LABDRIVE Tutorial

A Research Data Management and Digital Preservation Platform

|  |  |  |
| --- | --- | --- |
| **Teo Redondo** | **Antonio G Martinez** | **Maria Fuertes** |
| *LIBNOVA**Spain**teo.redondo@libnova.com**0000-0001-6465-7771* | *LIBNOVA**Spain**a.guillermo@libnova.com* | *LIBNOVA**Spain**mfuertes@libnova.com* |
|  |  |  |

**Abstract – LABDRIVE is a Research Data Management and Digital Preservation platform resulting from the ARCHIVER Project. It allows organizations to capture the research data they produce, helping them to properly manage, preserve and allow access to it, during the whole research data lifecycle.**

**The purpose of this tutorial is to introduce the main features of LABDRIVE as well as explain how it works through a tutorial (a guided demonstration).**

**Keywords – Research Data Management, Digital Preservation**

**Conference Topics – Innovation; Exchange.**

# Introduction

LABDRIVE [1] is a Research Data Management and Digital Preservation platform powered by LIBNOVA that focuses on scientific datasets. LABDRIVE allows organizations to transition from a siloed approach in which each series of datasets, departments or units is using multiple, disaggregated systems to keep content to a single repository that can adapt to the particularities of each dataset, unifying all content into a single platform.

The platform works for organizations both with a few gigabytes of data, to organizations managing several petabytes. Digital preservation principles are always present, so Data protection comes first. The platform is fully aligned with OAIS, ISO16363, and presents a variety of redundant checks and processes for safeguarding valuable research data.

LABDRIVE is primarily oriented towards research-intensive scientific and academic institutions that need to preserve research projects, working objects as well as associated tools (datasets, software tools, etc.).

With LABDRIVE, R&D organizations can keep the research data they produce for the long term, in a single platform. Researchers can manage their research datasets with the best tools, adopting good practices for digital preservation and also keeping code and data together in one single platform during the lifecycle, independently of functionality, protocols and featured needs.

# Main Characteristics of LABDRIVE

* **Metadata-driven, virtualized scalable storage**
	+ Digital Preservation Administrators can assign a specific Storage Policy to each Data Container in the platform (storage types, replicas, technologies, providers and integrity policies) to use at data container level.
	+ A single repository supports multiple storage providers and types (for very high volumes of content).
	+ Transition from one storage policy to another (even from a storage provider to another), fully managed by the platform.
	+ Virtualized storage so file paths remain unchanged when the underlying storage technology is changed.
	+ Extensible storage architecture (cloud object storage, CEPH, tapes, etc.).
* **Code-driven, advanced content management**
	+ LABDRIVE lambda functions can be defined by the organizations (or integrators) so the platform automatically processes the content using the logic defined.
* **Easy to use and powerful**
	+ Equally capable web interface and API, so users can easily manage the platform while power users can automate every process.
* **Strong digital preservation technology**
	+ Digital preservation principles always present: Data protection comes first.
	+ Fully aligned with OAIS, ISO16363, redundant checks and safe processes.

#### Tutorial Content

The contents would be divided into 3 blocks and would be roughly as follows:

1. **LABDRIVE Introduction**
* Architecture and overview
* How research content is to be organized
1. **LABDRIVE Configuration**
* Users and permissions
* Archival organization
* Container – concept and usage
* Metadata configuration
1. **LABDRIVE Operations**
* Create a data container
* Upload content
* Download content
* Introduction to metadata – concept and usage
* Searching
* File versioning and recovery
* Working with data containers
* LABDRIVE functions
* Storage mode transitions
* Advanced operations – Jupyter Notebooks & API usage

# REFERENCES

1. LIBNOVA LABDRIVE Public Documentation https://docs.libnova.com/labdrive