

DeGóis – The Portuguese National Academic CV Platform

Carlos Sousa Pinto^a, Pedro Leitão Lopes^b, Luis Pedro Valério^b, Luis Alfredo Amaral^a

^aCentro de Investigação ALGORITMI, Escola de Engenharia, Universidade do Minho, 4804 Guimarães, Portugal

^bEquipa DeGóis, Departamento de Sistemas de Informação, Escola de Engenharia, Universidade do Minho, 4804 Guimarães, Portugal

Abstract

Science and Technology Foundation (FCT) is the Portuguese authority responsible for the management of science and technology that deals with a large part of governmental funds to be allocated to researchers, in terms of grants and research projects financing. FCT owns DeGóis, a Portuguese academic CV platform derived from Lattes, the Brazilian system of curricula. Due to that, DeGóis inherited almost all the assumptions adopted for that platform without any alignment with the international recommendations or standards and best practices about usability in use today. This fact introduced several problems concerning the interoperability with other national and international systems and kept away many potential users. A consequence of this was the decision to develop a new version of DeGóis, as part of the national project PTCRIS (Portuguese Current Research Information System). In this article we describe the path that led to the development of the new version of DeGóis, highlighting the objectives, motivations, expectations, options, results, limitations and future work.

Keywords: CRIS; Interoperability; CERIF; CASRAI; ORCID; DeGóis.

1. Introduction

In 1999, the National Council for Scientific and Technological Development (CNPq) in Brazil launched the platform ‘Lattes’ as its national system to support Curricula and as a central component of its Science and Technology management system.

The success of the first years of use of the platform ‘Lattes’ in Brazil, added the will to share the acquired experience, motivated the establishment of an International Network for Science and Technology management, in Portuguese and Spanish language.

In December, 2001, a Delegation composed of officials from the CNPQ and the Stela group (responsible entity for the design and implementation of the platform Lattes) were in Lisbon invited by FCT. The two partners presented the platform Lattes, the international agreements and developments, as well as the prototype meanwhile developed, considering the Portuguese reality and data. This pilot experience served as an embryo for the platform DeGóis, which was later developed.

As a result of that mission, the Portuguese and Brazilian officials decided to cooperate in the domain of Science and Technology Management Systems and a collaboration agreement between the Academic groups, involved in the development of the systems (Stela Group/ Federal University of Santa Catarina – Information Systems Department / University of Minho) was established.

On 4th December, 2002, 12 countries signed the Florianopolis Declaration, which established the Network SCienTI joining Portuguese and several Spanish speaking countries in an international S&T network^{1,2}.

SCienTI intent is to contribute for the S&T development of the participant countries, by facilitating interchange mechanisms, complementary and socialization experiences, products, services and information systems in ST&I management support.

This Network was designed as a space of convergence between different actors in knowledge management for Science development and technological innovation, which can generate methodologies, tools, instruments and information on ST&I products.

2. Objectives and Motivations

DeGóis has a long history. The first version of the platform has been in use for more than a decade, either by university teachers and researchers or by grant applications sponsored by FCT³. The platform provided several web services and institutional indicators that were adopted by Portuguese high school institutions and its full-text repositories.

In order to obtain a wide range of indicators of scientific production of researchers, DeGóis had implemented a detailed scheme supporting structured information. However, the complexity of the process and the user interface were difficult to understand by the users. At the same time, information was lacking integration with other research systems.

At the same time FCT sponsored simultaneously another competing platform - FCTSIG - of a much simple usage with free text fields where users could fill in their data, by copying text from other documents such as docx or pdf formats. So, users were able to submit applications for grants or project financing in an easier way. However, the FCT could not treat such CV information as it lacked in structure. The achievement of productivity values and other indicators, to support funding decisions, was a not easy task.

Against this background, it was necessary to move towards a solution of just one CV platform, to solve or minimize the problems experienced by using the two existing platforms. It was decided to develop a new version of DeGóis as a unique national CV system, recognized and made mandatory by FCT for future applications for grants and other funding. These evidences and decisions lead to the development of a new version of the Portuguese national CV platform that ought to implement the following main characteristics: more user-friendly and easier to use; more intuitive; and more interoperable.

The new DeGóis is now perceived as an important part of a larger ecosystem of science and technology designed to support Portuguese researchers, institutions and governmental agencies.

3. Methodology

DeGóis is an academic CV platform that integrates the new Portuguese ecosystem of science and technology management. In order to design and implement a new version of DeGóis, with focus on interoperability and usability, it was used a goal driven process methodology (see Figure 1). The project started by identifying main concepts that could attend a state of the art analysis in the subject. This task ensured the analysis of several RIS (Research Information Systems) available in the software market and the exploration of available semantic and data dictionaries such as CERIF and CARSRAI dictionary. Other similar projects were reviewed in order to study new layout approaches that could benefit ergonomic and usability issues.

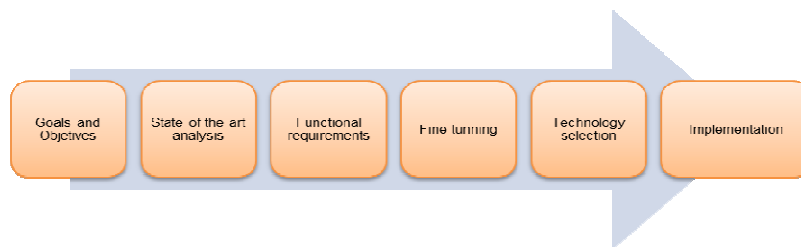


Figure 1 – Goal driven fine tuning process (adapted from Bauer, C. & Dey, A. K.⁴)

This research leads to the identification of the CASRAI project as a provider of guidelines for the development of the new version of the platform that could embrace the DeGóis and new international RIS concepts. CASRAI provides a dictionary of terminology and exchangeable business objects that can improve the transaction of research data between organizations⁵. The choice of CASRAI dictionary has allowed a more structured development that came to reflect positively on the final result. Concerning interoperability matters, ORCID was acknowledged as a potential main source of researchers metadata related to works, projects, identification, career, etc⁶. The links between ORCID and several scientific journal databases (e.g. PubMed, WoS, Scopus, etc.) could improve DeGóis to relate this data with researchers CVs. Further, the project identified a set of systems and databases in use in several state agencies related to the management of information on science and technology in Portugal. Such systems and databases have relevant information that can be incorporated into the CVs and it was essential to understand how this integration could be achieved.

The referred findings were discussed and analyzed in order to produce a functional requirements list that could be an input to the new software functionalities or product backlog. The backlog list was iterated through prototypes and functional analysis allowing to fine tune the importance addressed to each task. This process sorted out functionalities that could not contribute to final results or that could be redundant to system design in the context of the referred national ecosystem. Other aspects were attained, for instance, project constraints such as timing, team skills, project management methodology, etc., that could compromise the product's viability.

The product backlog and analysis were relevant inputs to selecting the appropriate software technologies that could assist the design and implementation of the new CV platform. Several APIs were tested and implemented that allowed DeGois to obtain various types of information, namely: CrossRef for funding and articles search; ORCID API for works, projects and funding synchronization, among others. The software architecture design employed the MVC (Model View Controller) model and it was implemented by using JSF (Java Server Faces) and Primefaces. These aspects allowed the project team to focus the application development on user's interface optimization and interoperability enhancements.

The project built a working product, the new DeGóis, that had iterated and incremented the new features during the development phase. Finally, usability and validation tests were performed, and questionnaires were designed to capture the reactions of the users involved in such tests.

In order to manage the work in progress, other working methodologies were selected and adopted by DeGóis team. Initially the project followed an iterative and incremental methodology that allowed producing functional analysis and working prototypes. However, due to the increased work log and iterations cost-benefit, it was realized that we should change to a more responsive methodology such as the Agile model. This change allowed a more dynamic control of feature developments and delivering quality.

5. The new DeGóis

DeGóis is part of the major Portuguese project PTCRIS whose central principle is "register once, reuse many". As part of the development process of the new version of DeGóis, a survey of the main existing Portuguese systems and databases was carried out. Those systems and databases can be found in various governmental agencies linked to issues of high education, science and technology, and statistical analysis of all the data they produce or collect. In that process various APIs were designed and developed, in order to allow those systems to access the information available at DeGóis. At the same time, DeGóis will receive, from those systems and databases, useful information for de CVs of the users.

In addition to the CASRAI dictionary, which guided the development of the new version of DeGóis platform, we have also considered other projects and the various APIs made available by those projects, in order to allow the access to the provided information.

We referred above a set of Portuguese official systems and databases containing or manipulating relevant information on scientific and technological context. Those artifacts can't share the information they deal with, because they were not developed with integration concerns. The new DeGóis allows that integration and will play the role of a national hub of scientific and technological information.

Also as a consequence of the decision-making of the adoption of ORCID as a central hub for the PTCRIS we realized that the CV created within the DeGóis platform would have to receive information from other sources, including local CRIS, national databases and manual entry information from the users, since the ORCID had no such information. This is the case of activities which are carried out throughout the researcher's academic career. As examples, we point out teaching activities, organization and participation in conferences and workshops, postgraduate students supervising, etc.

Comparing the relationship between DeGóis and ORCID, where ORCID is considered the hub of scientific works and projects, DeGóis can play a similar role in its interaction with Portuguese local or institutional CRIS. DeGóis can act as the source and the destination of information of interest to both systems. Figure 2 depicts a representation of the architecture involving ORCID and DeGóis in the context of PTCRIS where the interaction role of each platform is clear.

In the Figure 2, green arrows represent interactions already active involving the new DeGóis and the thin white ones, represent direct flows of metadata not going through ORCID or DeGóis.

Unidirectional arrows represent metadata flows from DeGóis to the target entities, from source entities to DeGóis and from RCAAP (Portuguese federation of full-text repositories) to ORCID.

Bidirectional arrows represent synchronization between DeGóis and ORCID related to metadata about publications, projects and personal data.

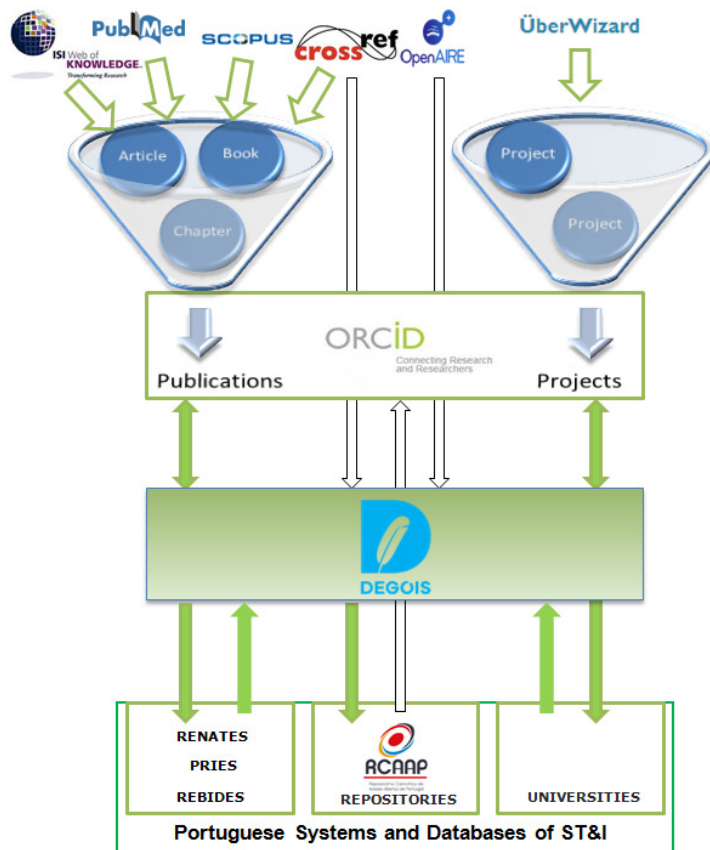


Figure 2 - DeGóis and their relationships with ORCID and other PTCRIS systems

The new DeGóis is a central part of the national project PTCRIS, started about 2 years ago and led by FCT. It was somewhat strange for us to realize that the same problem could be viewed with very different perspectives which influenced the development of new DeGóis. Figure 3 shows a representation of those two perspectives.

DeGóis team was greatly influenced by their history of more than 10 years of contact with users and has continued to focus on the viewing platform as a CV system centered on the user. On the other hand, FCT tended to see the platform as a subsidiary system to other components of PTCRIS, namely the one designed to support the financial decision making (scholarships and research projects financing) based on productivity indicators made available by DeGóis.

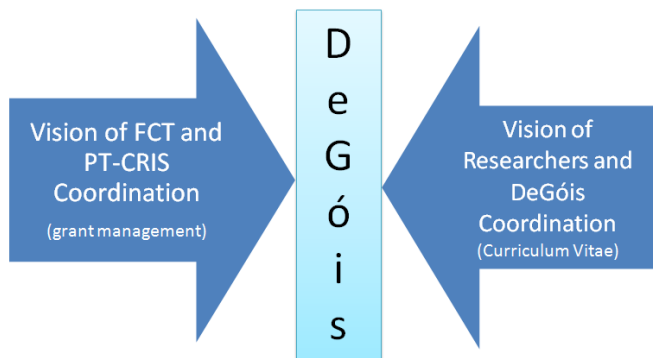


Figure 3 - Two different perspectives of the DeGóis platform

6. Interoperability mechanisms

To apply the basic principle of information reuse, DeGóis created several interoperability mechanisms to obtain information from various sources and provide their own information. These mechanisms have been developed in order to allow a faster fill of the CV.

ORCID⁷ as central hub of PTCRIS, presents itself as the most important source of information for DeGóis because it brings together not only the records from various international sources of information such as Scopus, Thomson Reuters, CrossRef, UberWizard, etc, but also the national sources that integrate PTCRIS ecosystem.

Due to the need of sharing the information between the national source members of PTCRIS with the purpose of its reuse, DeGóis implemented an information sharing mechanism with ORCID that enables the two sources to keep their records up to date. DeGóis is able to quickly collect ORCID information and integrate it into the curriculum, send information to ORCID and update the information in both systems. This mechanism is still endowed with automatisms that link information automatically when this is present in the two systems.

Crossref⁸ is another source of information that DeGóis uses. This organisation aggregates information from publications in the form of metadata, provided by more than 5000 publishers and editors⁹. It is therefore an essential source of information for DeGóis. This information allows, in most cases, the complete filling of a publication record in DeGóis that is not possible using only ORCID because the information available in that platform is not complete enough. Filling in the several fields of a work takes place in a simple way, through an autocomplete mechanism for the title of publication. This kind of mechanism can be extended using the name of a journal or the DOI of the publication, if known.

OpenAIRE¹⁰ is a European database of open access content. This data source is used by DeGóis for filling in information about projects. Also through a process of autocomplete by title of project the platform can make a complete filling of a project record.

Within the Portuguese context, the full-text repositories are also a source of information that DeGóis incorporates. As members of the PTCRIS ecosystem, they interact directly with ORCID and the synchronization mechanism of DeGóis collects the metadata about those works using the same process used for other publications available in ORCID.

This approach is an evolution of the previous solution for the same problem. That solution considered the communication from DeGóis to the full-text repositories and from the full-text repositories to DeGóis. In that approach, users had the possibility of sending directly the metadata of their works from the full-text repositories to the DeGóis platform and the full-text documents and its metadata from DeGóis to the full-text repositories. Being used for several years, that solution, however, didn't implement the mechanism of synchronization referred before.

Using several APIs, DeGóis also has the capacity to provide the information about CVs in xml format files to other national and international entities.

7. Limitations

The development of the new version of DeGóis platform was strongly influenced by the decision to follow the recommendations of CASRAI (Consortia Advancing Standards in Research Administration Information) project and to implement a synchronized integration with ORCID platform. These decisions brought several difficulties to the project. On one hand, CASRAI dictionary is not stabilized and therefore has been necessary to amend the DeGóis raised platform many times during the developing process. On the other hand, the platform ORCID viewed by DeGóis and PTCRIS as the central hub of scientific works does not concentrate all the information of a CV and also is not stabilized. The decision to synchronize information between DeGóis and ORCID made us use multiple versions of the API available for that platform. The limitations that this API presented in its version 1.2, and the promises of the release of version 2.0 in 2016, have led us to adopt such API Release Candidate versions, which even today, continue to evolve. This resulted in repeated situations of occurrence of unexpected errors, as consequence of the many changes that are being introduced by ORCID project in its new API version 2.0.

Although presented as compliant with CASRAI, ORCID just deals with a subset of information of CASRAI dictionary. Due to that, sometimes we did not find CASRAI relevant information in the set of elements we can access by using ORCID API.

Also because we were concerned with the possibility of doing the migration of CVs from the former DeGóis to the new version of the platform, and because the new DeGóis uses a very different data model, we had experienced a hard work of mapping the previous data model to the new one.

8. Outcomes

We must highlight the inclusion of a synchronization mechanism between DeGóis and ORCID allowing bidirectional flow of information between the two systems, regarding scientific works, projects, personal information and educational and professional background of users.

The information is synchronized, based on a synchronization algorithm developed under the aegis of the PTCRIS project.

DeGóis synchronization process introduced a significant improvement regarding the consistent update of the information in both systems.

The synchronization process between DeGóis and ORCID bases its operation in sharing of external identifiers of ORCID and DeGóis works and following this process, DeGóis presents a table to the user with unsynchronized works, and with the actions he can accomplish. Figure 4 shows an example of that table within the synchronization process.

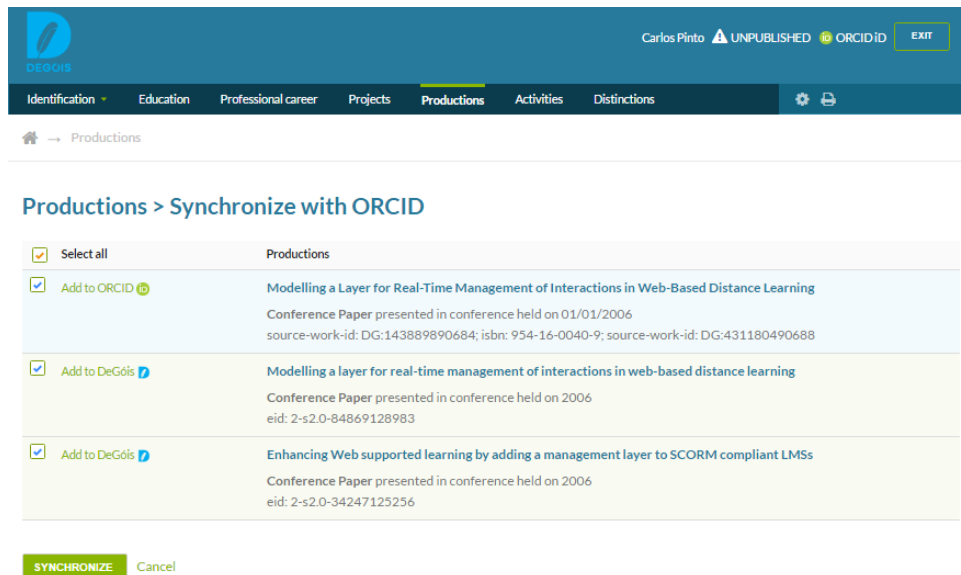


Figure 4 - Sync actions on unsynchronized works

As shown in Figure 4 one can see three works identified automatically by DeGóis which can be synchronized. If the user marks the checkbox "Select all" and presses the button SYNCHRONIZE at the bottom of the page depicted in Figure 4, two of them will be added to DeGóis CV and the third will be sent to ORCID CV.

Currently, printing a CV in the new version of DeGóis platform includes sending to the user an XML file with the result of that CV printing configuration. The availability of such a file, using a universal format, will allow further use of the information contained therein by other applications accessible to the user.

Also regarding the process of printing a CV, the platform allows to print it in a Microsoft "docx" format, which enables the user to change the content of its CV using Microsoft Word.

Our plans for 2016 include an improvement, a crucial matter to us, once it will represent a real possibility of great interoperability with other national and international systems. We refer to the availability of generating the information of a CV in CERIF-XML format.

We can't say that there are already many systems compatible with CERIF but taking this step, we will enable interoperability of DeGóis platform with other systems also adopting the standard¹¹.

DeGóis platform should therefore evolve towards higher levels of interoperability and should increase its international exposure as well. Its public viewing component, by implementing various search filters, will allow to find people whose CV include work in the areas of interest of the researcher by is using that facility, allowing contacts between these people.

9. Conclusions

We believe that the results obtained in carrying out this work are relevant, not only because it was concluded that the adoption of CASRAI dictionary allowed the development of a new version of DeGóis platform better structured, with better organization of information, but also made it more intuitive and therefore easier to use. At the same time, by adopting the CASRAI dictionary as a reference, we put the platform on a level that allows the exchanging of information with other systems which have adopted the same reference.

On the other hand, the ORCID is an increasingly used system, constituting a good solution to clarify the identification of the researchers in a worldwide dimension through the ORCID iD. This solves many of the problems related to the correct identification of the authors of the scientific production inserted in a CV.

Two usability tests were performed, and questionnaires were designed to capture the reactions of the users involved in such tests: a first test related to the platform still in use, and a second one, about the new version of DeGóis. In both cases we chose two types of users - one composed of people without any previous contact with DeGóis and other comprising users with experience in using the previous version of the platform. The idea was to realize how new users reacted to the use of both versions of DeGóis. The objective of involving people with previous experience with DeGóis was to determine, whether these users took notice of a significant and positive difference, when using the new version of the platform. In the end, all users have decided that the use of the new GeGóis version was significantly more effective. These findings mean that it was worth the work we carried out in the last two years.

Obviously, we are aware that these findings must be validated in a real usage scenario. In a couple of months we will start to have such feedback when the new version of DeGóis goes into operation.

Acknowledgements

This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2013.

The authors would like to thank the PTCRIS coordinator, João Moreira, and the FCT collaborator, Bruno Monteiro, for their contributions to the design of the new version of DeGóis and for the fruitful discussions that helped us clarify and improve the new version of the platform.

References

1. Amaral, L., Santos, L., and Bernardo, C.A., “Uma visão do Sistema Científico e Tecnológico Português”, 2002, in Actas da I Workshop da Rede SCienTI, Santa Catarina, Brasil.
2. Stallbaum, I., “Rede SCienTI já é realidade”, in Revista do MIC, 2003, Núcleo de Mídia Científica: Universidade Federal de Santa Catarina. p. 2.
3. Magalhães, S.T., L.D. Santos, and L. Amaral, “Curriculum DeGóis – A Scientific e-Portfolio, in ePortfolio 2006”, 2006, 4th International ePortfolio Conference, Oxford, England.
4. Bauer, Christine & Dey, Anind K., “Considering context in the design of intelligent systems: Current practices and suggestions for improvement”.*The Journal of Systems and Software*”, 2016, 112 26–47.
5. Haak, L., Baker, D. and Hoellrigl, T., “CASRAI and ORCID: Putting the pieces together to collaboratively support the research community”, 2014, *Procedia Computer Science*, 33 284 – 288.
6. Akers, K. G., Sarkozy, A., Wu, W. and Slyman, A., “ORCID Author Identifiers: A Primer for Librarians, *Medical Reference Services Quarterly*”, 2016, 35:2, 135-144, DOI: 10.1080/02763869.2016.1152139.
7. Haak, L., Fenner, M., Paglione, L., Pentz, E., & Ratner, H., 2012, “ORCID: a system to uniquely identify researchers. *Learned Publishing*”, 25(4), 259-264, doi:10.1087/20120404.
8. Brand, A., “CrossRef turns one”, 2001, *Learned Publishing*, 17(3), 225-230.
9. Crossref.org, “Publishers & Societies”, 2015, <http://www.crossref.org/01company/06publishers.html>, last access done at 17-05-2016.
10. Houssos, N., Jörg, B., Dvořák, J., Príncipe, P., Rodrigues, E., Manghi, P., & Elbæk, M. K., 2014, “OpenAIRE Guidelines for CRIS Managers: Supporting Interoperability of Open Research Information through Established Standards”, *Procedia Computer Science*, 33, 33-38. doi:10.1016/j.procs.2014.06.006.

11. Pinto, C.S., Simões, C. and Amaral, L.A. (2014). "CERIF – Is the standard helping to improve CRIS?", 2014, *Procedia Computer Science*, 33, 80-85.