

# J. Barbosa

## Master in Informatics and Computing Engineering

### UPBox: Armazenamento na Nuvem para Dados de Investigação da U. Porto

José Pedro Marques Barbosa

---

## Abstract

The new digital technologies boosted the generation of scientific data, on one hand, due to the digital storage capacity, on the other, due to the evolution of methods and research tools. The sharing of scientific data is essential in the discovery process and represents the dominant means by which researchers can earn credits for their work.

Several entities, such as universities and research communities, provide some infrastructures and services to improve data sharing and curation. However, the data deposit process on these infrastructures is complex and requires the joint effort of curators and researchers, whereby much of the generated research data doesn't reach these repositories.

The purpose of this work is to promote the sharing of research data in the University of Porto, bringing the researchers closer to the data curation process. This dissertation proposes a different approach to data curation, encouraging the participation of researchers in collaborative management of their research data. The adoption of the proposed methods will, in the future, accelerate and automate the submission and curation process into the University of Porto's repository.

According to the results of an audit conducted at University of Porto in order to survey the researcher's practices in the management of their data, it was verified that some researchers manage and share their research data through cloud storage applications like Dropbox.

The proposed solution, UPBox, is a service that supports the collaborative management of research data during the research process, keeping researchers in control of their data sets. This cloud storage service allows the researcher to annotate their data and, when appropriate, submit them for curation in order to be available in the University of Porto's repository. This simple and familiar platform works as a proxy between the researcher and the data repository, accelerating the submission process into the repository.

UPBox is a web application that allows the creation of data workspaces structured in directories and sharing them with other researchers. It is integrated with an annotation system, developed in a parallel project, DataNotes, which enables the annotation of files and directories.

To validate the acceptance of the solution, a group of researchers were requested to conduct a usability test. The test has evaluated the available features, identified improvements and new features and provided insight on the usefulness of the the developed prototype.

## Resumo

As novas tecnologias digitais impulsionaram a geração de dados científicos, por um lado devido à capacidade de armazenamento digital, por outro devido à evolução de métodos e ferramentas de investigação. A partilha de dados científicos é essencial no processo de novas descobertas e é a forma pela qual os investigadores ganham reputação pelo seu trabalho.

Várias entidades, como por exemplo universidades e comunidades de investigação, disponibilizam serviços e infraestruturas para auxiliar a curadoria e partilha de dados. Contudo, o processo de depósito de dados nestas infraestruturas é complexo e exige um esforço conjunto de curadores e investigadores, pelo que muitos dos dados de investigação gerados acabam por não chegar a estes repositórios.

O objetivo deste trabalho é promover a partilha de dados de investigação na Universidade do Porto, aproximando os investigadores do processo de curadoria. Esta dissertação propõe uma abordagem diferente à curadoria de dados, fomentando a participação dos investigadores na gestão colaborativa dos seus dados de investigação. A adoção dos métodos aqui propostos permitirá, no futuro, agilizar e automatizar o processo de curadoria e submissão no repositório da Universidade do Porto.

Segundo os resultados de uma auditoria feita na Universidade do Porto com a finalidade de efetuar um levantamento das práticas dos investigadores na gestão dos seus dados, verificou-se que alguns investigadores gerem e partilham os seus dados de investigação através de aplicações de armazenamento na nuvem, como por exemplo a Dropbox.

A solução proposta, UPBox, é um serviço que suporta a gestão colaborativa de dados de investigação durante todo o processo de investigação, mantendo os investigadores em controlo da organização dos seus conjuntos de dados. Este serviço de armazenamento de dados de investigação na nuvem permite ao investigador anotar os seus dados e, quando pertinente, submetê-los para curadoria, com vista a serem disponibilizados no repositório de dados da Universidade do Porto. Esta plataforma simples e familiar aos investigadores servirá, então, como intermediário entre o investigador e o repositório de dados, permitindo agilizar e facilitar todo o processo de depósito.

A UPBox é uma aplicação web que permite a um investigador criar áreas de trabalho por projetos, estruturadas em diretórios, e partilhá-las com outros investigadores. Está integrada com um sistema de anotação, desenvolvido num projeto paralelo, o DataNotes, que permite anotar ficheiros e diretórios.

Com o objetivo de validar a aceitação da solução, foram efetuados testes de usabilidade com investigadores. Estes testes permitiram avaliar funcionalidades e identificar melhorias a realizar, bem como tirar conclusões acerca da utilidade do protótipo desenvolvido.

## Jury

- Chair: Rui Carlos Camacho de Sousa Ferreira da Silva
- External Examiner: José Luís Brinquete Borbinha
- Supervisor: Maria Cristina de Carvalho Alves Ribeiro
- Date: 8/2/2013

From:  
<https://web.fe.up.pt/~jlopes/> - JCL

Permanent link:  
<https://web.fe.up.pt/~jlopes/doku.php/students/201302j-pedro>

Last update: **01/07/2013 11:11**

