Recursive engagement: the public as data analysts and outreach creators

Charles William Kalderon, on behalf of the ATLAS Collaboration

Lund University, Sweden

E-mail: william.kalderon@cern.ch

Two recent outreach projects are making use of public communities to enhance and build upon the first phases set up by physicists. "ATLAScraft" is a recreation of ATLAS and the wider CERN complex in Minecraft. The basic layout was provided, but school students subsequently researched and created the experiment and subdetector models and their own mini-games to explain various aspects of the LHC and detector physics to others. "HiggsHunters" asked the public to search for displaced vertices in event displays, during which time a pool of trusted members arose in the associated discussion. A second phase also involved schoolchildren, with groups analysing the data this generated, through which they can both learn the principles of scientific research and contribute directly to it.

1. Introduction

Typically, outreach material and projects are created by members of the scientific community, and disseminated and sustained through their continued efforts. This talk discussed two ATLAS [1] outreach projects that involved members of the public beyond their being consumers: in the case of *ATLAScraft* members of the target audience were heavily involved in the main content creation, while in the case of *HiggsHunters* they formed a community to sustain the project and a separate group acted as data analysts.

2. ATLAScraft

2.1 Minecraft and the base map

Minecraft [2] is a creative multiplayer sandbox game, with Lego-like appearance in large customised worlds. Over 100 million copies have been sold, making it the second-highest-selling video game ever, after Tetris. It has free-build and adventure modes, and allows custom maps to be created and distributed for use by other players.

ATLAScraft aimed to release a fun, educational virtual museum of ATLAS to the Minecraft world, targeting the main players of the game: principally schoolchildren aged 8-15, but also their families and other adults.

2.2 Student development

The base map, of the whole wider Geneva area encompassing the LHC ring and CERN sites, was created by the project leaders, and a multiplayer server hosted at Oxford University. The topography came from LIDAR [3] data, with a resolution of 1 meter. Public buildings and roads were imported from SITG [4] and OpenStreetMap [5] while CERN buildings and tunnels were provided by CERN maps, and the ATLAS detector geometry was that obtained by the *ATLASrift* [6] developers. Finally, feature manipulation engine software was used to convert map data to Minecraft code.

Following this, the creation of the ATLAS museum was done by school students from four secondary schools (ages 11 - 16) in the UK. Through this, with support from the Institute for Research in Schools (IRIS) [7], an organisation whose mission is to give school students a taste of real research, the students learned about ATLAS and high energy physics (HEP).
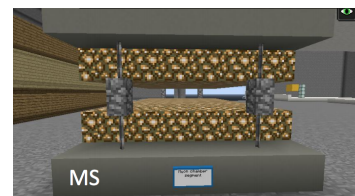
Different detector parts were assigned to each school (some results seen in Fig. 1), with the initial task being to create an interactive Minecraft area which "explains to students your age how that detector section works", following which they were free to populate other parts of the map.
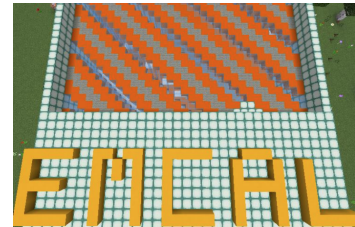
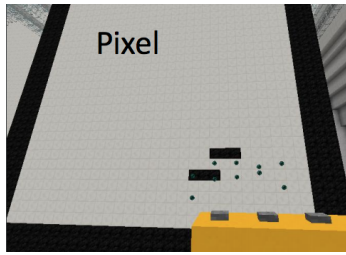(a) Electromagnetic calorimeter
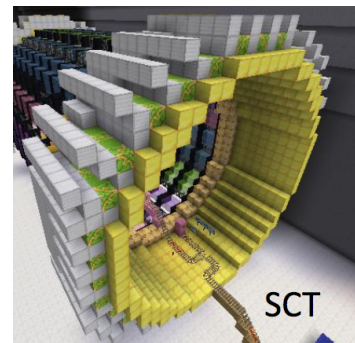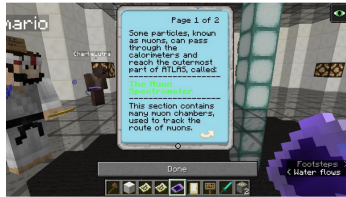
(b) Semi-conductor tracker

(c) A muon chamber

Figure 1: Individual ATLAS subdetectors created by the students

(a) Charge collection in the pixel

(b) Instructions for the muon drift tube game

Figure 2: Examples of games created by the students

The end result is a game whose aim is to visit different parts of the detector and museum and collect a special block for each of the particles of the Standard Model. Players 'spawn' in the CERN globe, which acts as a headquarters for the game and where there is a brief introduction to HEP and *ATLAScraft*. There are reproductions of the ATLAS control room, and general information on ATLAS and the LHC in the surface buildings around it, as well as some offices and canteens. In the ATLAS cavern is a full-size version of the ATLAS detector, in addition to a 5:1 scale version elsewhere for easier exploration, with different subdetectors being made of different in-game materials.

There are then numerous games and simulations, all student-created, describing and explaining the operation of various parts of the LHC and ATLAS. For example, charge capture in pixel detectors is explained via catching small falling blocks on a plate, and muon drift tubes through following hovering spheres (see Fig. 2). There are also links to the *HiggsHunters* (section 3) website.

2.3 Release

The "virtual museum" and base maps were released in late November 2017, with both full and lightweight versions, on the *ATLAScraft* website [9]. This was accompanied by a press release from Microsoft [10] and a trailer on YouTube [8] – some stills are shown in Fig. 3.

The map was also targeted as an educational tool for teachers, and as an outreach tool, as a follow-up after primary school (5-11 years old) visits. Since release, there have been 2000 map downloads.

Possible future plans include the construction of more LHC experiments, since the full LHC ring is present in the map, either following same model with more schools or opening to the general public. There is also the option to make a mobile and/or tablet version, which Minecraft supports, since currently *ATLAScraft* is desktop-only. However, there is currently no active work ongoing on future developments.

Figure 3: Stills from the *ATLAScraft* trailer [8]

PoS(LHCP2018)303

3. HiggsHunters

3.1 Zooniverse and the initial project

Zooniverse [11] is a large platform for citizen science, hosting projects in many areas of science and the humanities where visual inspection and classification by large numbers of people can benefit researchers' understanding of their data. Many of the science projects are in the area of astrophysics and astronomy, following from the initial project, Galaxy Zoo [12], in 2007.

The aim of *HiggsHunters* was to create a particle physics project with the Zooniverse platform. Citizen scientists were provided with specially-processed ATLAS event displays (example in Fig. 4), and were tasked with identifying secondary vertices within them.

The physics aims of this task were two-fold: first to search for new scalars in Higgs decays ($h \rightarrow aa, a \rightarrow bb/\tau\tau$) as described in ref. [13], termed 'baby higgs

Figure 4: Example *HiggsHunters* image, emphasising the inner detector for identification of displaced vertices.

bosons'; and second to directly examine images of data with a large number of human eyes, with citizen scientists asked to mark "weird" features in images. The images consisted of events with a dimuon pair consistent with a Z boson with moderate transverse momentum, i.e. characteristic of the "Higgsstrahlung" *hZ* production process, both from data collected in 2012 and from simulation of the *hZ* process and the $h \rightarrow aa$ model with a variety of *a* lifetimes and masses.
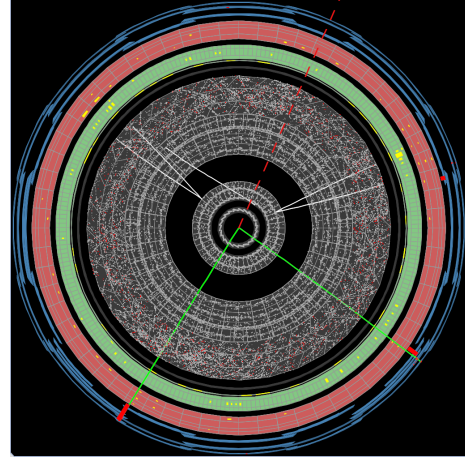

(a) Number of images examined
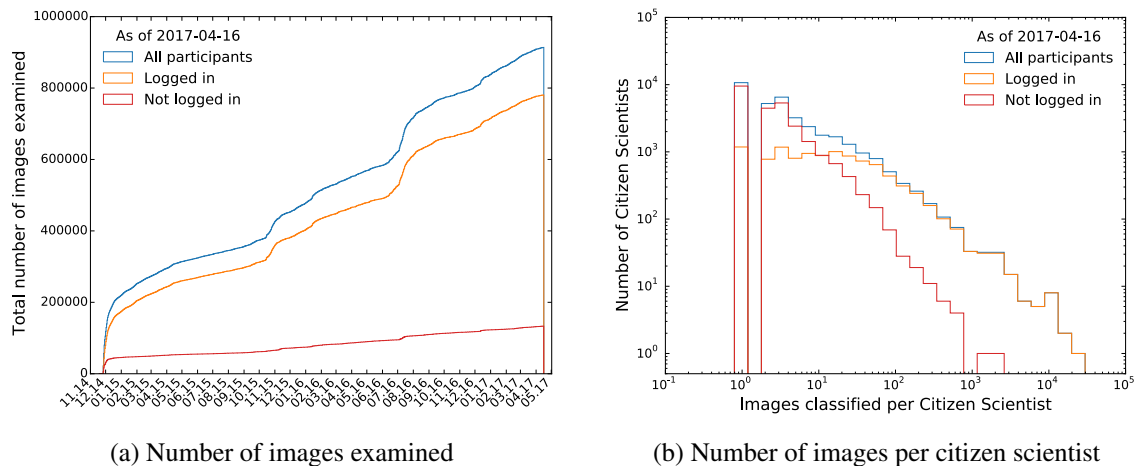
(b) Number of images per citizen scientist

Figure 5: Image classification and participant numbers over time. From [14].

In the 2.5 years following launch in November 2014, nearly 1M images were examined by more than 30,000 people – there are 85,000 images in total, each kept in circulation until there are between 20 and 40 views per image. Most people viewed only a few images, but over 1000 classified more than 100, and the most dedicated person classified over 25,000 images (Fig. 5).

3

3.2 Initial analysis of clicks

In a first analysis of the clicks data [15], clicks from multiple people within a small area of an image were clustered together and the resulting points compared to known secondary vertex locations in simulation. This clustering was subject to several arbitrary parameter choices, e.g. selecting only clicks where the clicker had recorded more than a certain number of tracks, requirements on the fraction of people examining an image whose clicks contributed to a given cluster, the number of clicks in cluster and the cluster radius parameter.
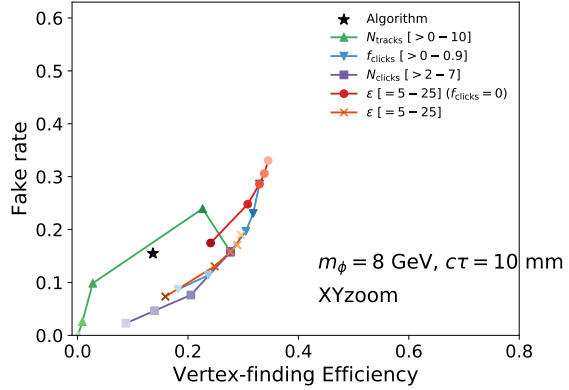
Preliminary examination of performance (Fig. 6) revealed that the combination of citizen scientists' clicks was competitive or better than existing vertex-finding algorithms (as of 2015, which required 3 tracks) in a substantial fraction of models.

3.3 Talk forum

An important part of Zooniverse projects is the associated 'Talk forum' [16]: for *HiggsHunters* there were 20,000 comments from 1345 citizen scientists, and detailed help pages emerged from enthusiastic participants.

Early on in the project, there were several groups of clicks marking 'weird' objects, which were followed up in the Talk forum, and identified apparent jets of muons (Fig. 7). This unexpected feature turned out to be the result of punch-through jets [17] (where some hadronic particles reach the muon system and leave multiple hits there), and the community quickly learned what these were and tagged them appropriately from then on.

This growth in experience is shown in Fig. 8 – use of the "weird" tag plateaus as experience grows, and the term "bun-

(a) Varying clustering parameters for one signal model and image view.

(b) Summary of all signal models and views, comparing the best clustering for citizen scientists (marked with an eye) and the computer algorithm (marked with a computer).

Figure 6: Vertex-finding efficiency and misidentification ('fake') rate of citizen scientists' clicks, compared to the vertex-finding algorithm in use at the time. From [15].

dle" is overtaken by the technical term "jet". Other technical terms adopted include "track", "punchthrough" and "higgszcandidates", while in other cases the citizen scientists developed their own vocabulary: "#diametric" for a dijet event and the aforementioned "#bundle" for "jet". This aspect, as well as the results of a survey of the citizen scientists, is discussed further in ref. [14].

3.4 Public data analysis

In another partnership with IRIS [7], UK schoolchildren were given the clicks data to analyse, in order e.g. to optimise the notion of what a cluster is by considering the accuracy of individual citizen scientists. In this way, both aspects of the search - the 'image processing' and the 'data analysis' - can be citizen-led. This aspect launched in September 2017, and the first results of this effort were presented in July, with some students calculating e.g. the relative accuracy in different image projections and others developing data processing methods and customising clustering algorithms.
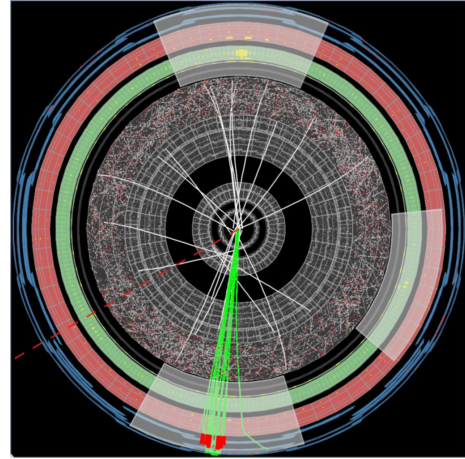
Figure 7: Example *HiggsHunters* image of an apparent muon jet.

4. Summary

Two ATLAS outreach projects expanded their scope beyond the "finished product" to engage more and different groups. In *AT-LAScraft*, there was substantial creation of material by the target audience, with targeting of girls-only schools ensuring a balance of student developers was achieved despite 70% of Minecraft players being male. In *HiggsHunters* continuing support for new participants was provided by experienced participants on a web forum, and analysis of data is being performed by science students in schools.

Figure 8: Cumulative number of comments matching particular words as a function of date. From [14].

Both of these projects were well received by participants in both of their phases: all students involved in the *ATLAScraft* development said they were more interested in physics, and a large majority of the *HiggsHunters* citizen scientists who responded to the survey stated they had read, watched or studied more about science as a result of their involvement.
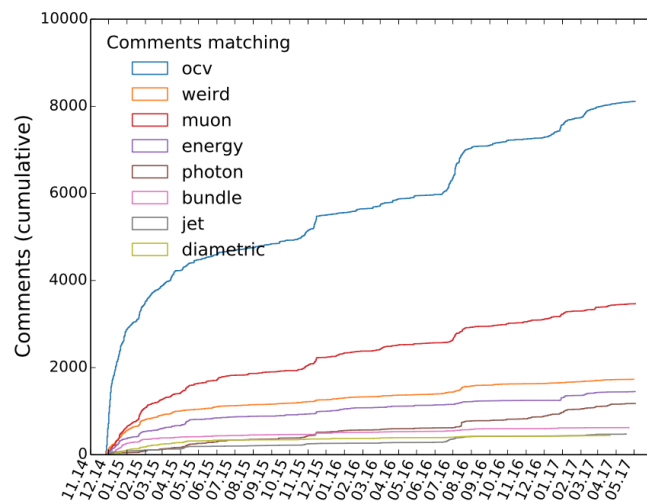
References

[1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[2] Microsoft, *What is Minecraft?*, https://minecraft.net/en-us/what-is-minecraft/.

[3] *Light detection and ranging (lidar)*, https://www.ngs.noaa.gov/RESEARCH/RSD/main/lidar/lidar.shtml.

[4] *Le système d'information du territoire à Genève (SITG)*, https://ge.ch/sitg/.

[5] *OpenStreetMap*, https://www.openstreetmap.org/.

[6] *ATLASrift website*, https://atlasrift.web.cern.ch/.

[7] *The Institute for Research in Schools*, http://researchinschools.org/.

[8] *ATLAScraft trailer*, https://www.youtube.com/watch?v=7RuAej2TJJg.

[9] *ATLAScraft website*, https://atlascraft.web.cern.ch/.

[10] Microsoft News Centre UK, *UK pupils recreate CERN lab and Large Hadron Collider in Minecraft*, https://news.microsoft.com/en-gb/2017/12/13/uk-pupils-recreate-cern-lab-and-large-hadron-collider-in-minecraft/.

[11] C. Lintott, L. Forston, A. Smith et al., *Zooniverse*, https://www.zooniverse.org/about.

[12] C. J. Lintott et al., *Galaxy Zoo : Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey*, *Mon. Not. Roy. Astron. Soc.* **389** (2008) 1179 [`0804.4483`].

[13] M. J. Strassler and K. M. Zurek, *Discovering the Higgs through highly-displaced vertices*, *Phys. Lett.* **B661** (2008) 263 [`hep-ph/0605193`].

[14] A. J. Barr, A. Haas and C. W. Kalderon, *Citizen scientist community engagement with the HiggsHunters project at the Large Hadron Collider*, *Research for All* **2** (2018) 359.

[15] A. J. Barr, C. W. Kalderon and A. C. Haas, *'That looks weird' - evaluating citizen scientists' ability to detect unusual features in ATLAS images of LHC collisions*, `1610.02214`.

[16] *HiggsHunters Talk forum*, https://talk.higgshunters.org/.

[17] *HiggsHunters Talk page: 'Muon jet followup'*, https://talk.higgshunters.org/#/boards/BHH0000007/discussions/DHH00001o1.

PoS(LHCP2018)303