plan: The semantic approach

The goal was defined as:

“We must improve the visibility of the SIGMA’s institutions providing them scientific output Portals and experts’ guides with semantic approach, easing the search and discovering of the information.

This information must be linked with the OpenAccess repository of the institution and the tools must be provided in the Cloud.”
One of the goals of the SIGMA strategic plan: The semantic approach
How to select the best tool?: Analyzing the Spanish scenario
Spanish Scenario

2018 CRUE survey answered by 62 of 76 Spanish universities

- The most used option is the development of own applications.
- It is worth highlighting the presence of the Universitas XXI, GREC and SIGMA Research.
- In the section of others, Fundanet (Semicroll), DspaceCRIS (Dspace), Widi and Open Vivo are cited, in general, open source solutions.

CRUE survey October 2018: ESTADO DE LA CUESTIÓN DE LOS CRIS EN LAS UNIVERSIDADES ESPAÑOLAS (REBUIN)
Spanish Scenario
2018 CRUE survey answered by 62 of 76 Spanish universities

- Only 18% are housed outside the university in a SaaS system or similar.
- 71% of the solutions are hosted in the campus
- 9% of mixed solutions (SaaS/no SaaS)
- 2% don't know

Most of the solutions are not in the Cloud (SaaS model), that is the second part of the SIGMAs goal.

CRUE survey October 2018: ESTADO DE LA CUESTIÓN DE LOS CRIS EN LAS UNIVERSIDADES ESPAÑOLAS (REBUIN)
Spanish Scenario

2018 CRUE survey answered by 62 of 76 Spanish universities

- 33% don’t have a Public Portal for the Research.

CRUE survey October 2018: [ESTADO DE LA CUESTIÓN DE LOS CRIS EN LAS UNIVERSIDADES ESPAÑOLAS (REBUIIN)](https://example.com)
Spanish Scenario

2018 CRUE survey answered by 62 of 76 Spanish universities

- Another surprising finding was that 40% don’t have the CRIS linked with the IR.

Almost half of the CRIS systems are not linked with the openAccess Institutional repositories.

CRUE survey October 2018: ESTADO DE LA CUESTIÓN DE LOS CRIS EN LAS UNIVERSIDADES ESPAÑOLAS (REBUIN)
Spanish Scenario

2018 CRUE survey answered by 62 of 76 Spanish universities

- Another finding was the certainty that the CRIS information is used to **improve the research visibility**.
There is no homogeneity in the use of solutions, great percentage of them are own developments, mainly hosted in the university, not connected with the IR, and some not provide public portals.
Spanish Scenario

We analysed the tools showed in the first graphic:

- Drac
- Universitas XXI
- Cientia
- Grec

In order to define possible collaborations.
None of the solutions analysed fits at a high percentage with our strategy. The majority doesn’t contain a semantic approach or are not provided in SaaS, among others. Seeing that no one Spanish solution fits completely with our strategy, we started to analyse solutions at international level.
How to select the best tool?: Analyzing the international scenario
There is no homogeneity in the use of solutions.

- Almost the most used option is the development of own applications.
- Pure(Elsevier) it’s widely used specially in the UK, US and Australia.
- The rest use heterogeneous solutions
International Scenario

2017/18 OCLC / euroCRIS survey Practices and Patterns in Research Information Management

- A wide range of installations are hosted in campus infrastructures

Most of the solutions are not in the Cloud (SaaS model).
International Scenario

2017/18 OCLC / euroCRIS survey Practices and Patterns in Research Information Management

- Most of the activities are very similar of the Spanish scenario
There is also no homogeneity in the use of solutions and great percentage of them there are also own developments, mainly hosted in the university.
We analysed the tools showed in the first graphic:

- Pure
- Symplectic
- DSpace-CRIS
- Converis
- VIVO

We decided to focus on DSpace-CRIS and VIVO, for their OpenSource model.
One of the solutions analysed fits at about 80% with our strategy. The semantic approach, the ontology, the functionalities and the possibility to install in the cloud made that we choose VIVO. VIVO is member-supported, open source software and an ontology for representing scholarly. Composed by members of world-class universities.
SIGMA’s experience with VIVO
• We found that VIVO semantic model and functionalities fits almost 80% with the Spanish model for research, and is evolving and adapting to the new trends and have a great community behind.

• We start doing a test uploading information to an installation of VIVO (download from GitHub), and see how the information looks like.

• We could upload a lot of information in the ontology. This test satisfied us and finally we decided to use VIVO.
The next step was to start collaborating with the VIVO community, so since March of 2018 SIGMA participates in the Leadership group, LG. This group establishes priorities and is responsible for making strategic decisions.

During 2018 in VIVO we collaborate in LG with the roadmap of the product definition. It was a great opportunity for SIGMA to align their strategy and goals with the advance of the VIVO product.

We also collaborate in the creation of the Steering group, SG, a group that will bring issues to the Leadership Group for decisions, in which SIGMA is also involved.
To implement a new Experts guide, we followed the next steps:

1. **Study and analyse** the VIVO ontology.

2. **Compare the SIGMA CRIS model with the VIVO ontology.** Entities and relations that exist in the SIGMA CRIS and not in the VIVO model. The VIVO ontologies gives answer to the US scholarly model that is slightly different from the European model. We found that VIVO ontology fits almost 80% with the European model for research (and therefore, the Spanish model).
A new Experts Guide

To implement a new Experts guide, we followed the next steps:

3. **Extend the VIVO ontology** with the properties and entities for the Spanish model that are not in VIVO. The result was 10 entities and more or less 50 properties are not in the VIVO ontology. I.e:

<table>
<thead>
<tr>
<th>VIVO</th>
<th>Extended Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher - people</td>
<td>vivo-local#personID</td>
</tr>
<tr>
<td></td>
<td>vivo-local#universityPPC</td>
</tr>
<tr>
<td></td>
<td>vivo-local#urlPPC</td>
</tr>
<tr>
<td></td>
<td>vivo-local#namePPC</td>
</tr>
<tr>
<td></td>
<td>vivo-local#urlPhotoPPC</td>
</tr>
<tr>
<td></td>
<td>vivo-local#expertGroupPPC</td>
</tr>
<tr>
<td></td>
<td>vivo-local#mainResearchPPC</td>
</tr>
<tr>
<td></td>
<td>vivo-local#departmentPPC</td>
</tr>
<tr>
<td></td>
<td>vivo-local#surnamePPC</td>
</tr>
<tr>
<td></td>
<td>vivo-local#rankingPPC</td>
</tr>
<tr>
<td></td>
<td>vivo-local#expertisePPC</td>
</tr>
<tr>
<td>Estate - Organisation</td>
<td>vivo-local#orgID</td>
</tr>
<tr>
<td></td>
<td>vivo-local#tipoEstamento</td>
</tr>
<tr>
<td>Journal articles</td>
<td>vivo-local#resPubliID</td>
</tr>
<tr>
<td>Book chapters</td>
<td>vivo-local#resPubliID</td>
</tr>
<tr>
<td>Books</td>
<td>vivo-local#resPubliID</td>
</tr>
<tr>
<td>Thesis</td>
<td>vivo-local#resPubliID</td>
</tr>
</tbody>
</table>
A new Experts Guide

To implement a new Experts guide, we followed the next steps:

4. **Develop a process to upload** the information in VIVO in one step. This new process is now in beta version.

5. **Develop a new frontend angular for the Experts guide** that aligns the visual requirements of the current Experts guide of the SIGMA universities, so the process to implement VIVO will be do it as an improve of the current tools. This frontend uses the search engine of VIVO and its ontology.
A new Experts Guide

- The guide is in Multilanguage
- We can use some topics to search (Boolean connectors, literal words or phrases, approximated words, etc…)
- It searches in the whole researchers profile and scientific outputs upload in VIVO ontology from SIGMA CRIS (assuring the data quality)
- The experts are sorted by their surname. It could be sorted by relevance (not implemented yet)
- Shows the same style as the institutional web of the university
- Shows the institutional photo only if the researcher gives his consent.
- Very fast searches
- Link to the researcher profile though name of button ‘+’.
A new Experts Guide

- Access to the complete researcher profile from the CRIS system.
A new Experts Guide

The new Experts guide improvements over the current SIGMA Expert guide are:

• More powerful model (ontology) that provides scalability and more easy evolution of the tool (provide by VIVO or new evolutions)
  
• More powerful search engine
  • Search in all the researchers information, easing the discoverability
  • More possibilities for the search (booleans, literal and approximated words…)
  • Very fast and efficient searches

• More possibilities of sorting the results: for relevance, citation, impact… of the researchers

Search Tips

- Keep it simple! Use short, single terms unless your searches are returning too many results.
- Use quotes to search for an entire phrase — e.g., "protein folding".
- Except for boolean operators, searches are not case-sensitive, so "Geneva" and "geneva" are equivalent.
- If you are unsure of the correct spelling, put ~ at the end of your search term — e.g., cabbage~ finds cabbage, steven~ finds Stephen and Stefan (as well as other similar names).

Advanced Tips

- When you enter more than one term, search will return results containing all of them unless you add the Boolean "OR" — e.g., chicken OR egg.
- "NOT" can help limit searches — e.g., climate NOT change.
- Phrase searches may be combined with Boolean operators — e.g. "climate change" OR "global warming".
- Close word variations will also be found — e.g., sequence matches sequences and sequencing.
- Use the wildcard * character to match an even wider variation — e.g., nano* will match both nanotechnology and nanofabrication.
- Search uses shortened versions of words — e.g., a search for cogni* finds nothing, while cognit* finds both cognitive and cognition.
Conclusions and next steps

**This is our first pilot** and we start working on implementing it in one of our universities.

We are also working in the evolution of this tool until we get the **complete research output Portal implemented, using VIVO**.

These tools are for the reuse, visibility and dissemination of scientific information, so the maintenance of the CRIS is done by other tools. **VIVO will be for showcasing research information.**
The CERIF2VIVO project
The CERIF2VIVO Project

SIGMA is involved in the definition of the CERIF2VIVO project.

This project will be a collaboration between: euroCRIS, VIVO and SIGMA and open to other collaborations.

The goal of this project is to define an interface to upload information to VIVO always the same way independent of the source, so, the standard CERIF, seems the best option.

We have had a first experience with CERIF in the collaboration with CSUC to upload information in the PRC (Research Portal of Catalan Universities). Even though PRC only uploads a subset of the information that is stored in a CRIS (articles, books, book chapters, projects and thesis), it was a good experience.
The CERIF2VIVO Project

To move ahead with this project, we need to avoid one of the current drawbacks of CERIF, namely its complexity. So a parallel refactoring of the model is planned.

This refactoring proposes to divide the CERIF model in a CORE model that will contain the basic entities, and then the definition of specific areas of entities that completes the model and that should be defined by experts in the area.

At the same time the CERIF refactoring takes place, the mapping to VIVO will be defined.
Conclusions
Conclusions

• Our participation in the VIVO community it has been a great experience and an opportunity to evolve and be able to provide to our universities solutions align with the solutions used by world-class universities (that leads the international rankings) that are in the VIVO community, using and adopting their best practices at some levels.

• We want to be a reference in the use of the tool, so we could collaborate closer with it’s evolution and use.

• We are engaged with the collaboration between these two relevant organisations that are euroCRIS and Duraspace and hope that we can do relevant advances.
And an invitation… VIVO conference for the first time in Europe!