PROCEEDINGS OF SCIENCE

The search for new long-lived particles at CERN with the REINFORCE EU Citizen-Science project

C.Kourkoumelis*

Department of Physic, National and Kapodistrian University of Athens, Campus,157884 Zografou, Athens ,Greece

E-mail: ckourk@cern.ch

The REINFORCE EU project (Research Infrastructures FOR Citizens in Europe) engages and supports citizens to cooperate with researchers and actively contribute to the development of new knowledge for the needs of science and society. The overall aim is to bridge the gap between them, and reinforce society's science capital.
REINFORCE targets citizens of any age, without requiring prior physics or computing knowledge, and offers four "discovery demonstrators" from different areas of physics: Gravitational Wave Astronomy, Neutrino Physics, High-Energy Collider Physics, and the interplay of Cosmic Rays with Geoscience and Archaeology. The infrastructure of all demonstrators is based on Zooniverse, the most popular citizen-science platform, which is expected to reach a large number of volunteers, and motivate them to play a part in frontier scientific research.
After a brief description of the demonstrators comprising REINFORCE, we present in detail the demonstrator titled "Search for new particles at CERN", which engages citizen-scientists in searches for new elementary particles produced by the high-energy proton-proton collisions at the LHC of CERN. To make this possible, the demonstrator adopts a three-stage architecture which will be described below.

The 22nd International Workshop on Neutrinos from Accelerators (NuFact2021)
6–11 Sep 2021
Cagliari, Italy

*Speaker

https://pos.sissa.it/

PoS(NuFact2021)047

1. Introduction

The REINFORCE project [1] aims to engage citizens in the research activities of Large Research Infrastructures through citizen-science activities, in order to increase science-awareness and to positively influence attitudes towards it.

The aim is to engage non-expert volunteers in citizen-science projects, ranging across all age groups and in high numbers. No prior computing knowledge is required of the public. The project's design is built around the so called "demonstrators", which attract and motivate volunteer citizens to become a part of, and contribute to, frontier research science topics. The demonstrators are hosted in the Zooniverse platform [2], the most widely used citizen science platform with over 2M users.

REINFORCE has four "discovery" demonstrators:

- Gravitational Wave (GW) noise hunting: Its aim is to get citizens help the scientists in the search for GW by recognizing and categorizing different sources of noise.
- Deep Sea Hunters: Its aim is to get citizens study sources of bioluminescence and acoustic noise detected by hydrophones in deep sea where the KM3NeT neutrino telescopes will be installed.
- Search for New Particles at the Large Hadron Collider of CERN: This is the demonstrator which will be explained in detail below.
- Cosmic Muon Images: Citizen scientists will be engaged in Geoscience and Archaeology studies by observing traces of atmospheric muons.

In addition to the above four demonstrators, an important objective of REINFORCE is to extend the data handling and research to the scientific community which is sense-disabled (especially visually-disabled) and senior citizens.

2. The "New Particle search at CERN" demonstrator

The purpose of this "discovery" demonstrator is to allow citizens to explore the cutting edge research done at CERN. It is designed to use data recorded by the huge general purpose ATLAS experiment [3] at the Large Hadron Collider (LHC) of CERN. ATLAS, together with its "sister" experiment, CMS, announced, in July 2012, the discovery of the Higgs boson [4],[5]. This was a major scientific achievement of thousands of highly specialized researchers who over decades designed, constructed and operated cutting-edge detector systems, along with the dedicated LHC accelerator itself. The discovery of the Higgs boson established the validity of the Standard Model (SM), the theory that describes the known elementary particles and the interaction forces between them. For the first part of the "discovery" stage with real data, the volunteer citizens are invited to study the Higgs boson through its decay two photons.

Furthermore, the data corresponds to the hard scattering of proton constituents (i.e. quarks, gluons), which can produce all of the known particles, even heavy ones, that decay instantaneously, or potentially (due to the unprecedented energies available) to new particles, predicted by theoretical models and which go Beyond the Standard Model (the so-called BSM theories). The production of such particles may be rare, but this can be countered with the expected statistical size of the LHC data, considering that they will continue to be recorded for at least two more decades. In addition to this, the scientific impact from new discoveries may well surpass even the discovery of the Higgs boson. In the second part of the real data "discovery" stage, the volunteer citizens are motivated to contribute to the search for new particles predicted by such BSM theories, as well. Their search could lead to a discovery, which would be a direct proof of new physics and would highlight a path for future research.

In order to facilitate the citizens' analysis of events, we have chosen to give them a virtual representation of the products of the collisions of the LHC beams and ask them to visually inspect and analyse those images according to our instructions.

3. The different stages of the demonstrator

We have divided the citizens' study of events in a three stage process. All stages of the demonstrator have been uploaded on the Zooniverse platform [6]. For each stage a tutorial together with an extensive help session (including a video) is provided in the Zooniverse platform.

In the ATLAS experiment a large variety of algorithms has been developed to automate the task of identifying displaced vertexes or identify different particle traces –tasks which are necessary for the specific searches- with varying degrees of accuracy depending on a large number of factors. As these are tasks that, under certain conditions, can be easily performed by humans, our goal is to compare the accuracy and efficiency of the automated identification methods to that of human volunteers. The primary reason for using simulated events at the first two stages is that this is the only way of knowing with absolute certainty, the exact position of each displaced vertex or the identity of the particle tracks. The same events will be analyzed by both the algorithm and the citizen scientists allowing us to make an accurate and impartial comparison based on the known "true" information.

The data gathered at the first two stages are a large and important part of this project so the citizens should not treat it as a practice session for later stages. Each identification that the citizens perform is an important data point for us that will help in our goal of comparing humans and computers more accurately. Indices measuring one against the other will be provided in order to determine the areas where each method is superior.

3.1 Stage 1: Identification of Displaced Vertices

The proton-proton collision take place at the center of the ATLAS detector, defined as the main interaction point (IP). Theories which predict the existence of long-lived particles or decays of the Higgs boson into two photons –one (or both) of which convert in a pair of electrons-positrons can produce a Displaced Vertex (DV), namely a point with two of more tracks located away from the IP. In stage 1 we use a simulated sample of events to have the user practice in locating DVs. The displayed candidate tracks are obtained from the set of tracks that have been reconstructed in the ATLAS Inner Detector, satisfying the quality criteria used in ATLAS experiment searches for long-lived particles [7] as well as, few additional requirements to facilitate the identification by the human eye. ATLAS Inner Detector is installed around the center of the ATLAS, right after the LHC beam pipe. It consists of three different sub-detector technologies, all designed to accurately record traces of charged particles produced at the collision point and moving outwards. The users inspect stationary images displaying reconstructed tracks of charged particles (the tracks are drawn in different colors so that the

PoS(NuFact2021)047

users can associate them between the two views of the detector) in both views: transverse (left) and longitudinal (right) as shown in Figure 1.

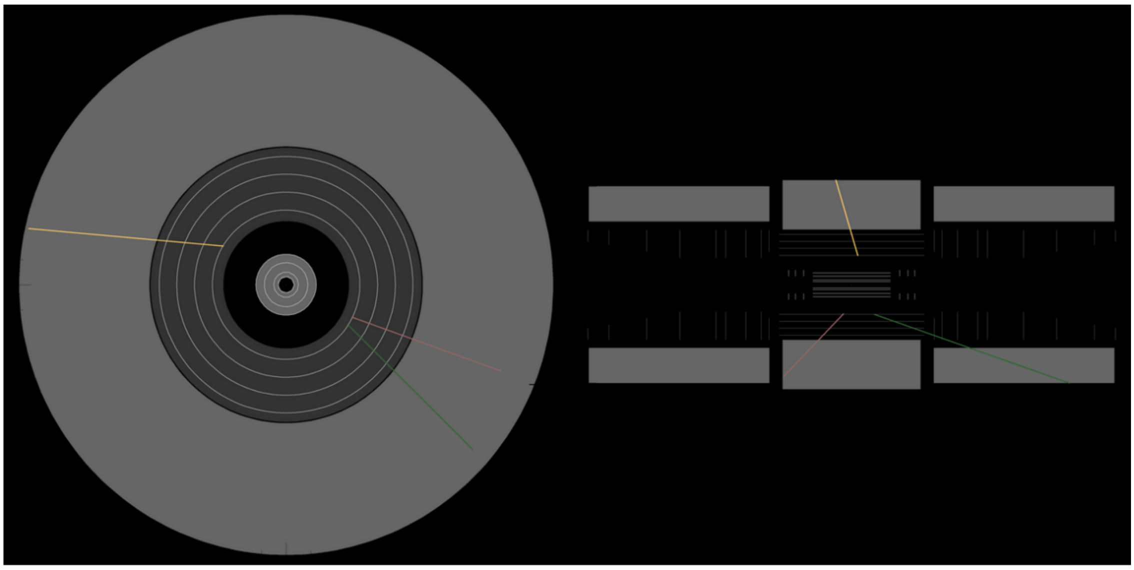

Figure 1: A displaced vertex event displayed in the Zooniverse platform (both projections shown).

The task that citizen-scientists have to carry out, is the identification of the coordinates of any DV in each event, using the mouse pointer. The citizens' selection is internally assessed by Zooniverse platform based on the "truth" information of the location of the DVs. Zooniverse presents an immediate feedback to the citizens after each processed event, before progressing to the next event as shown in Figure 2.

Figure 2: The feedback of a displaced vertex search on the Zooniverse platform for stage 1

After an event is viewed a number of times (each one is called classification) the project manager can specify a limit and “retire” the event.

In order to compare human performance with automated algorithms we have developed a high purity sample of simulated events corresponding to various scenarios of new particles with DVs and an algorithm that extrapolates the tracks and automatically looks for intersection points. Specifically, the algorithm loops on all possible pairs of 3-dimensional tracks and estimates their distance, i.e. the length of the shortest line segment connecting them. If the track distance is of the order of a few millimeters, and if the shortest line segment connecting them lies between the center of the detector and the first measured point (hit on the detector) of either track, then its middle is considered as a possible common origin.

3.2 Stage 2: Identification of particle traces

In order to study the real events, the citizens need to identify the different traces which certain particles (electrons, muons, photons and converted photons) leave in the ATLAS detector. To do so, the citizens have to perform one step further and instead of simply examining images, they should interact with the HYPATIA [8] event display in order to perform in-depth analysis of the event traces. HYPATIA (HYbrid Pupil’s Analysis Tool for Interactions in ATLAS) is an event display for ATLAS events either real or simulated. For stage 2, simulated events are chosen and a dedicated HYPATIA layout has been developed as shown in Figure 3.

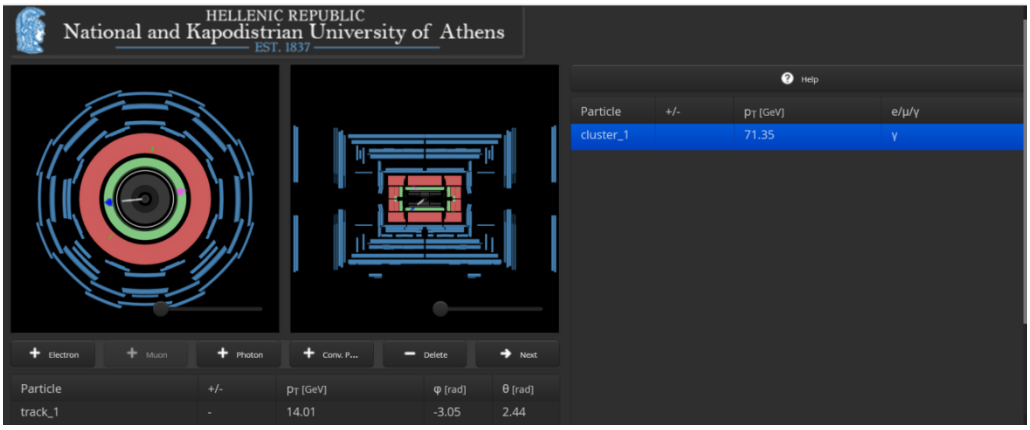

Figure 3: An event displayed by the HYPATIA event display as used in stage 2

As in stage 1, citizen-scientists are provided with a tutorial which shows the typical signatures that electron, muons, photons and converted photons leave in the ATLAS detector as well as an extensive help section, with a description of the ATLAS detector and detailed instructions on how to recognize each particle Their selections are automatically logged by HYPATIA for further analysis and feedback.

As already mentioned, this stage uses simulated events because this allows us to know with certainty what particle created each track or cluster. Each identification answer will be compared to the “truth” information from the simulated events. We have developed a machine-learning algorithm which uses exactly the same information as the one available to the users in order for this comparison to be fair and impartial. The algorithm is based on Boosted Decision

Tree classifiers [9], which provides, for each candidate particle, weights corresponding to the different particle hypotheses.

3.3 Stage 3 with real data

3.3.1 Introduction to stage 3

Both stages 1 and 2 have trained the citizens in the identification of DVs and of particles traces necessary for them to proceed to the third stage which involves further interactions with the event displays and calculation of kinematic qualities in order to make possible discoveries.

The difference between stage 3 and the previous two stages is that here the citizen scientists use samples of **real** events collected by the ATLAS detector at the LHC Run II, which are part of the 13 TeV Open Data released by ATLAS [10]. Since in these stages we need some additional information for the events, we have asked and obtained formal approval from the ATLAS collaboration to use the real data needed for these searches.

Stage 3 is divided into two parts.

- The 3a part called "Study of the Higgs bosons" where the citizen scientists study the decays of Higgs bosons to photons.
- The 3b stage "Discovery of Long Lived Particles (LLP)", where the users need to select events which they believe correspond to one of the long lived particles predicted by theories of "new physics" (Beyond the Standard Model).

In both parts the citizens interact more with the HYPATIA event display, selecting traces of particles and calculating kinematic quantities, such as the mass of Higgs boson or possibly, the sought-after particles. For the needs of each one of the 3a and 3b stage we have written up different dedicated layouts of the HYPATIA tool.

In stage 3 the citizens have the opportunity to visualize real events, interact, analyse them and possibly make discoveries according to certain guidelines given to them.

3.3.2 Stage 3a – Study of the Higgs bosons

In stage 3a, citizen-scientists need to identify, in each event, energy clusters in the ATLAS electromagnetic calorimeter (which measures the energy of electrons and photons). The citizens, following their training in stage 2, they have to identify photons or converted photons that may originate from the decay of a Higgs boson. For each pair of selected photons, HYPATIA calculates the corresponding invariant masses as shown in Figure 4. In addition to the selection of the photons or converted photons of an event, the citizen-scientists are given guidelines how to rate the event using a star rating system ★☆☆☆☆, according to its complexity and its probability to include a rare Higgs production mode.

The citizen-scientists' selections and event ratings are recorded by HYPATIA. At the end of the project, statistical analysis of the invariant mass histograms will follow, in order to look for differences between the three event cases: events with two unconverted photons, events with an unconverted and a converted photon and events with two converted photons. Finally, discussion on cases in which citizens have shown particular interest (through the star rating system) can follow in the project's discussion board.

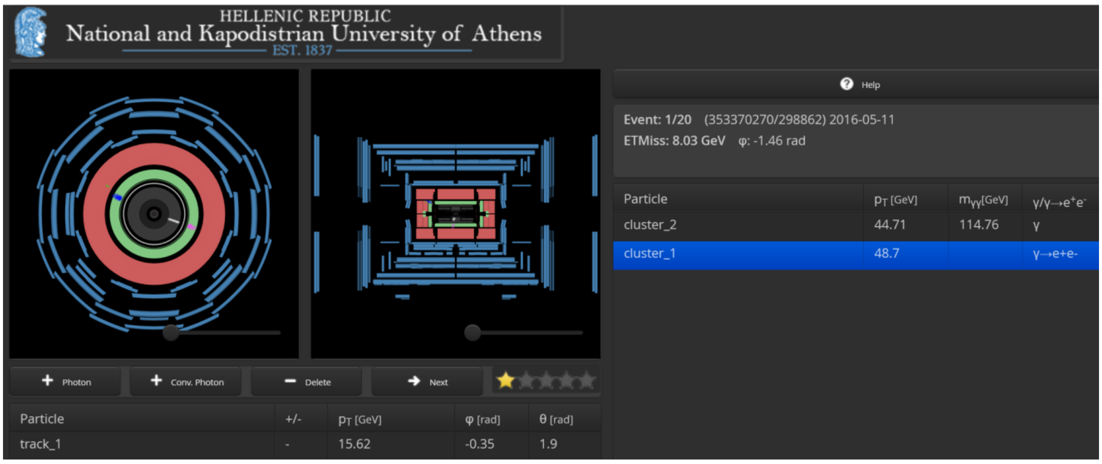

Figure 4: A collision event displayed by HYPATIA in stage 3a. Two projections of the entire ATLAS detector are shown; transverse (left) and longitudinal (right).

3.3.3 Stage 3b – Discovery of Long Lived Particles (LLP)

In stage 3b, citizen scientists will be searching for LLPs through the identification of DVs. Following the vertex identification, they will proceed to select muons originating from the DV and study their kinematic properties, as well as properties of the candidate displaced vertex (e.g. its invariant mass and transverse distance from the beam-collision point), which are provided by HYPATIA, as shown in Figure 5. Furthermore, they are asked to rate each event, using the star-rating system ★☆☆☆☆, according to how consistent they believe the event is with the production and decay of an LLP. This way the citizens will be directly involved in potential discoveries without requiring any high level physics or computing skills.

As in stage 3a, citizen-scientists' selections and event ratings are recorded by HYPATIA. At the end of the project, a statistical analysis of the kinematic quantities populated with the measurements of all of the participating citizens will be performed, in order to identify any evidence of the possible existence of LLPs. Furthermore, discussion on cases in which citizens have shown particular interest (through the star rating system) can follow in the project's discussion board.

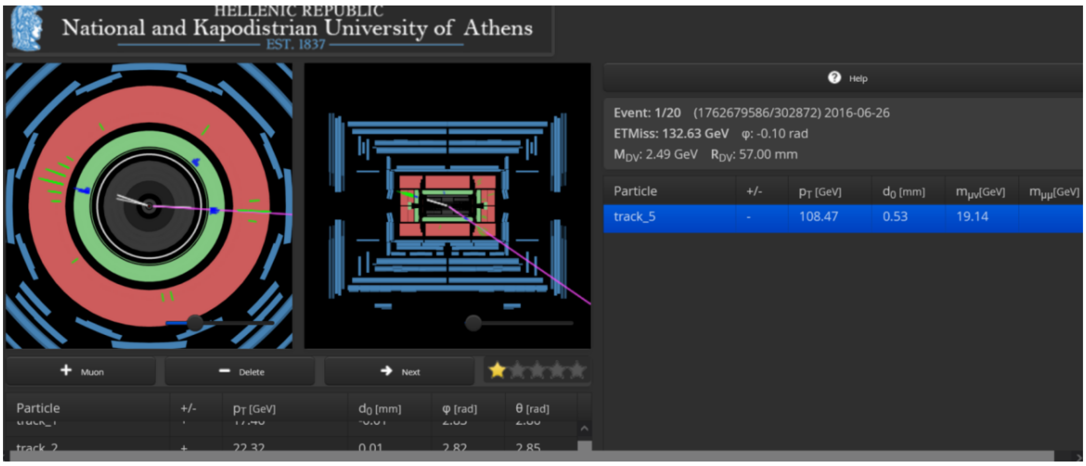

Figure 5: A collision event displayed by HYPATIA in stage 3b. Two projections of the entire ATLAS detector are shown; transverse (left) and longitudinal (right).

4. Conclusions

The “New Particle Search at CERN” was launched on the Zooniverse platform at the end of October 2021. Up to now it has about 50,000 classifications in all three stages together. At the end of the project -around next summer- we will perform an overall assessment of the users’ classifications and a comparison with our custom-made algorithms for stages 1 and 2 (simulated data). For the two parts of stage 3 (real data) we will carefully study all highly rated events to look for possible discoveries.

5. Acknowledgements

This project has received funding from the European Union’s Horizon 2020 project call H2020-SwafS-2018-2020 funded project Grant Agreement no. 872859

References

[1] The web site of the REINFORCE project: https://www.reinforceeu.eu/

[2] The web site of the Zooniverse platform: https://www.zooniverse.org/

[3] The ATLAS experiment web site: https://atlas.ch/

[4] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with ATLAS at the LHC*, Phys Lett B. 716 1 (2012) 1-29

[5] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett B. 716 1 (2012) 30-61

[6] The web site of the “New particle search at CERN” demonstrator described in this contribution: https://www.zooniverse.org/projects/reinforce/new-particle-search-at-cern

[7] ATLAS Collaboration, *Search for long-lived massive particles in events with a displaced vertex and a muon with large impact parameter in pp collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. D 102 (2020) 032006

[8] C.Kourkoumelis and S.Vourakis, *HYPATIA-an online tool for ATLAS event visualization*, Phys.Edu. 49 (2014) 21

[9] The Machine Learning algorithm description: https://root.cern/manual/tmva/

[10] ATLAS Collaboration, *Review of the 13 TeV ATLAS Open Data release*, PUB-OTRC-2020-01, 22nd January 2020