

Help us build a vibrant, collaborative global commons

Donate

[Creative Commons](#) > [Blog](#) > [Open science](#) > Making data and tools available for the world to see: Arturo Sanchez of CERN on why ATLAS uses CC0 data

Find news by topic

Making data and tools available for the world to see: Arturo Sanchez of CERN on why ATLAS uses CC0 data



Jennie Rose Halperin and Noam Prywes
November 2, 2016

ATLAS Open Data website introduction video



According to [Arturo Rodolfo Sanchez](#), a member of the [ATLAS](#) community and outreach team, “The [large hadron collider](#) is running now at 13 TeV. This is an energy level never seen before in a collider.” This exciting development is built on the power of open science – at ATLAS, data sharing and an open, innovative approach to information collaboration has become a fundamental part of this important scientific community.

This year, ATLAS [decided to release the data from 100 trillion proton-proton particles to the public under CC0](#), the first release of 8 TeV data. More than 3000 scientists from 174 countries work on ATLAS, and more are joining every day. At the CERN open data platform educational portal, scientists, educators, and science enthusiasts can access the work of thousands of scientists working together to hunt for the Higgs-Boson particle and other important scientific discoveries.

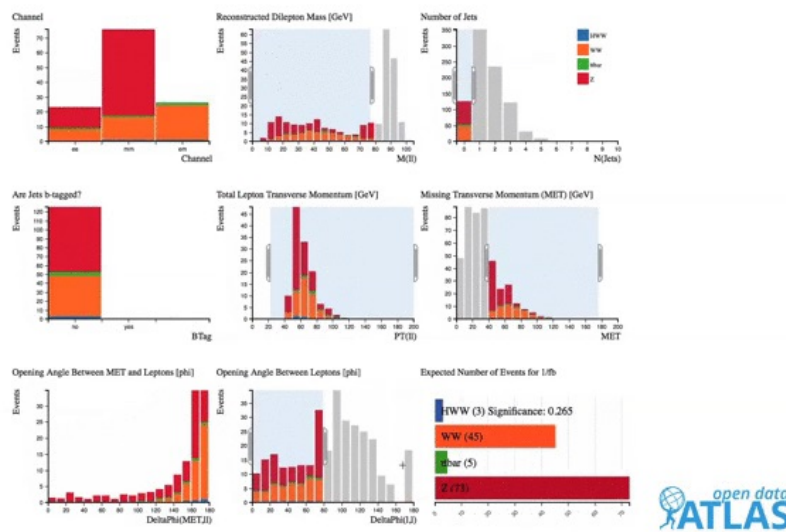
Sanchez’s vision of science is open, and he believes that CERN’s is as well – working with Creative Commons, he describes a new kind of research organization built with the power of community. Though the 7000 ton ATLAS detector in the large hadron collider lives “[100 meters below a small Swiss village](#),” the data moves far beyond the confines of the institution, providing insights and experimentation to the entire world.

This interview was conducted with the assistance of [Noam Prywes](#), a post-doctoral fellow at Massachusetts General Hospital.

Why is open data and open science important to CERN? Why have you chosen to use CC0 for this dataset in particular?

Open Data, open software and open hardware are very important for us! It is part of our policy in the [ATLAS Collaboration](#) and the other [Large Hadron Collider \(LHC\) experiments](#). This is important for us because we are a scientific community and our main goal is to look for answers as humankind, not as an institution. We are also funded by taxpayers – CERN as an organisation and facility, and the experiments like ATLAS ([part of the LHC](#)) use public sector funding.

Independently of the [member country](#), most of them have as policy/law to publicly release any final result, publication, dataset, and conclusion that public funding research institutions generate. In ATLAS, we develop resources (datasets and tools) that can be used mainly for educational projects carrying out by ATLAS and not-ATLAS members. Of course, this is not a restriction! We don't want to limit the use that a person (educator, scientific, artist, etc...) could have with the data.



The "Histogram Analyser" allows to make data cuts and selections directly from the browser. (Image: ATLAS Experiment/CERN)

There are a lot of people out there with many different ways of thinking, so who knows what can be possible or not possible with those resources? This is why we went for the CC0 license for the datasets released by ATLAS on its Open Data project. The same has been done by the CERN Open Data project. I can complement my answer by mentioning several projects from CERN or CERN groups:

- [Root](#): A modular scientific software framework
- [CERN Open Data project](#) for the release of experiment data and resources
- [ATLAS OpenData and educational resources](#), our platform!
- [CERN Open Hardware Project](#)
- [Invenio](#) and [Indico](#): Very successful projects born at CERN: and
- [Open Lab](#): Another training and educational project where public and private sector come together

What's the relationship between your initiative and other open data and open access initiatives in scientific communities? How are you working together? Is there anything unique about your relationship to open access that's different from other open science initiatives?

As you can see, the CERN community is keen on the involvement of a high number of people, countries, institutions and research fields involved. Therefore, any project that includes two or more groups working at CERN or in CERN-hosted experiments is already an international enterprise!

Let me give you the ATLAS example: we are an experiment with ~3000 members coming from more than 120 universities around the world. Many of outreach activities. This leaves us with the possibility of having a professor in a North American university using public data to write some code to train her new master's student. At the same time, an ATLAS college in a German university is running a complete laboratory course in particle physics using the ATLAS public data together with a combination of public software and custom code. Meanwhile, a group of Latin-American ATLAS members are presenting public seminars and running exercises for high school students using public apps and public ATLAS data.



ATLAS experiment detector under construction in October 2004 in its experimental pit; the current status of construction can be seen on the CERN website.[1] Note the people in the background, for comparison. [Nikolai Schwerg CC BY-SA 3.0](#)

Coming back to your question, we are working together with other communities and sharing as much as we can! Different communities in the high energy physics (HEP) sector have meetings and conferences to share their experiences, knowledge, and research with other teams. I don't think there is anything unique in the way we are doing Open Data and Open Source, in fact, it is this constant feedback between communities that helps to find common frameworks, platforms and even ways to develop and deploy resources. Our community is global and our audience is global, but the approach is in fact local. It is important for us to understand the difficulties and limitations in each region: it is not the same to teach HEP to students in the United States to those in Venezuela. The languages, resources, culture, and differences in the academic systems are now part of our fine tuning when writing projects and documentation.

Since CERN is so international, how do you choose how you release data and publish research? Is open access a more acute concern because of national boundaries? What about funding sources? Are there countries that demand open access as a precondition for money? Has that influenced scientists from different locales?

The way to release data is in a worldwide common framework: on a web platform, with a lot of files to create the best documentation possible.

holes and make the web and user interfaces better every day.

The fact that CERN is a multinational organisation with so many funding governments and institutions consolidate the openness of the research and the resources products of those. Many legal aspects are taking into account and I am do not know all the [details](#), but the spirit is to share and be as useful as possible.

CERN is in such an individual position in terms of the science it does, so what kind of innovative measures are you taking to publicize this science? How are you highlighting the work that scientists and communities are doing with the published data?

We have been working very hard in the communication side by using every possible media out there to communicate results, activities, tutorials, and even how physicists spend their time. This is done by the CERN community and included in each of the experiments now. Our presence in social media is strong (at least for a scientific community!) and more and more people are aware of what we do and why it is important. Students around the world come to visit CERN and the experiments, and [some others visit the place virtually](#). In the case of the data, the challenge right now is to use the power of the media and the web in order to explain how to use it. Developing easy but still powerful user Interfaces is the key! With a lot of energy and ideas we are trying to reach more people every day, even with the limited resources that we have.

I am reaching the end with the beginning of this story – the ATLAS Open Data platform. In the outreach group we are learning and developing tools and protocols that help us disseminate the data publicly, trying to prove to ourselves and the members of the experiment that there is interest to use those datasets and resources by the international community.

Our aim is getting more data out there! We want to make that data and tools available for the world to see.

CATEGORY: [Open science](#)

SHARE: [!\[\]\(e78f798d4ea5c530c9db49e7d26e6b95_img.jpg\)](#) [!\[\]\(034433b90593e82e5460e34e3ed48e9b_img.jpg\)](#) [!\[\]\(5f24500834b50a8307ffe63e419281a9_img.jpg\)](#) [!\[\]\(8502790a2fc8970abb55aa7f7a7a6bae_img.jpg\)](#)

Thank you to our supporters!

THE WILLIAM AND FLORA
HEWLETT
FOUNDATION



THE BRIN WOJCICKI
FOUNDATION

Google



and thousands of donors like you!

[SEE ALL SUPPORTERS](#) ▶



Except where otherwise [noted](#), content on this site is licensed under a [Creative Commons Attribution 4.0 International license](#). [Icons](#) by The Noun Project.

[Contact](#) | [Privacy](#) | [Policies](#) | [Terms](#)



