## A proposal for Transversal Computer-related Strategies & Services

### for Scientific and Training efforts to the Latin American Strategy for Research Infrastructures

 $\mathsf{WP} \rightarrow \mathsf{https://doi.org/10.5281/zenodo.3614109}$ 

Arturo Sánchez, Università degli Studi di Udine, ICTP, INFN & ATLAS @CERN July 6, 2020

## Introduction

This WP looks to start a discussion and **to address as early as** possible aspects that are of significance for almost any scientific endeavour in terms of Computation and Science Reproducibility

It is looking to give a view of the different and vital services, protocols, tools and know-how relative to the Scientific Computing (SC) and Information Technology (IT) **for scientific endeavours and capacity-building projects in Latin America** 

The proposal of a transversal services and protocols targets the design, development, deployment, distribution, training, publication and citation, proper accreditation and dissemination of scientific experiments, data, processes, software, documentation, results and resources using world-leading protocols and industrial standards under the Open Access philosophy  $\downarrow$ 



Mexico

# **Review of ideas**

A review of scenarios and proposals that can be seen under an umbrella of a **"SC+IT Hub"** 

Also, considering effective usage of current hybrid spaces (physical and virtual) that contains well-known industrial and academic resources and practical ideas

And how to deploy those ideas at current and future experiments and research teams in the region.



**Figure:** A simplified view of multiple experiments and the ways they *currently* access services and tools in the areas of Scientific Computing (SC) and Informatics Technology (IT).

## Review of ideas: Harmonisation I





#### The top of the production chain

can be harmonised with the usage of common services for the publication of the final datasets, tools and results. Something that is almost field-independent nowadays

many of these common resources are already in place. Open Access, Open Source, Open Data, Open Software and Open Hardware define current and future scientific endeavours. **This environment is beneficial for long terms and fully supports Open "X" projects that we can also profit and contribute** 



#### The bottom of the production chain

can be harmonised in multiple experiments. Also the ways they can access services and tools through common Service Provider Platforms. Common in large-scale experiments, and well-known industries

# Harmonisation II

Using global industrials standards by design. Education and Training will be crucial at the moment of keeping those protocols, but it will pay off in the long run

As result of those implementations (gradual for current experiences and by design in new ones) The reproducibility of science by use of Open Standards Access is a relevant goal also

Using services developed on public and/or generic platforms. A current new field for capacity build and jobs/employment creation is easily imaginable →

Important to mention that the deployment, support, consultancy and delivery of computational services (notably Cloud Computing, Big Data, storage, DOI & licence, ML and AI as a Service) are a very profitable and growing industry.

**Figure**: A resulting simplified view of multiple experiments and the ways they can access services and tools through common Service Provider Platforms in the areas of SC and IT highlighting the area of usage of ALL needed services.



### Reproductivity of the Science's Reproducibility



#### **Publications**

- Publication systems are evolving but at the same time, they are well standardised. So, a system of Publication as a Service (PaaS) is relevant for small and medium sizes research groups to get their results out as efficient as possible.
- This includes the revision, edition and event the design of the articles to have the best impact.

#### **DOI & Access**

- DOIs are absolutely essential for the modern track of scientific and technical contributions.
- They will be a measurement of the impact of the research and in a modern way, where final papers and patterns are not the best way to release software tools, developed protocols, preliminary findings, blueprints and documentation.

#### Repositories

- Local, regional, Cloud and "Cold" Repositories for final datasets, Jupyter notebooks, software pipelines, Docker containers and Virtual Machines, etc.
- They also are relevant for the documents when other publication media is not suitable due to experiment embargos, privacy legislation and cybersecurity.

**Common Services and Infrastructure**