ARCHIVED ETHERPADS FROM THE 2015 #MOZSPRINT AT CERN

Index:

1. Ideas

(Original archive: https://public.etherpad-mozilla.org/p/MozScience15-ideas)

CERN-hosted sprint:

Location: CERN Ideasquare
Time: 4./5. June 2015, 9.00-18.00
Participation:

- in person: add you name as a participant to the project you want to contribute to, if you don't have a CERN access card, contact us in advance to organise your access card
- remotely: please contact the project lead of the project you want to contribute to for details

Snacks and coffee will be provided, we might ask you for a small contribution to the "kitty"

Please sign up at https://indico.cern.ch/event/397376/, too!!

Agenda (draft):

Thursday, 4th:
10.00 - 10.30 Welcome, introduction, set-up of working groups and spaces
10.30 - 12.30 Hacking
12.30 - 13.30 Lunch (flexible of course)
13.30 - 17.30 More hacking
17.30 - 18.00 Short report to the rest
Drinks :)

Friday, 5th:
10.00 - 10.30 Welcome, introduction, set-up of working groups and spaces
10.30 - 12.30 Hacking

10.00 - 12.30 Parallel session on Open Science tools in the presentation square
12.30 -13.30 Lunch (flexible of course)
13.30 - 17.00 More hacking
17.00 - 18.00 Presentation of sprint results
Final party :)

For any other questions, please contact:
Achintya Rao/ @RaoOfPhysics
Patricia Herterich/ @pherterich
Laura Rueda/ @espacial

Confirmed projects:

- Crowd-sourced Cosmic Rays
    - Project lead: Achintya Rao
    - Participants:
        - Achintya Rao
        - Hugo Day
        - Anirudha Bose (remotely)
        - Alexander Brown
        - James Devine
        - Sophie Redford
        - Sahal Yacoob

- GitHub Science Badges
    - Project lead: Laura Rueda
    - Participants:
        - Laura Rueda
        - Patricia Herterich
        - Jacopo Notarstefano
        - Karel Ha
        - Mateusz Susik
        - Jan Stypka
        - Ioanna Rompou
        - Nalin Chhibber (remotely)

- Enhancing existing/new LHC event displays:
    - Project lead: Tom McCauley
    - Participants:
        - Tom McCauley
        - Adam Wegrzynek
        - Harri Hirvonsalo
        - Jiri Manasek
        - Nisia Thornton (remotely)
        - Eamonn Maguire (Thursday)
        - Carita Logren

        - Luke Barnard
        - Rachel Prudden (Met Office)
        - Michael Saunby (Met Office)

- E3 - Extreme Energy Events real-time monitoring
    - Project lead: Ines Knaepper
    - Participants:
        - Ines Knaepper
        - Daniel Dobos (remotely)
        - Brice Copy
        - Leonardo Milano (only Friday afternoon)
        - Audrius Mecionis

- Testing of the Geotag-X prototype
    - Project lead: Ellie Cervigni
    - Participants (note that you can also just contribute for a few hours):
        - Ellie Cervigni
        - Brice Copy
        - Nairit Sur
        - Eva Tolosa

Other ideas

These might get picked up spontaneously:

- Etherpad-Indico integration: https://github.com/ether/etherpad-lite
    - Just like webcasts are integrated with Indico at CERN – same unique identifier etc. – each meeting could have associated Etherpad automatically created, for joint note-taking.
    - CERN SSO could protect sensitive notes
    - Also useful for every other organisation that uses Indico!
    - We can also document the stand-alone Etherpad set-up to simplify life for others, or make some other non-CERN-specific contributions…
- http://hypothes.is integration with CERN services —- particularly for PDFs on CDS and Indico?
- Education session on existing open-source collaborative tools, including ones (being) deployed at CERN
    - https://coauthoring.cern.ch, which uses ShareLaTeX
    - ?
- Learning from others
    - https://thewinnower.com/posts/archiving-and-aggregating-alternative-scholarly-content-dois-for-blogs – relevant to ZENODO?
- Contributing to standardisation for scholarly communication?
    - https://github.com/scholmd/scholmd/wiki – use-cases we've experienced that are missed?
    - Other (similar) projects? Missing metadata fields?

- CERN Open Data Portal:
    - Integrating analysis examples from CMS that can be performed with currently released datasets? Being discussed internally already. Also see: https://github.com/cernopendata/opendata.cern.ch/issues/787
    - Any other matters?

2. Open Cosmics

(Original archive: https://public.etherpad-mozilla.org/p/mozscience15)

Vidyo room:
https://vidyoportal.cern.ch/flex.html?roomdirect.html&key=Y71x6cflijd0ZKrxesei6NpQ9Q
GitHub organisation: https://github.com/opencosmics
CERN e-group for discussions:
https://e-groups.cern.ch/e-groups/Egroup.do?egroupId=10172161

Motivation
A common data format and storage mechanism to allow cosmic rays detected from multiple experiments to be analysed collectively. In turn this would facilitate the development of cosmic ray detection applications and software. See Diagram 1 in the lnk below:
https://docs.google.com/presentation/d/1WyhnIJdfYHwRHbvGNl9OQCkQgDxIg_s8M_JG0VM__OE/edit?usp=sharing

Background
There are lots of different student/citizen science
cosmic ray experiments either in progress or being planned. At the moment each is developing it's own data structures and storage systems. Investigation of existing cosmic ray experiments will form part of the activity.

Use-cases/Applicability
Such a system could be applied to Cosmic Pi, Crayfis, HiSPARC, ERGO, assuming everyone wants to join in!
1) A cosmic ray detector encounters a cosmic ray particle from a shower and uploads it to a common database.
2) The database can then be queried for all events within a certain physical region, within a certain time frame, providing raw data for the reconstruction of the shower.
See diagram 2:
https://docs.google.com/presentation/d/1WyhnIJdfYHwRHbvGNl9OQCkQgDxIg_s8M_JG0VM__OE/edit?usp=sharing

Feasibility
Part of the exercise is to ascertain the feasibility, already validated in principle by discussions between Cosmic Pi and Crayfis (see notes at the bottom). Cosmic Pi data format is also provided below.

Some of the data will need to be filtered (at source, or via an intermediate level, for example Crayfis), this is experiment specific.

Sprint deliverables
Investigate existing cosmic ray projects (HiSPARC, Crayfis, CosmicPi, ERGO etc.)
Review feasibility for a common platform across some or all projects
Investigate potential solutions for a Cosmic Ray open repository (e.g. CERN Open Data)

Time required
2 days
1/2 Day for investigation of platforms
1/2 Day for Feasability based on platform information
1/2 Day for drafting of common format specification
1/2 Day for example implementation

Skills required
Python, understanding of data structures and database principles.
See implementation in Diagram 3:
https://docs.google.com/presentation/d/1WyhnIJdfYHwRHbvGNl9OQCkQgDxIg_s8M_JG0VM__OE/edit?usp=sharing

Post-sprint follow-up(s)
Integration with Cosmic Pi platform application

Data examples:
HiSPARC:
- http://www.hisparc.nl/en/hisparc-data/
- http://data.hisparc.nl/show/stations_by_country/
- Download data http://docs.hisparc.nl/publicdb/data_access.html#downloading-via-python

Global data format outline:
(per event)
High level
1) Position (GPS coordinates) + bound (accuracy)
2) Timing (Converted to UTMS) + bound (accuracy)
3) Energy (MeV) + bound (sensitivity)
4) Direction vector + range
5) Detector name/reference -> needs to be unique.
Low level
6) Detector specific metadata

Other projects:
- https://quarknet.i2u2.org/group/international-muon-week
- http://ippog.web.cern.ch/resources/2014/international-cosmic-day-2014

Misc. notes:

Do we add the project to http://www.mozillascience.org/collaborate/ ?

- Standard kit – schematic?
- DAQ / analysis tools?

Cosmic Ray projects:
HiSPARC http://hisparc.nl/en http://data.hisparc.nl/show/stations_on_map/
ERGO http://www.ergotelescope.org/
Crayfis http://crayfis.io/
CosmicPi http://cosmicpi.org
RICH http://www.insidescience.org/blog/2015/04/13/particle-physics-cheap
CrowdMag: http://www.ngdc.noaa.gov/geomag/crowdmag.shtml
DECO http://wipac.wisc.edu/deco
...
Deliverables:

- Data specifications: format (JSON?), metadata etc.
- Set up a test server that can collect data
- Crayfis has a protocol-buffer format, maybe it can be used by CosmicPi
- Overview of all existing projects with commonalities and differences (outreach task)
-
-

Notes from meeting on 2015-04-24T17:00+02

- Cookies!
- What are the goals of the sprint?
- Standardisation for data: format, storage, analysis?
- Coincidence cannot be done by all detectors, i.e. those that have a single scintillator block or CCD/CMOS
- Kyle: Can you drop details of how you collect data from the phones? And point to more documentation…
- "Where to store the data is not the stumbling block": Amazon?
- Data privacy issues: how to address?
- ID-tagging: can you identify a person from the data you take?
- Getting the CosmicPi kit FCC/CE certified
- The extra components would need to be certified as safe to use in an educational setting.
- Certification costs need to be determined
- Goal of CosmicPi: get hardware+API

- For Crayfis, it's important to also know when the system was measuring, but no particles were seen. This is for scientific analysis later.
- Timing: how accurate?
-
-
- Timing: NTP vs Navstar (¿por que no los dos?)
- Energy ranges for CosmicPi (low energy) and Crayfis (high energy)?
- Merch! A phone-case or sticker that blocks your camera
- What about using CCDs from old cameras? Is this possible?

Cosmic Pi Data Raw Format
Things that cosmic pi records as part of an event:
0) Cosmic Pi serial number
1) GPS position, lat/long/alt
2) GPS time stamp
3) 2 channel energy reading (2x12 bit) from SiPMs via ADC's
4) Temperature (possibly 2 channels)
5) Orientation (3-axis accelerometer)
6) Pressure (12 bit), possible direct altitude conversion
7) Relative humidity % 6 bit precision.
8) Magnetic field (3 axis magnetometer)
9) How long have I been switched on for?
10)The cryptographic signing key, based on hardware token.

Cosmic pi goals
Short term: Working hardware
Medium term: Kickstarting, Draw pretty pictures like this
http://www.aspera-eu.org/images/stories/Media/MEDIAPICTURES/HR/cosmicrays2.jpg
Long term: Distributed data processing and storage on Pi.

Crayfis goals:
-

3. GitHub Science Badges

(Original archive: https://old.etherpad-mozilla.org/pojjcLQEVD)

Mozilla Science Global Sprint, CERN Site

Project idea: 'GitHub science badges'
---
aka Matraz

GitHub repo: https://github.com/MSusik/matraz

Create a visualisation for scientific GitHub repositories that shows how re-usable a repo is

We went for 4 criteria (to start with, the brainstorming below shows that there are more things that could be included in the future but were not feasible to implement in the scope of this sprint):

- the license chosen for the repo
- is there any documentation?
- is there a contact person for questions
- does the repo have a DOI assigned.

Visualisation: http://i.imgur.com/SMa3BV4.gif

Full flask = re-usable scientific code as our 4 criteria are fulfilled

Brainstorming we started from:

a free-style modification of Jacopo's idea :)

Motivation

Researchers are increasingly publishing, sharing and reusing software. Sometimes it is difficult to identify how open, extensive or used a project is. Providing a visual representation (badge-style) could prove helpful to describe simply and effectively such details. Moreover, it could encourage citation and reuse by highlighting the persistent identification of the software.

Description, applicability and use-cases

GitHub is one of the most widely used repository and version control services. Although other alternatives exist (e.g. Bitbucket), for the scope of this sprint, we will only consider GitHub.

GitHub provides an open API to retrieve the information of each repository. We will use the retrieved information to extract the most important details and build the badge.

The badges could look like:

```
-------------------------------------
| License                         |  << color coded?
-------------------------------------
| Some metrics                    | << e.g. # lines / # contributors?
-------------------------------------
| DOI of the latest issue         | << simplify citation
-------------------------------------
```

Documentation
Contact info

Possible metrics for the re-usabilty of code

- license (see http://choosealicense.com/licenses/ for options and develop a rating for their openness)
- does it compile/ pass tests (e.g. is there a Travis integration...)
- Is a DOI minted to facilitate citability?
- How much metadata is available (this needs some discussion about the metadata we count on how we will measure the "much": length of the readme file?)
- # of open issues (or better resolved issues as this might be a better measure for good documentation)
- does the project have milestones? (sustainable?)

For impact, a lot of work is done by http://sciencetoolbox.org/ already
-> can we integrate their work? do we need to expand? how?

Technology/Skills required

Python... ?
Wrappers https://developer.github.com/libraries/

Ideas from: https://www.openhub.net

http://www.mozillascience.org/contributorship-badges-a-new-project/
https://github.com/mozillascience/PaperBadger
http://paperbadger.herokuapp.com/

Time required and feasibility

2 days, starting with defining the concrete specs and then putting it into practice

Deliverables

Docker based web app, the entire system can rely on a docker image to calculate the code 'score' on the fly and store it to some kind of knowledge base for reference [concept similar to CodersCrowd]

-

Post-sprint follow-up(s)

-possible integration/refinement as part of the THOR project (https://thorproject2014.wordpress.com/)

4. Event Displays

(Original archive: https://public.etherpad-mozilla.org/p/MozScience15-eventdisplays)

Day 2 (5.6.2015)

Everyone at CERN seems to have converged on a Google Cardboard application with Luke and Carita doing most of the heavy lifting yesterday and Luke doing most of it today.

The result is on http://cern.ch/cms-cardboard (alpha)

How-to and behavior:

- Go to the url on your phone in Chrome (recommended)
- Open an event via the file button
- Rotate your phone to landscape mode
- Press the binocular button
- Insert phone into your cardboard and enjoy
- If you lean back it zooms in; if you lean forward it zooms out
- If you are adventurous press the "movie" button before you strap on the cardboard: you should see some proton bunches come and collide and tracks appear and disappear

There was also a pull request from Michael at the Met Office allowing for setting the framerate.

Day 1 (4.6.2015)

Tasks:

Recording an animation

- Implement a way for an user to define start and end state using the UI of iSpy (user selects what is displayed at start state and at the end state).
- Store these into a JSON-file that defines the .ig-file and the event inside .ig-file that these object belong to.
- The JSON-file should also specify transitions (camera angles, panning, etc.) between the start and end state.
- --> Make iSpy interpret contents of the JSON file and render the animation.
- There could be additional states between the start and the end file. (each one having the transitions between states defined)

- Check https://github.com/cms-outreach/ispy-webgl/blob/master/js/animate.js for references on how Animation-button is implemented. These "instructions" for iSpy should be defined in the JSON-file.

CMS-visualization for Google Cardboard **(Luke)**

- https://www.google.com/get/cardboard/
- Here's a demo how it should look like: https://github.com/tpmccauley/cms-cardboard
- WIP version http://cms-cardboard.web.cern.ch/cms-cardboard/

Import models to iSpy in different formats (not just .obj)

- Currently it is possible to import models into iSpy in .obj-format. Three.js support importing of models in different formats.
- http://threejs.org/docs/ (scroll to "Loaders" for more documentation / references)
- http://threejs.org/examples/#webgl_loader_obj (search "loaders" for more examples)
- Collada import might be the most useful?

Animation of tracks, "explosions", etc.

- ???

Export SVG-image from iSpy

- Check http://blog.felixbreuer.net/2014/08/05/using-threejs-to-create-vector-graphics-from-3d-visualizations.html
- svg-crowbar has been used succefully: http://nytimes.github.io/svg-crowbar/, https://github.com/NYTimes/svg-crowbar

Documentation / code commenting

- Document and comment for developer friendliness

Loosely-coupled object configurations [low_priority?]

- Currently file https://github.com/cms-outreach/ispy-webgl/blob/master/js/objects-config.js defines how an object should look like.
- The file also holds the keys of the objects that iSpy expects to be in an .ig-file ("CSCSegments_V2", "CSCSegments_V3"). If a new object if added to the .ig-file the must be defined here also, or else iSpy doesn't know how to draw it.
- Differences on how to draw a "CSCSegments_V3" vs "CSCSegments_V2" are negligible :) and basically this kind of version change that doesn't involve changes to code that does the actual drawing (https://github.com/cms-outreach/ispy-webgl/blob/master/js/objects-draw.js), shouldn't need code changes at all.
- Goal is to specify how objects should look like in the .ig-file or a JSON-file with the same name as the .ig-file. (Any alternatives?)
- --> What about version changes that DO involve changes to https://github.com/cms-outreach/ispy-webgl/blob/master/js/objects-draw.js ? Are there going to be any?

RZ and RPhi views **(Eamonn)**

- Reproduce the functionality of standalone iSpy, where you can quicky get RZ and RPhi-views of the event. (Need screenshots of the views as examples...)

Different views for Experts and Public (Show advanced functions-button / -checkbox)

- Currently iSpy is more or less for Expert users (physicists?)
- For public-users (wider audience) abbreviations like "CSC" are not informative and in order for iSpy to be more attractive to general public help / info menus and windows should be implemented.
- Canvas version of iSpy http://opendata.cern.ch/visualise/events/CMS has the required functionality (check the red question marks next to elements of treeview).
- --> Implement this (or similar) functionality to WebGL-version of iSpy
- We don't want the UI of iSpy to be cluttered or provide "not needed information" to expert users, so this user-friendliness should be disabled for expert users. Maybe a checkbox in settings-panel? ("Show advanced-view", or similar?

Motivation

Get involved developing an application used to create images and animations of LHC collision events as seen in the CMS detector.

Typically event displays are bespoke applications created by the experiments and used by the experts. These days, with modern browsers and open-source code and open data anyone can help with development and visualize some real collision events.

Background

The browser-based event display for the CMS experiment at the LHC (http://ispy-webgl.web.cern.ch/ispy-webgl/) is open-source, open-access, and uses a simple JSON based format for its input. In November last year CERN released LHC data via an open data portal (http://opendata.cern.ch). Some of the data was released in event display format so there is some open data to play with.

Use-cases/Applicability

The display is used by CMS in physics masterclasses aimed at high-school students and to create images and animations of collisions for public dissemination of physics events e.g. http://cms.web.cern.ch/news/lhc-delivers-low-energy-collisions-cms-and-other-experiments The display could be adapted for other uses such as public exhibitions, mobile devices, etc.

Feasibility

Feasible!

Sprint deliverables

Some ideas:
- Create native apps for mobile devices
- Improve geometry handling: currently we suck in some large-ish files that describe the geometry; is there a better way? Almost certainly.
- RZ and RPhi views?
- Picking? Correlated picking (i.e. pick an object in the canvas and show it's info in the table)? DONE
- Improve animation config? Use key frames?
- Improve object configuration?
- Another idea: expand/improve application for Google Cardboard
https://github.com/tpmccauley/cms-cardboard

Time required

Skills required

A bit of knowledge about experimental particle physics and the CMS detector would be useful but not required.
Otherwise, knowledge of JavaScript, HTML, CSS, and WebGL (three.js) is required.

Post-sprint follow-up(s)

Keep involved via github?

Resources:
https://github.com/cms-outreach
https://github.com/cms-outreach/ispy-webgl
https://github.com/cms-outreach/ispy-online
https://github.com/cms-outreach/ispy

Misc. notes:

- Evtl. kick-start the Total Event Visualizer w/ MediaLab and some people from the experiments (Tom, Ana) as a tool for Open Education/Masterclasses? – Patricia
- Or work on enhancements to already available and open-source iSpy (for CMS only, though)? I've proceeded to add text on this above.

5. Extreme Energy Events

(Original archive: https://public.etherpad-mozilla.org/p/MozScience15-extremeenergyevents)

Motivation
Explosions and other extreme energy events: so the world knows now

Use of explosives especially in urban areas is the defining phenomenon of the last and current century. It and other extreme energy events, is the stuff of everyday global news. But still today and in terms of the nature and extent of the destruction, such events remain in the domain of eye witness accounts or official statements. E3 would permit with multi technology detector systems a real time monitoring on an entirely objective basis of explosions and other extreme energy events in a given context. By collecting scientific hard evidence for the watching world, E3 aims to reduce the excesses of violence by making the story of an attack publicly available as it happens, improve the information situation for peace-keepers, humanitarian agencies, journalists as well provide new possibilities with respect to prosecution of war crimes.

Background
OUR OBJECTIVE is to provide the means to monitor explosions and other extreme energy events on a real time, objective and on-line basis.

BY designing and building appropriate technology in the form of detectors, and coupling these detectors with the necessary software for analysis, telecommunications and the end-product on-line.

WITH THE GOAL OF making available to any interested parties the capacity to:

- detect the nature and extent of active hostilities and other extreme energy events;
- monitoring of cease fires;
- facilitate needs assessment and security requirements for humanitarian organisations with respect to extreme energy events;
- provide accurate reporting of extreme energy events on an objective basis;
- provide people whether affected populations, schools, companies etc. with - objective information pertinent to their own security and well-being;
- provide the basis for evidence collection for crimes including war crimes;
- carry out research with respect to extreme energy events.

Use-cases/Applicability
How would real-time , objective and open source information about extreme energy events (especially explosive events) bring about change? Who would the end-users be of E3?

- Peace-keeping / enforcing bodies could reliably monitor conflicts and cease-fire agreements.
- Humanitarian / relief organisations would be able to use the information to assess damage, needs of civilians and their own operational security.

- The media would have a more accurate and non-biased picture of events relying less on “reports.”
- The general population would know what is happening and where and so be able to make their own security decisions if, for example, they are moving from one town to another.
- Airlines would be able to verify areas of active conflict and so change flight paths accordingly.
- Lawyers would be able to use the information as “evidence“ in the pursuit of prosecuting war crimes.
- Academic and organisations including those orientated around conflict and peace-building could use the information for a wide variety of research projects.

Feasibility

Sprint deliverables

- a first version prototype of the web user interface demonstrating the user experience and data visualisation of the E3E data (sample data?)

Time required

- 48 Hours

Skills required

- design
- web development
- data visualisation
- programming
- data analysis

Post-sprint follow-up(s)

- CERN Summer Student
- CERN Webfest
- THE Port hackathon at CERN

Misc. notes:

6. GeoTag-X

(Original archive: https://public.etherpad-mozilla.org/p/MozScience15-geotagx)

What is the GeoTag-X project?

GeoTag-X relies on volunteers analysing photographs taken in disaster-affected areas. Thanks to their work, we should be able to gather data that will be crucial to humanitarian response efforts.

What are we trying to do for the Mozillla Global Sprint 2015?

Test the GeoTag-X platform!
For this we need as many volunteers as possible to get online analysing photos and giving us their feedback.

For the sprint we aim to have all 81 photos of the Winter Shelter application http://geotagx.org/project/wintershelter/ analysed by 20 pairs of eyes. This doesn't mean you have to go through all these photos: what we need from you is to analyse as many photos as you feel like – even just 2 or 3 would be extremely helpful.

Who can contribute?

Anyone can be a volunteer – there are no particular skills required to participate in our project.
All you need to start contributing is a pc or tablet with an Internet connection.
This means that you can do it wherever you want, whenever you want.

We will be at the CERN Ideasquare if you want to come and work with us during the Sprint. Just make sure to register!
http://indico.cern.ch/event/397376/

How can you contribute?

All you have to do is create an account and answer some questions about some photos. Don't worry: it will take no longer than 2 minutes to register and 1 to 3 minutes to analyse each photo.

Are you ready? Follow these steps and start contributing.

1. Go on the GeoTag-X project website http://geotagx.org/

2. Create your account http://geotagx.org/account/signin

Note: Is that mandatory? No, but it would be very helpful for us. We would like to create a community of experienced volunteers.

Now you are ready to contribute!

3. Go to "Projects" and pick up one.

4. Hit the play button to get started!

We highly recommend you complete the tutorial before starting the project, as it will teach you what to look for.

5. Send us your comments and feedback!
Via the feedback link on the website, or on github
https://github.com/geotagx/geotagx-theme/issues

Some things to keep in mind
If you are not sure about your answers, don't worry: each photo is presented to at least 30 people and we take the majority answer.

If you are puzzled, take a look at the help.

You can see your contributing progression in the "Community" page.

What are you waiting for? Go to the GeoTag-X project webpage and become one of our volunteers!

Want to do more?

There are other ways to participate to our project:

- All of our code is on github https://github.com/geotagx and everything is open source. Help us by contributing to the project and submitting any issues you find, or requests/ideas for improvements
- If you have an idea for a project on GeoTag-X, you can send us an email at eleanor.rusack@unitar.org .
- If you find a Flickr Albums containing relevant photos, send them to us at geotag-x@cern.ch and we will load them in for analysis. (n.p. Geotag-X stores URLs of photos, not the photos themselves)

Motivation behind Geotag-X

In disaster situations or other humanitarian crises many photos are taken by different people on the ground: the local population, field workers working in the response effort, the media etc. These photos are a potential gold mine of data that could be used in not only the immediate response effort but also future recovery and preparedness work.

We hope that by developing a series of questions for analysing the photos based on field-based needs assessments we can extract meaningful, relevant, and structured data from these photos, and link this data with that collected in the field. The manpower needed to process the incredible number of photos coming out of these humanitarian crises makes

this impossible for a single organisation. Therefore we are turning to the crowd, asking them to help analyse the photos.

Current methods of collecting data in the field during a humanitarian crises can be limited in that they are very time consuming and do not cover areas that cannot be accessed. Photos can be taken quickly and over a much larger area that a traditional field-based assessment, they can also be taken by people in areas that the response teams cannot reach. We hope that these photos, taken by people in the field and then rapidly analysed by the crowd, could provide a new source of data to complement the traditional field assessment and fill in gaps left by some of these limitations.

The final aim is to have a open source tool and associated analyses developed that can be taken by anyone working in a humanitarian crisis and quickly and easily adapted to their needs.

Background

We have built a platform, GeoTag-X, for analysing photos, and a series of test projects. However now we need to see if this idea really works. Can we get anything useful out of these photos? Is the crowd able to learn to do the analysis successfully and are the results reliable? To do this we need volunteers! The system is still in beta mode and we need as many volunteers as possible on there analysing photos so we can collect enough data to assess it's suitability as a tool in disaster response.

So for this data sprint we have two aims:

- **To analyse all of the photos in two projects on Geotag-X: the Emergency Winter Shelter project and the Somalian Drought project.**

- **Have as many people as possible fill in our survey about their motivations and experience using GeoTag-X.** This will help us understand how we can motivate and engage volunteers in the long term

Use-cases/Applicability

An example of where this sort of tool could be useful is our Emergency Shelter Assessment

- http://geotagx.org/app/wintershelter/

If Geotag-X became operational it would be a tool used during field assessments of shelter. Locals on the ground, as well as field workers, could easily and quickly take photos of shelter as they move around an areas. The photos can be sent to Geotag-X and very quickly analysed by our volunteer community. Data produced would cover a much larger area than possible in a normal field survey and therefore would complement the extensive data collected in the field by also giving an indication of the how well shelters are prepared for winter in areas not covered by the field survey.

Feasibility

This data sprint is an important part of determining the feasibility of a tool such as Geotag-X.

Sprint deliverables

Encourage people to analyse as many photos as possible over the two days so that we can collect as much data as possible for the project.

Time required

One photo takes between 1 and 3 minutes to analyse so you can spend as little or as much time as you like on this.

Do it from home or come into one of the spaces to work with others. We will be at CERN to answer questions.

Skills required

Not much, just the ability to use a web browser. We have very short tutorials for each application to show you what you should be doing, then you can get right into the analysis. Don't worry if you feel uncertain about what you are doing (this will definitely happen in the beginning!), we ask at least 30 people to analyse each photo. So if you make a mistake with one its OK.

Post-sprint follow-up(s)

Data produced will be summarised and made available via the Humanitarian Data Exchange and raw data can be accessed via Geotag-X's API. Results from the data and the surveys will be used to help us improve the site and volunteer engagement. We will publish the results via blogs and online articles so you know the outcome.

Misc. notes:

- http://geotagx.org/app/wintershelter/
- http://geotagx.org/app/category/somalidrought/

Hi guys, reporting in from Palais des Nations this morning. Anyone there? [REDACTED] or [REDACTED] on Skype. Thanks!

Hi - just registered for GeoTagX and was presented with what I assume is the pre-survey. Some comments:

- Some of the first questions are hard to answer without having used the tool, yet you're given the survey at registration time. Perhaps pop up the survey after someone's done 3 or 4 classifications, or else remove those questions?
- Only male and female genders allowed as responses, perhaps add "Other" and "Prefer not to say"?

Tried out the Winter Shelter project - comments:

- Tutorial is good, though would be good to have more than one tutorial image available
- Would be good to have a progress / status counter to help "gamify" the task e.g. displaying how many photos / what percentage photos you've tagged, flashing up a thank you message if you've tagged 10 photos - this should appear on the page that you're tagging things on.
- Website seems to be grinding to a halt as all three of us at the Edinburgh site are trying to use it at the same time

Thanks for these comments! I have added them to our github issues list for the project template - you can find them here: https://github.com/geotagx/geotagx-project-template