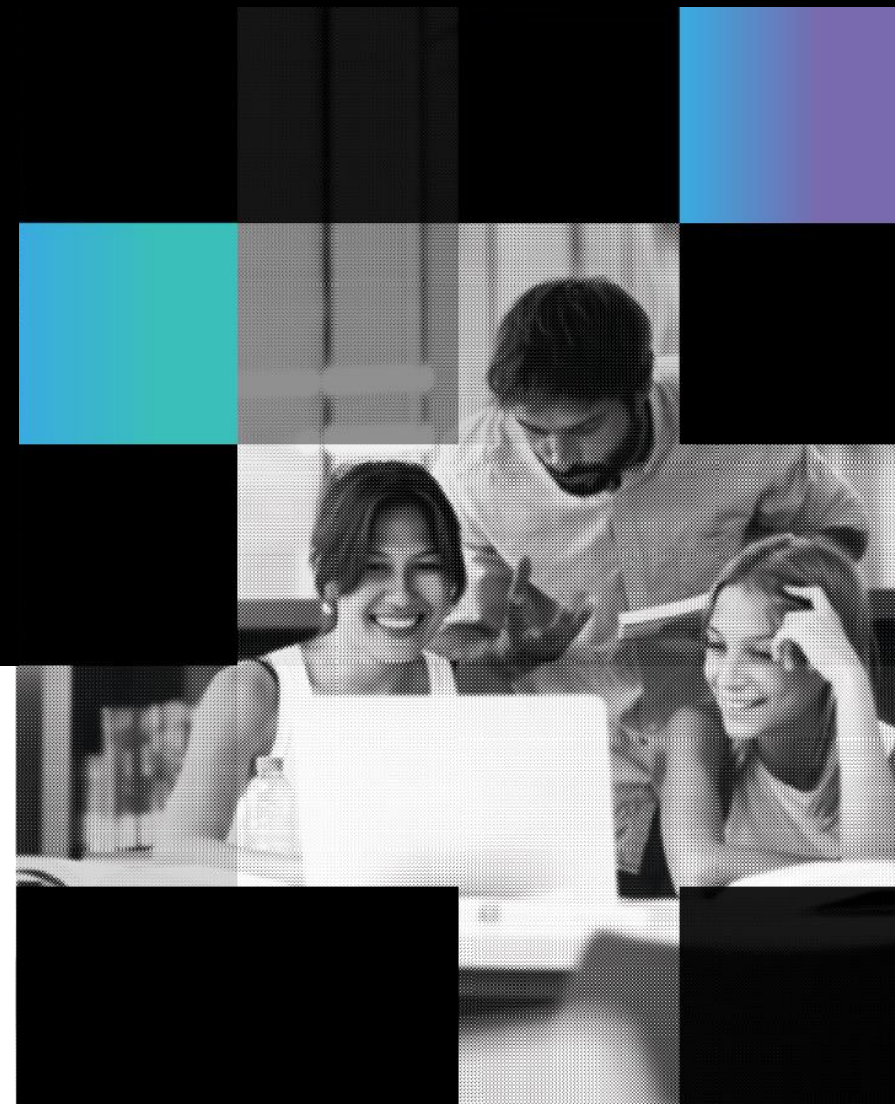


Tecnologia para
o conhecimento

Migrating DSpace from 5x to 7x: How we did it in our national open access service - RCAAP

Paulo Graça (1); Fernando Ribeiro (1); Raquel Truta (2); Ricardo Saraiva (2); Paulo Lopes (1)
(1 – FCT/FCCN; 2- University of Minho)



Repositório Aberto

Universidade Aberta Scientific Repository

[Search](#)

Sobre o Repositório Aberto

O Repositório Aberto é o repositório institucional da Universidade Aberta e foi construído com o objetivo de armazenar, preservar, divulgar e dar acesso à produção intelectual da UAb em formato digital. Pretende-se reunir, num único sítio, o conjunto das publicações científicas da UAb contribuindo desse modo para o aumento da sua visibilidade e impacto e garantindo a preservação da sua memória intelectual.



Who we are

FCCN - Digital services from the Foundation for Science and Technology (FCT), aims to **contribute to the development of Science, Technology and Knowledge in Portugal.**

Provide a set of shared **digital services** that aim to support the development of teaching and science in Portugal, contributing to the excellence of national teaching and research.



Communities served

+600 000

Students, teachers,
researchers and staff



SESP

Portuguese Higher
Education System

RAE

Extended Network
from Education

**Public
administration**

SNCT

National Science
and Technology
System

MECI

Ministry of Education,
Science and Innovation





Framework

Action pillars



Connectivity



Security



Knowledge



Computing



Colaboration



Inovation





fccn.pt/conhecimento/rcaap

The national open access initiative

Service that aims to **store, preserve and promote access to scientific knowledge produced in Portugal.**

RCAAP is managed and executed by FCCN (FCT's unit), with the support of the University of Minho.

Numbers in 2023

Entities with
repository

146

Aggregated
documents

965 878

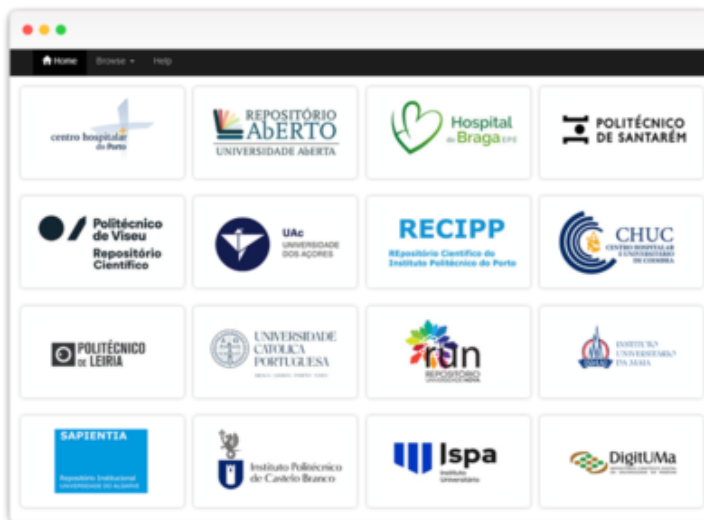
Journals

225

Downloads

> 26 Millions

About SARI – Institutional Repository Hosting Service



- **Started in 2008** with 5 repositories based on **DSpace** version 1.8 provided on a **Software as a Service** model;
- Currently has **28 hosted** repositories (one of them represents more than **98 institutions**);
- Each one with their **own** individual corporate **identity**;
- Configurations and parameters according each institutional structure and policy of self-archiving;



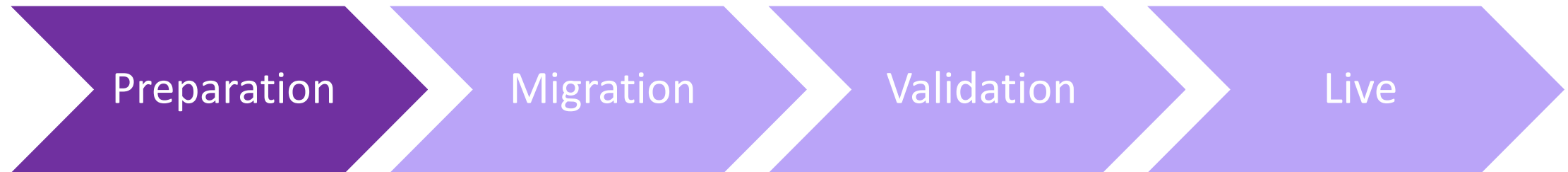
Migrating from Dspace 5X to 7X



Motivation

- DSpace 5X and 6X are no longer supported
- New solution available – DSpace reshaped
- Obsolete infrastructure (old CentOS 6.7) – new available
- Docker opportunity brings isolation and less hassle with dependencies
- PTCRIS National interoperability requirements

Migration plan - methodology



Preparation - DSpace contributions

Pages / ... / DSpace Working Groups

DSpace 7 Entities Working Group (2018-19)

Created by Tim Donohue, last modified on Jul 03, 2023

The screenshot shows the GitHub repository page for 'DSpace / DSpace'. The repository is public and has 40 issues and 112 pull requests. The page also shows a navigation bar with links to Product, Solutions, Open Source, and a pull request link. The repository description mentions the DSpace 7.0 release and ongoing development.



39

Created Pull requests

Deliverables

This working group will

1. Assess the **early prototype of Configurable Entities** and its alignment with guidelines/needs.
2. Define an implementation roadmap for DSpace 7 Configurable Entities (based on the assessment of early prototype)
3. Implement Configurable Entities, per defined roadmap (by creating a Pull Request & documentation, to be reviewed by DSpace 7 Working Group (2016-2023) and DSpace
4. Draft a post-DSpace 7 roadmap for Configurable Entities (to be reviewed by Steering and Leadership)

Once the above deliverables have been completed, this working group will merge with the DSpace 7 Working Group (2016-2023). At that point, the Configurable Entities



Preparation

- **DSpace 7++ Package**
- Infrastructure Setup
- Test Environment Preparation
- Definition of SARI's specificities
 - General Configurations
 - Interface Definitions
 - Integration Settings (Sherpa, Ciencia Vitae, OpenAIRE API, CiencialD, RENATES)
 - User Migration to CiencialD (Shibboleth authentication)

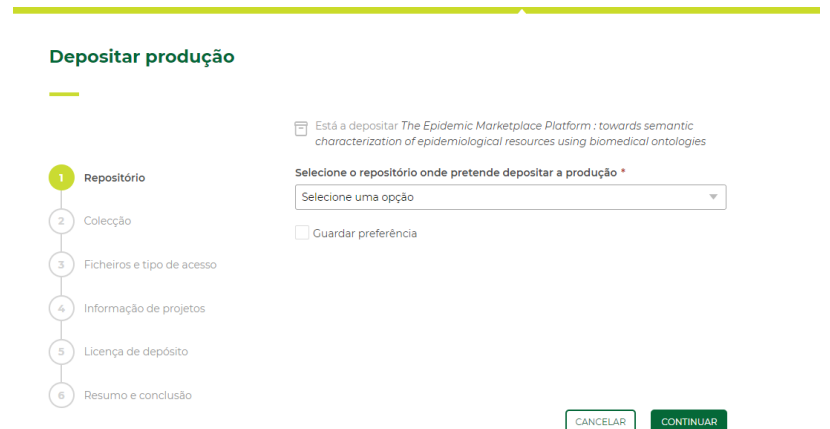
DSpace 7++

- Code on **DSpace additions module**
- Supported Languages (PT and EN)
- Specific Controlled Vocabularies
- **Entity oriented** based on OpenAIRE 4 guidelines
- Specific forms:
 - Publications
 - Thesis
 - Learning objects
 - Persons
 - Organizations (funders)
 - Projects
- COAR Access rights based on DSpace's resource policies
- CienicalD authentication (Shibboleth)



Dspace 7++ - integrations

- ReCAPTCHA (Google v3)
- Google Analytics V4
- ORCID public API (people search)
- Ciencia Vitae API (people search)
- Ciencia Vitae deposit (via Swordv2 with specific field mapping)
- Sherpa Romeo
- OAI (OpenAIREv4 schema)
- RENATES (Thesis and Dissertations integration)



Depositar produção

Está a depositar *The Epidemic Marketplace Platform : towards semantic characterization of epidemiological resources using biomedical ontologies*

Selecione o repositório onde pretende depositar a produção *

Selecione uma opção

☐ Guardar preferência

CANCELAR CONTINUAR



Repositório Aberto

Repositório Institucional da Universidade Aberta

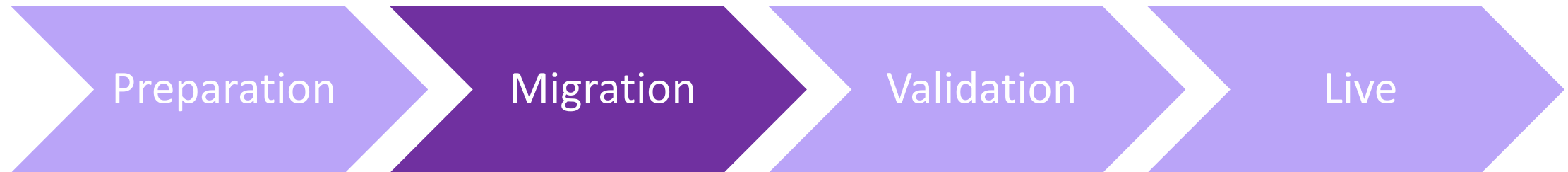
pesquisar no repositório...

Pesquisar

Sobre o Repositório Aberto

O Repositório Aberto é o repositório institucional da Universidade Aberta e foi construído com o objetivo de armazenar, preservar, divulgar e dar acesso à produção intelectual da UAb em formato digital. Pretende-se reunir, num único sítio, o conjunto das publicações científicas da UAb contribuindo desse modo para o aumento da sua visibilidade e impacto e garantindo a preservação da sua memória intelectual.

Migration plan - methodology



DSpace 7 new architecture

- Each SARI is based on **microservices** (Docker containers) contained in a single Virtual Machine instance:
 - Content
 - Database Server
 - Frontend Server
 - Backend Server
 - Index Server



DSpace 7 migration – Migration phase

- Content Migration
 - Assetstore (Files) + thumbnails
 - UMinho Stats to DSpace Solr Stats
 - Licenses
 - Database (through FlywayDB scripts)
 - Existing DS5 items will be considered Publications
 - Entities of type Person created (from Author Profile)
 - Entities of type Project created (from OpenAIRE 3 Guidelines)
 - Entities of type Organization created
 - Migration to new metadata field types and schemas (degois.publication.title to oaire.citation.title)
 - New vocabularies:
 - COAR types
 - OECD - Sustainable Development Goals
 - OECD - Fields of Science

DSpace 7 migration – Content migration

Pesquisa na autoridade de nomes

Procurar por identificador

2D18-A6F4-20B9

Lista de Resultados

Graça, Paulo
'2D18-A6F4-20B9' (não em Autoridade)

Informação sobre o autor



Nome
Graça, Paulo

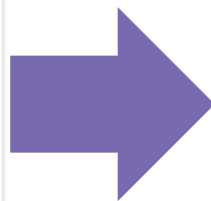
Variantes

Afiliação

Identificadores
2D18-A6F4-20B9  0000-0002-3503-4812 

[Apresentar mais resultados](#)

[Aceitar](#) [Cancelar](#)



RCAAP
Repositório Científico de Acesso Aberto de Portugal

Comunidades & Coleções

Importar pessoa de

Paulo Graça

Resultados da pesquisa

A mostrar 1 - 1 de 1

 Graça, Paulo
<https://qa.cien>

Pré-visualização dos dados pessoais a importar

Os seguintes metadados foram importados de uma fonte externa. Serão pré-preenchidos quando iniciar o depósito.

Identificador:
<https://qa.cien/vitae.pt/6614-C5AE-3CF7>

Título:
Graça, Paulo

Acrónimo:
Paulo Graça

Apelido:
Graça

Nome:
Paulo

CIÊNCIA ID
6614-C5AE-3CF7

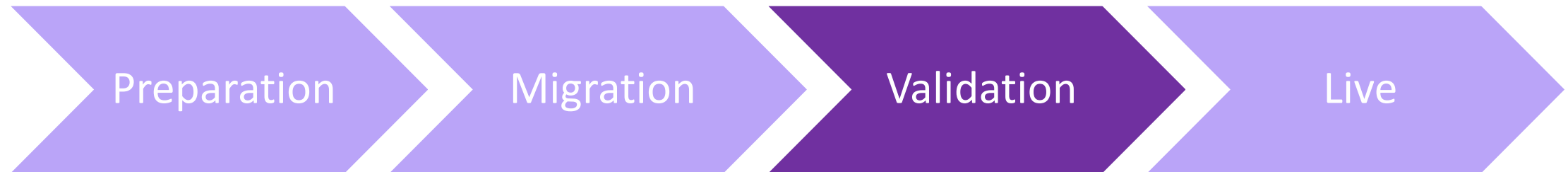
ORCID:
0000-0002-3503-4812

Researcher ID
A-1009-2008

Scopus Author ID
12345678900

[Iniciar depósito](#)

Migration plan - metodologia





DSpace 7 Validation phase

- Validation tasks
 - Rely on base DSpace 7 tests
 - Running our build tests (currently based on github actions)
 - Comparing content verifications before and after
 - Perform our regression tests
 - Perform security tests
 - Perform load tests
- Migration **Test Pilot** with 2 hosted institutions in a controlled environment

>> Any feedback was used to improve the process and, also, DSpace



DSpace 7 testing Pilot - Purposes

- Validating Migration Processes;
- Validating the Integrity of Migrated Information;
- Validating the Key Backend Interface Functionalities (with repository managers);
- Validating the Frontend Interface;
- Validating the Suitability of the New Platform for Internal Institutional Processes.



DSpace 7 testing Pilot – Results

- Significant **improvement** in the **graphical interface**, making it visually more appealing and functional
- **Enhancement of management** and search capabilities by Entities: Publications | People | Projects | Funding.
- **Enhancement of searches**, making them more comprehensive, rigorous, and intuitive, with the ability to filter by community or collection.





DSpace 7 testing Pilot – Feedback

- **Need** for more **comprehensive statistics** than those provided by DSpace 7.
- Occasional system **slowness**, although improvements have been made with cache management configurations to enhance performance.
- **Need** for **additional tools** for extraction of metadata from imported PDFs
- and cross-reference Projects / Funders with authors (extend relationships)

Migration plan - methodology





Live plan

Calendário 2024

	S	D	S	T	Q	Q	S	S	D	S	T	Q	Q	S	S	D	S	T	Q	Q	S	S	D	S	T	Q	Q	S	S	D	S	T	Q	Q	S	S	D	
Janeiro			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31					
Fevereiro						1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29				
Março							1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
Abril			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30						
Maiο					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31			
Junho	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30								
Julho			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31					
Agosto						1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31		
Setembro		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30							
Outubro				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31				
Novembro							1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
Dezembro		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31						

Lote 1
Lote 2
Lote 3
Lote 4
Lote 5
Lote 6



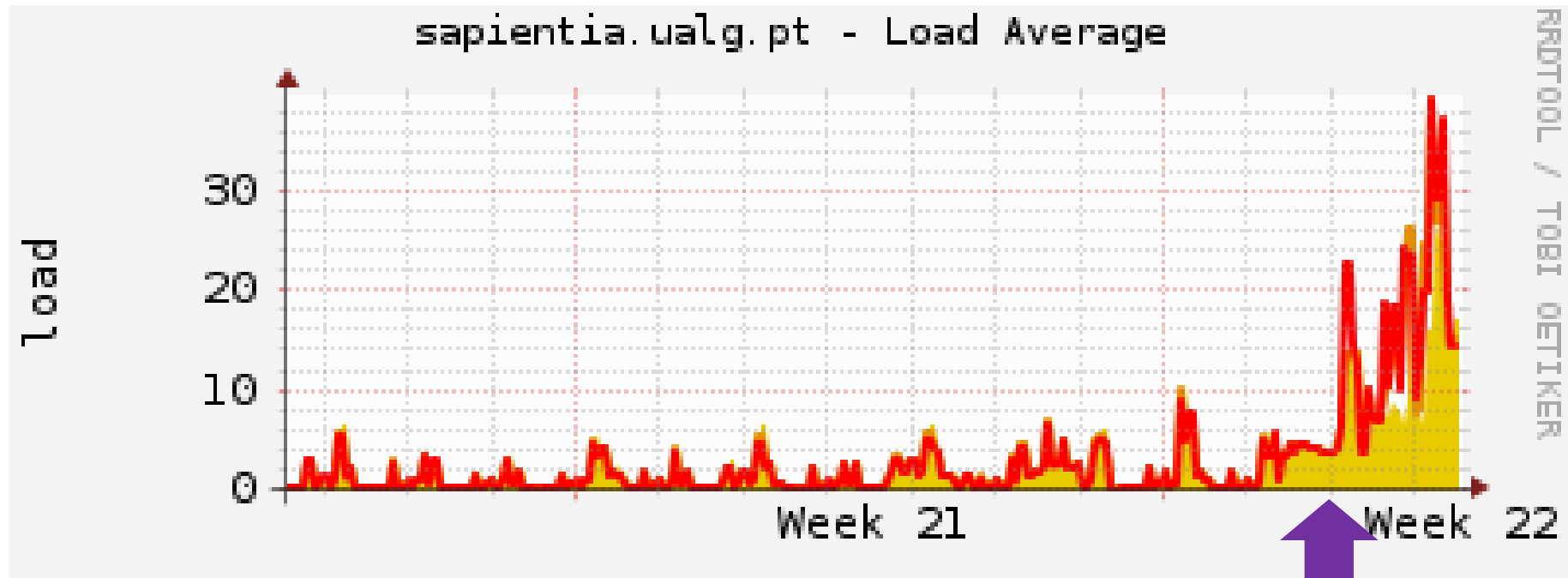
Conclusions and future work



It's been a long journey



But it's not over yet



When migration occurred



We shared with you our journey and we will share our resources

RCAAP's DSpace angular theme

Instalation

Clone DSpace Angular version 7.6.1 project (replace [dspace-angular-7X] with dspace-7.6.1, the specific DSpace Angular version) source directory on your system.

```
git clone --branch=[dspace-angular-7X] https://github.com/dspace/dspace-angular.git dspace-angular
```

Place the Theme content at your dspace-angular/src/themes/rcaap folder:

```
cd dspace-angular/src/themes
git clone https://github.com/fccn/rcaap-dspace-angular-theme.git rcaap
```

Also, to activate the theme, you should replace the Theme's provided angular file, so, firstly we remove current files:

```
mv ../../angular.json /tmp/
```

we replace them with the provided files:

```
mv rcaap/.config/angular.json ../../angular.json
```

We also need to replace some of the source base components and replace the file [dspace-angular]/src/themes/eager-themes.module.ts in order to add the required features:

```
rsync -r --remove-source-files rcaap/.src/ ../
```



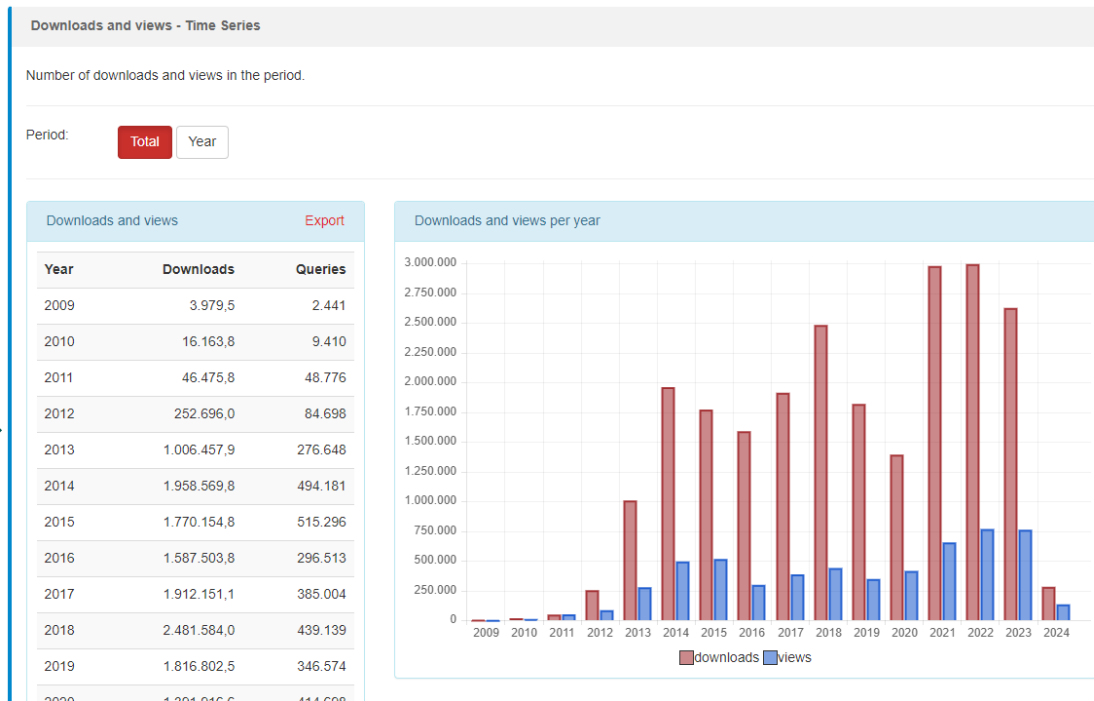
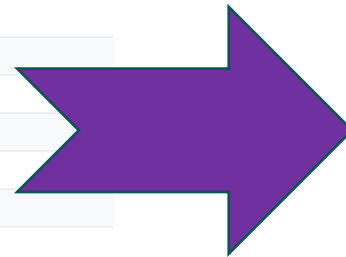
Future plans – Help to improve DSpace statistics

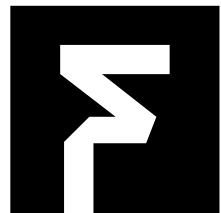
Home · Statistics

Statistics for Repositório Aberto

Total visits

	views
Transitando de um ensino remoto emergencial para uma educação digital em rede, em tempos de pandemia	40427
Modelo pedagógico virtual da Universidade Aberta : para uma universidade do futuro	36860
Conceitos e exercícios de programação: utilizando Linguagem C	19866
Citar e referenciar : APA 6ª ed.	18332
Condições de vida e de trabalho na Inglaterra da Revolução Industrial	17305
Paradigma qualitativo e práticas de investigação educacional	14958
Como vender a alma ao diabo em 4 lições ou o diabo como tentador da humanidade	14603
O contributo da auditoria interna para uma gestão eficaz	14060
APA 7th: diretrizes para a realização de citações em texto e referências bibliográficas de acordo com o Publication Manual of the American Psychological Association (7th edition)	13320
Metodologia da investigação : guia para auto-aprendizagem	13102





Tecnologia para
o conhecimento

Thanks!

Paulo Graça (FCT|FCCN)

<https://www.fccn.pt/>

<https://projeto.rcaap.pt/sari/>

<https://www.rcaap.pt/index.jsp?locale=en>

