PTCRIS: Planning and implementing a national CRIS ecosystem

João Mendes Moreira
https://orcid.org/0000-0002-9081-2728

Cátia Laranjeira
https://orcid.org/0000-0003-3952-8294
Outline

- About FCT
- About PTCRIS
  - Why
  - How
  - What
    - PTCRISync
    - Org Ids
    - Funding/Projects
    - Outputs
    - Ciência Vitae (CV)
- Conclusions
Foundation for Science and Technology in an Nutshell
Fundação para a Ciência e a Tecnologia (FCT) is the Portuguese national funding agency for science and research.

FCT is a publicly-funded government agency, under the responsibility of the Ministry for Science, Technology and Higher Education.

FCT’s vision

- To establish Portugal as a global reference for research and innovation
- To ensure knowledge generated by scientific research underpins social and economic development
FCT supports the whole spectrum of funding instruments
Main pillars

CONNECTIVITY
COLLABORATION
KNOWLEDGE
COMPUTING
SECURITY
Scientific Information

Mission:
1) To ensure the community access to sources of scientific information of recognized prestige and quality;
2) To promote, support and facilitate open access to Portuguese scientific production:
3) To facilitate management and access to information on national scientific activity
PTCRIS in a nutshell
Why?

Scientific community

- **Simplify** the management and reporting of information on scientific activity

Managers

- **Optimize** the evaluation of the scientific practice, the **administration** of the information systems on the same and report

Public

- **Facilitate the discovery** of innovative ideas
- **Foster** the approach between **Science and Society**.
The problem

Unnecessary effort, incomplete and unreliable information!
Ecosystem of Science and Technology

SoS = System Systems = ecosystem
Interoperability

System A

- PIDS
- Data model
- Semantics/vocabulary
- Other standards

API

System B

- PIDS
- Data model
- Semantics/vocabulary
- Other standards

Profiles
Mission

Ensure the creation and sustained development of a regulatory framework (protocols/standards) and infrastructures to implement an integrated national information ecosystem (PTCRIS) to support the management of scientific activity.
National cooperation
Cooperation / “Lobby”
The process

Prospection  Analysis

Adoption  Implementation

Kit / Framework  Tests

Pilot

Standard / Protocol
The methodology

Time / adoption of regulatory framework

System A
- Standard 1
- Standard 2
- API

System B
- Standard 1
- Standard 2
- Standard 3
- API

Platform 1
- Standard 2
- Platform 2

Platform 3
- Standard 3
Standards

Regulatory framework

Bussiness Profiles

Technical Profiles

data models

(model and exchange format)

Ids

People, Orgs, Fund., publications.

Semantics/ Vocabulary

CERIF

VIVO

COAR

PTCRIS
Ecosystem of Science and Technology

 Researchers | Affiliations | Organizations | Funding / projects | Outcomes / outputs | Infrastructures

CV | Grant Management | Local CRIS

Research Portal | Business Intelligence

SoS = System Systems = ecosystem
What – PTCRISync – Motivation

"A output should be recorded once and reused multiple"

Feb 2014, FCCN Summit
João Mendes Moreira
PTCRISync concept

https://vimeo.com/251179542
Average implementation time

4 to 6 weeks
PTCRISync Demo

https://vimeo.com/251312395
Team

Alcino Cunha
Library design
0000-0002-2714-8

António Lopes
Integrator & developer
0000-0003-3045-0

Bruno Monteiro
UX Designer
0000-0002-3300-7

João Mendes Moreira
Project Leader
0000-0002-9081-2

Luís Pedro
Integrator & developer
0000-0002-3936-6

Nelson Madeira
Integrator & developer
0000-0002-4575-4

Nuno Macedo
Library design & developer
0000-0002-4817-9

Paulo Graça
Developer
0000-0002-3503-4
Want to know more about PTCRISync?

Videos

– **Teaser**
– **Demo**
– **Tech**

[GitHub](#)
[Wiki](#)

E-mail: [ptcrisync@googlegroups.com](mailto:ptcrisync@googlegroups.com)
(Some CERIF entities and their relationships – adapted from http://proj.badc.rl.ac.uk/moles/attachment/wiki/CRIS/cerif-main-structure.jpg)
Why?

• Problem:
How many scientific papers from Lisbon University are indexed in WoS?

Number of scientific papers in WoS

Univ Lisboa

Technical sheet: WoS publications between 2010 and 2017 whose affiliation of the Corresponding Author is the University of Lisbon excluding data from the former Technical University of Lisbon, Colleges and Centers. Courtesy of Inês Fonseca.
Difficult to unequivocally identify organizations operating in the HE and Science system

Impact on the association of ORGs to person/projects/outputs

Formulation of inaccurate reports/documents
- Difficulties of analysis
- Difficulty in decision making
Aim

Unique persistent identifiers

Associate accurately ORGs to Persons and their info

Info flow in the ecosystem

DATA

Accurate Complete Updated
How?

Define a **UNIQUE** and **PERSISTENT** OrgID to adopt according with best national and international practices

Promote the use of OrgIDs by adopting an **AUTHORITY TABLE**

Ensure data accuracy regarding Portuguese organisations
What? – implementation timeline

- **2015**
  - State of the Art Study
  - Adoption of ISNI+ (Ringgold’s authority table)
  - Set up of a reconciliation tool – Google Refine

- **2016**
  - Definition of principles and guidelines
  - Reconciliation of ORGs DB (HE and Research)
  - Specification of an OrgNR

- **2017**
  - Reconciliation of ORGs DB (Corporate/Research-IPCTN)
  - Integration with CV

- **2018**
  - Mapping ISNI+/National Codes
  - Other integrations...
  - Integration with RCAAP
  - Reconciliation of ORGs DB (Corporate/Research-ANI)
What?
– PTCRIS services for ORGs management –

1. Search API
2. FTP access to the authority table (Ringgold’s DB)
3. Reconciliation Service
4. Mapping ISNI+/national code
PTCRIS services for ORGs management:
1. Search API

https://api.cienciavitaes.pt/docs/
PTCRIS services for ORGs management:
1. Search API
PTCRIS services for ORGs management: 2. FTP access to the Authority Table

- PT ORGs

**User Manual**
- Metadata schema
- Ringgold Classification
- Ringgold ORGs Types
- Data snapshot

**Access conditions**
- Ask for access credentials: ptcris@fccn.pt
- Sharefile platform

**Data formats**
- json
- xml
- csv

- PT ORGs and others (~500k ORGs)

Available upon signing of agreement
PTCRIS services for ORGs management:
3. Reconciliation Service
PTCRIS services for ORGs management: 3. Reconciliation Service

Import wizard > Import thesis and dissertations

- Select all

- Add

- Education level
  - Portugal
  - Other

- Education level - Portugal *
  - Doutoramento

- Institution/Organization *
  - Universidade Nova de Lisboa Faculdade de Ciências e Tecnologia

Enter and select an institution or organization (eg. Universidade do Porto)
PTCRIS services for ORGs management:

4. Mapping ISNI+/national code

<table>
<thead>
<tr>
<th>ringgold</th>
<th>ISNI</th>
<th>DGEEC</th>
<th>FCT</th>
<th>NIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>56060</td>
<td>0000000122872414</td>
<td>1512</td>
<td>50013</td>
<td>508910099</td>
</tr>
<tr>
<td>70904</td>
<td>0000000121913202</td>
<td>4555</td>
<td>500745684</td>
<td></td>
</tr>
<tr>
<td>70988</td>
<td>0000000406321876</td>
<td>4539</td>
<td></td>
<td></td>
</tr>
<tr>
<td>86719</td>
<td>000000101903701</td>
<td>757</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89237</td>
<td>0000000406209518</td>
<td>50005</td>
<td>506134466</td>
<td></td>
</tr>
<tr>
<td>112044</td>
<td>0000000403923425</td>
<td>50012</td>
<td>500825840</td>
<td></td>
</tr>
<tr>
<td>112046</td>
<td>0000000403923732</td>
<td>283</td>
<td></td>
<td></td>
</tr>
<tr>
<td>112047</td>
<td>000000107569687</td>
<td>50014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>112054</td>
<td>000000121078200</td>
<td>693</td>
<td></td>
<td></td>
</tr>
<tr>
<td>112057</td>
<td>0000000403924698</td>
<td>295</td>
<td></td>
<td></td>
</tr>
<tr>
<td>112148</td>
<td>0000000403934941</td>
<td>50008</td>
<td>502854200</td>
<td></td>
</tr>
<tr>
<td>112150</td>
<td>000000102798114</td>
<td>50021</td>
<td>504547593</td>
<td></td>
</tr>
<tr>
<td>112331</td>
<td>0000000403952787</td>
<td>4627</td>
<td>506346064</td>
<td></td>
</tr>
<tr>
<td>121429</td>
<td>0000000122228143</td>
<td>904</td>
<td>124</td>
<td>506030636</td>
</tr>
</tbody>
</table>

Code DGEEC (HE)

Code FCT (Research units)
Future Work – Phase II

• ORG IDs Management System – workflow proposal

![Diagram showing ORG IDs Management System with labels for DGES, FCT, API PTCRIS, Organisations Management, API Registries, User Management and Access, Workflow, Updates, CV, Grant Mng, BI, ISNI, Ringgold, ORCID, Scholar One, ARIES, International Systems, National Systems, Syncronization.]
(Some CERIF entities and their relationships – adapted from http://proj.badc.rl.ac.uk/moles/attachment/wiki/CRIS/cerif-main-structure.jpg)
Why?

- **Difficult** to associate each researcher’s contribution to a final scientific product
- Almost **impossible** to associate each funder’s contribution to a final scientific product

Contribution of the different players to a final scientific product shouldn’t appear isolated and disconnected
Why?

- Funding is dispersed throughout **multiple databases** (more or less handcrafted)
- Most part, lack an associated PID
How?
How?

Using **Persistent Identifiers**: allowing for search and reuse

Collaborating in **international initiatives**
Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?

Access to information

Connects all players of the Scientific Production Network (federated IDs)

Allows for services interconnectivity

Insert once, Use multiple

Provides monitoring and compliance to OA

Access to information

Funding Database

Connector

National Funding Database

Open Access

DaaS

Insert once, Use multiple

Open Access

What?
What?

Redaction of a perspective study

State of the funding system within the Portuguese research system

- 3Q2017
- 1Q2018
- 2Q2018
- 3Q2018
- 2019
What?

National and International Funders

NATIONAL
- Portugal2020
- FCT
- FCG
- FC

INTERNATIONAL
- EU
- EMBO
- EEA
- NSF
- HFSP

Nonprofit corporations
What?

ETL system

Funding DB

FCT DB

National DBs

International DBs

Extract

Transform

Load

API

MySQL

CERIF
Adopting **P-O-P-F profile of CERIF**: one language, many voices
Establishing anti-telomerase terapeutics for invasive melanoma in zebrafish

Project/scholarship ID
116821

Reference
PTDC/SAU-ONC/116821/2010

FundRef URI
http://www.fct.pt/apoios/projetos/consulta/volegal_proje
to.html?IdProjecto=116821&udEleemConcurso=4172

Funder
FCT - Fundação para a Ciência e a Tecnologia, I.P.

Funder’s country
Portugal

Funding program
3599-PPCDT

Funding amount
179,505,00 €

Start date
2012-02-01

End date
2015-01-31

Date awarded
2011-12-07
Future Perspectives
Now and Onwards

<table>
<thead>
<tr>
<th>State of the Art</th>
<th>Characterization</th>
<th>DB - Phase I</th>
<th>DB - Phase II</th>
<th>Reconciliation</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redaction of a perspective study</td>
<td>Defining guidelines</td>
<td>National funding</td>
<td>International funding</td>
<td>To ecosystem’s DBs</td>
<td>Funding information provider</td>
</tr>
<tr>
<td>State of the funding system within the Portuguese research system</td>
<td>Adoption and adaptation of the P-O-P-F project</td>
<td>Major Portuguese funders</td>
<td>Major international funders</td>
<td>Integration with other PTCRIS services: RCAAP and CIÊNCIAVITAE</td>
<td>DaaS, Search</td>
</tr>
</tbody>
</table>

- **3Q2017**
  - 3Q2017
  - 1Q2018
  - 2Q2018
  - 3Q2018
  - 2019
RCAAP integration with PTCRIS
Some CERIF entities and their relationships – adapted from http://proj.badc.rl.ac.uk/moles/attachment/wiki/CRIS/cerif-main-structure.jpg
Mission: promote, support and facilitate Open access to the national scientific production

Electronic services
- Portal RCAAP

Support services
- 28/52 IRs
- 55 Shared IR
- 16/76 Journals

Communication, dissemination and training
- ~500K docs
- >20M downloads

> 129 Institutions
Why integrate?

• To maximize open access practice by providing appealing, interoperable and innovative electronic services

• To increase the visibility of scientific production

• To facilitate deposit process and information circulation and reporting
How – Strategy

• IR and National Portal adopt PTCRIS regulatory framework (protocols and standards)

• Scenario A (short/medium term)
  – Adapt Dspace 5
  – Adapt RCAAP portal (new)

• Scenario B (medium/long term)
  – Lobby / Direct participation on Dspace 7 roadmap in order to early include CRIS functionalities
How – Integration
IR integration roadmap

July / 2018

- Ciência ID Authentication
- Authors IDs
- OpenAIRE 4.0

November / 2018

- IR deposit from external CRIS systems
- Author Claim
- Organizations IDs
National Portal integration roadmap

July / 2018

- Import from CV based on author ID
- OpenAIRE 4.0 aggregation
- Funding data management
- Author profiles

November / 2018

- Integrated search for entities (P-O-P-F)
- CERIF-XML data exposure
- Ciência ID Authentication
- PTCRIS Sync
Results – DSpace

- IR Deposit with author ID
Results – DSpace

- OpenAIRE 4.0 guidelines
Results – New RCAAP Portal

• OpenAIRE 4.0 guidelines aggregation with authors IDs
Results – New RCAAP Portal

- Data exposure (CERIF POPF profile)

```xml
<Funding xmlns="urn:xmns:nl:pop-project:cerif-profile-1.0-EXPERIMENTAL"
         xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
         xsi:schemaLocation="urn:xmns:nl:pop-project:cerif-profile-1.0-EXPERIMENTAL ./schemas/p-o-p-profile-schema.xsd">
  <Project id="55095">
    <Title xml:lang="en">[From relativistic shocks to ultra intense lasers: igniting discovery in High Energy Density Science]
    <Acronym>POCI/FIS/55095/2004</Acronym>
    <StartDate>Thu Sep 01 01:00:00 WEST 2005</StartDate>
    <Amount>86000.0</Amount>
    <Funders>
      <Funder>
        <OrgUnit id="240624">
          <Acronym>FCT</Acronym>
        <Name xml:lang="en">Foundation for Science and Technology</Name>
      </Funder>
    </Funders>
  </Project>
</Funding>
```
Future work

• Improvements on developed functionalities;
• Dspace 7 development collaboration;
• New services based on systems integration.
(Some CERIF entities and their relationships – adapted from http://proj.badc.rl.ac.uk/moles/attachment/wiki/CRIS/cerif-main-structure.jpg)
Why?
Vison / Misson

Porque a nossa vida é a Ciência

Racionalidade, simplificação, desburocratização

ESTUDANTE ID

CIÊNCIA ID

CIÊNCIA VITAE
Your CIÊNCIAVITAE

Facilitates the management of your scientific curriculum, enhancing your work visibility.

CREATE YOUR ACCOUNT

One CV  Customise  Integrate  Promote
One CV
– Register once, reuse often –

- P2020
- FCT
- Affiliation Management
- HE
- Site
- Observatory
- Thesis jury
- DGES
- A3ES
- Census
- DGEEC
- Inquiries
- Europe
- ERASMUS
- PTCRIS
  - Data as a Service
    - Research Portal
    - Observatory
Your CIÊNCIAVITAE

Facilitates the management of your scientific curriculum, enhancing your work visibility.

CREATE YOUR ACCOUNT

One CV  Customise  Integrate  Promote
Customised CV
– Adapt according to your needs –
Your CIÊNCIAVITAE

Facilitates the management of your scientific curriculum, enhancing your work visibility.

CREATE YOUR ACCOUNT

One CV  Customise  Integrate  Promote
Integrated CV
– Gather your info from multiple sources –
Your CIÊNCIAVITAE

Facilitates the management of your scientific curriculum, enhancing your work visibility.

CREATE YOUR ACCOUNT

One CV  Customise  Integrate  Promote
Public CV

Find CVs

Search in 93 researchers with published curriculum
Minimum of 3 characters for search

Found 2 curricula

Cátia Laranjeira
Cátia Laranjeira.

Cátia Laranjeira teste

View curriculum
Demo – Creating a CIÊNCIA VITAE

Create your CV on CIÊNCIA VITAE

One CV  Customise  Integrate  Promote
Ciência Vitae: quick and easy

39% (22-62%)

10 min (8-19 min)
Conclusions

• Managing research information effectively = managing CRIS ecosystem
• A CRIS ecosystem requires solid foundations (CRIS for managing basic entities)
• A regulatory framework is key to insure integration / interoperability
• Standards organizations are key for PTCRIS
• The definition and adoption of a regulatory framework can be made in a step by step approach
• Kits and frameworks can speed up its adoption
Future work

• Release national CV platform: Ciência Vitae
• Continue the development of structural pillars
• Move on with:
  – regulatory framework
  – adoption in national CRIS
  – adoption in local CRIS
• Develop new services
  – Research Portal (VIVO?)
  – Business Intelligence
Questions ?