ICRI sustainable connected cities

Research on Humanitarian Mapping:
The Community in the Data

Martin Dittus · ICRI Cities, UCL · @dekstop

27th January 2016, Oxford Internet Institute

Humanitarian OpenStreetMap Team

MISSING MAPS

Which “community”?

HOT has many actors.

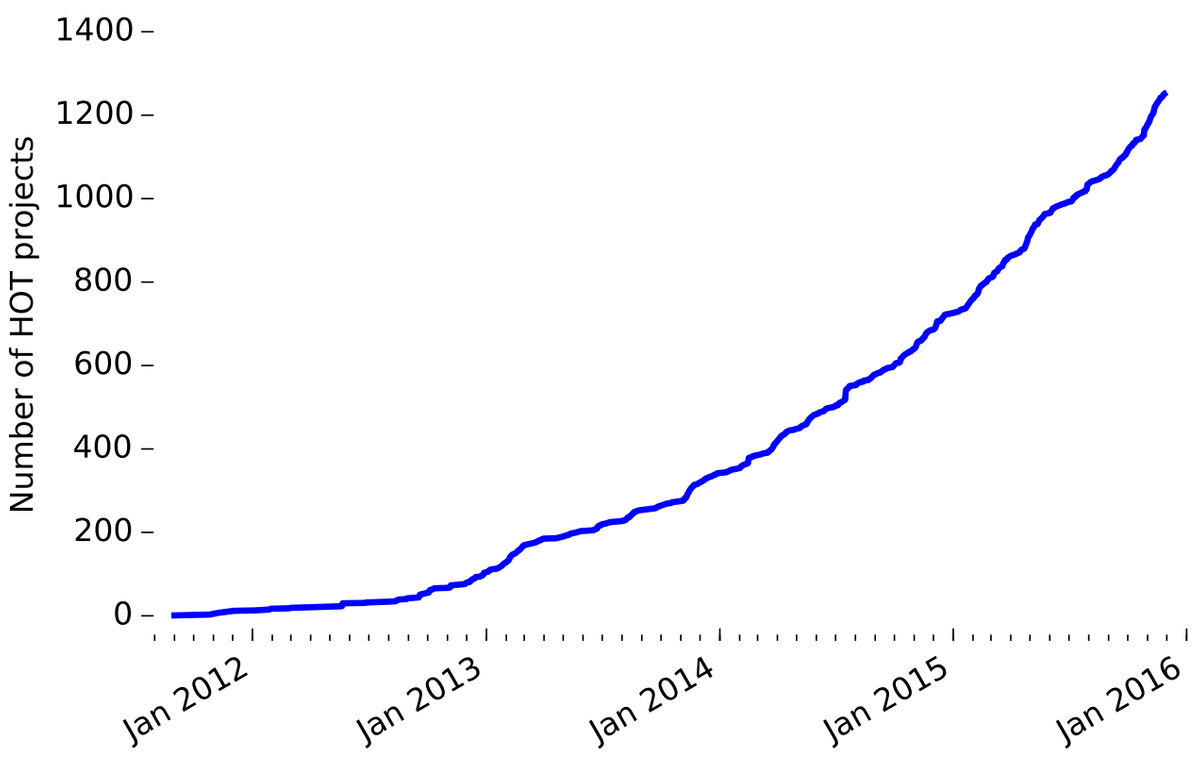
Aid organisations, field workers, local residents: data gatherers, beneficiaries, … many parties are involved, online and offline.

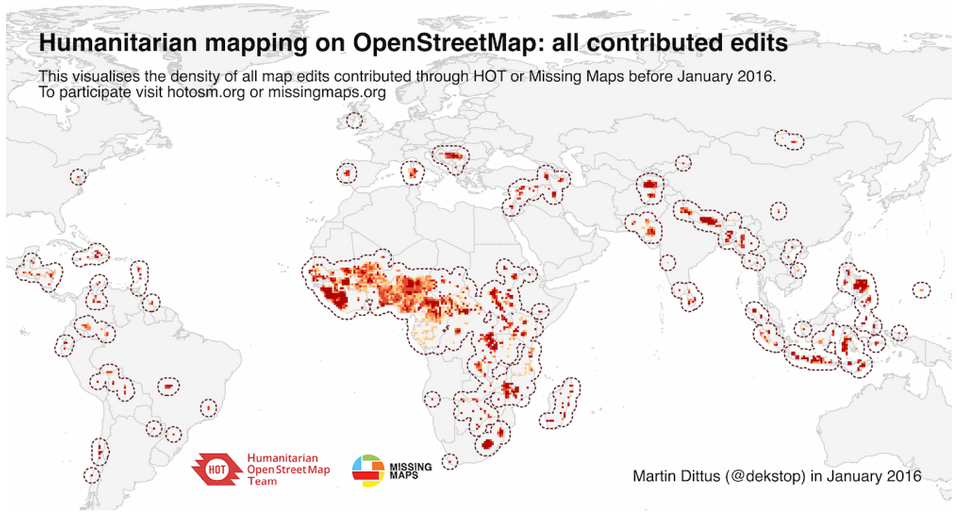
My work is focused on the largest constituency: the **thousands of online contributors tracing imagery** on the HOT Tasking Manager.

For a simple reason: **all their contributions are public**.

HOT contributor activity to date…

- 1,000 projects have received contributions,
- 20,000 online contributors have participated,
- 120M changes were made to the map,
- involving an estimated 150,000 hours of volunteer work.

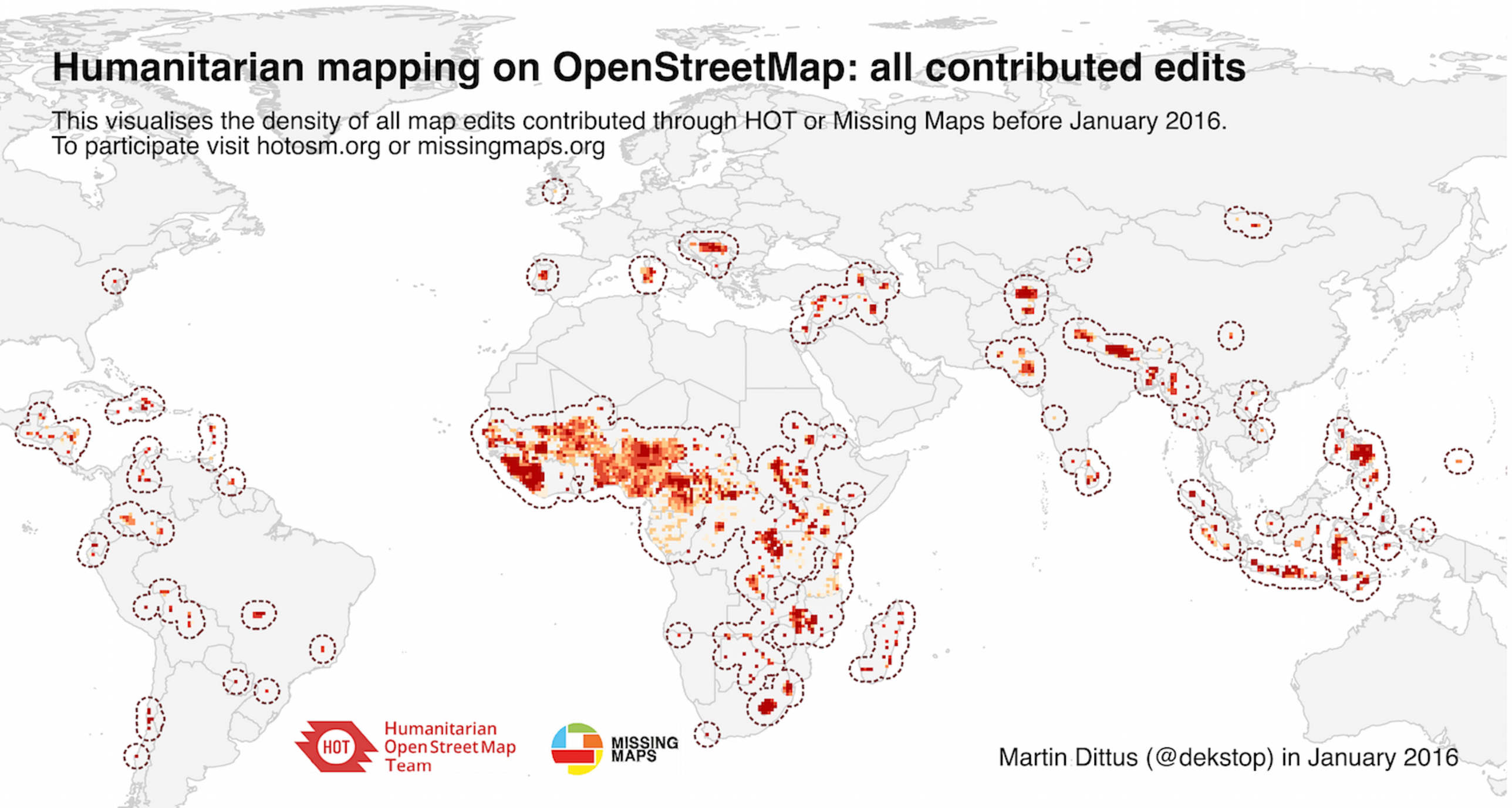




Martin Dittus · ICRI Cities, UCL · @dekstop

Humanitarian mapping on OpenStreetMap: all contributed edits

This visualises the density of all map edits contributed through HOT or Missing Maps before January 2016.
To participate visit hotosm.org or missingmaps.org

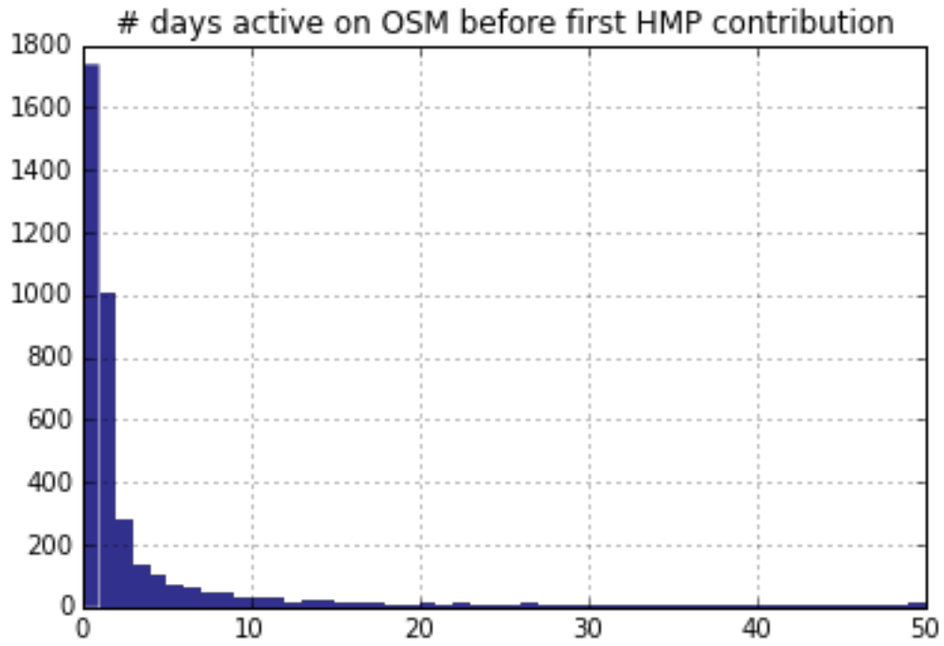


Martin Dittus (@dekstop) in January 2016

Most contributors are new to OSM.

- 72% sign up with a new OSM account.
- Mean prior experience: 22 days of OSM contributions.
- Median: 0 days.

… i.e., a long-tail distribution of prior OSM experience.



Humanitarian OpenStreetMap Team

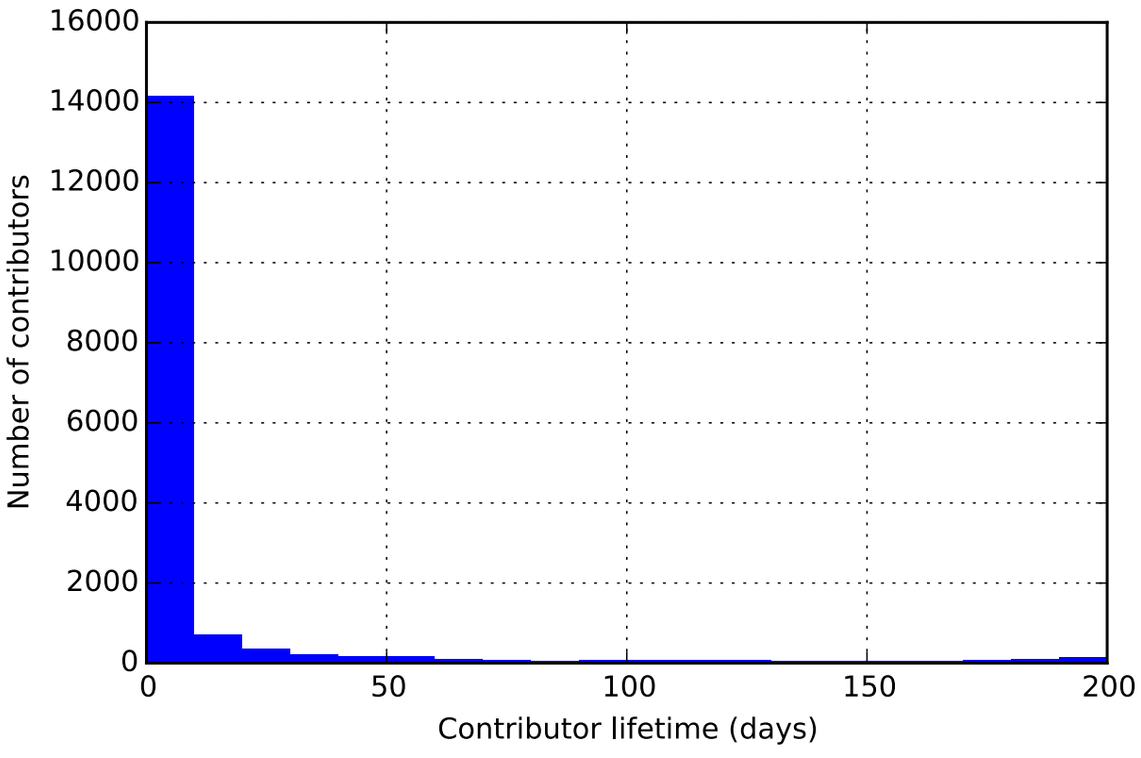
MISSING MAPS

ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop

High contributor turnover.

- Median contributor lifetime: 1 day (top 20%: >16 days)
- Median effort: 70 minutes of work (top 20%: >225 mins)



May sound small for a volunteer org. For an online community it's massive!

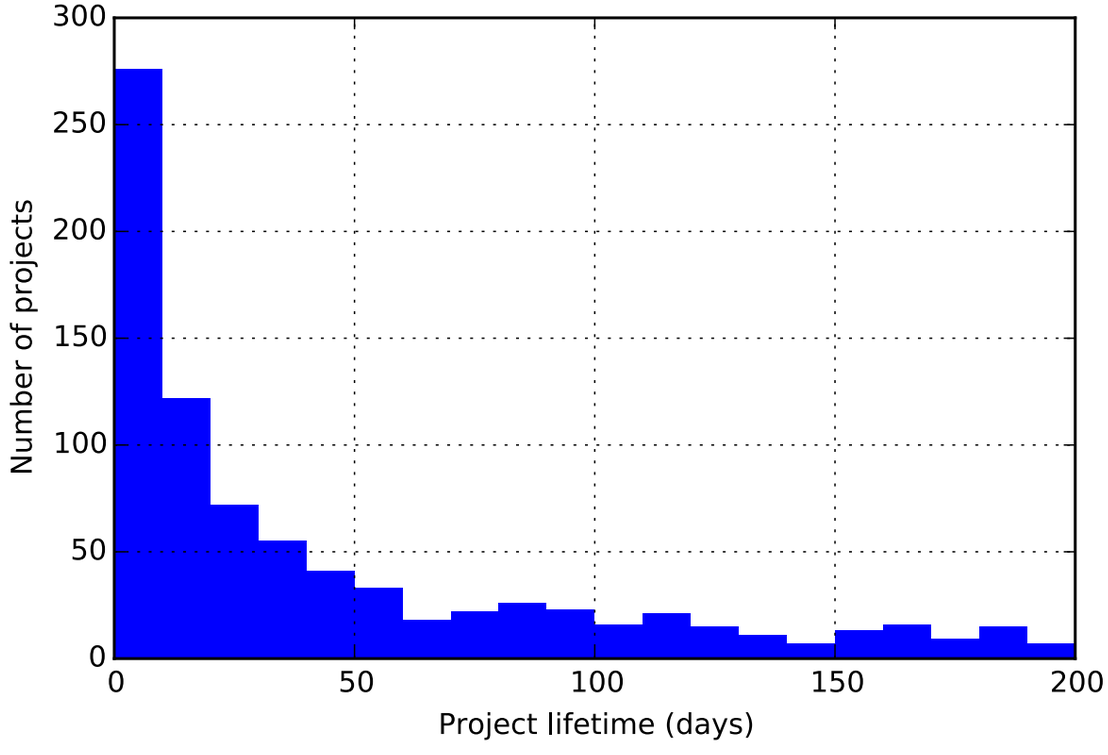
Humanitarian OpenStreetMap Team · MISSING MAPS · ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop

This extends project lifetimes.

- Median project lifetime: 34 days (top 20%: >176 days)

The ability to **distribute effort across a large number of participants** allows for types of work that otherwise would be infeasible



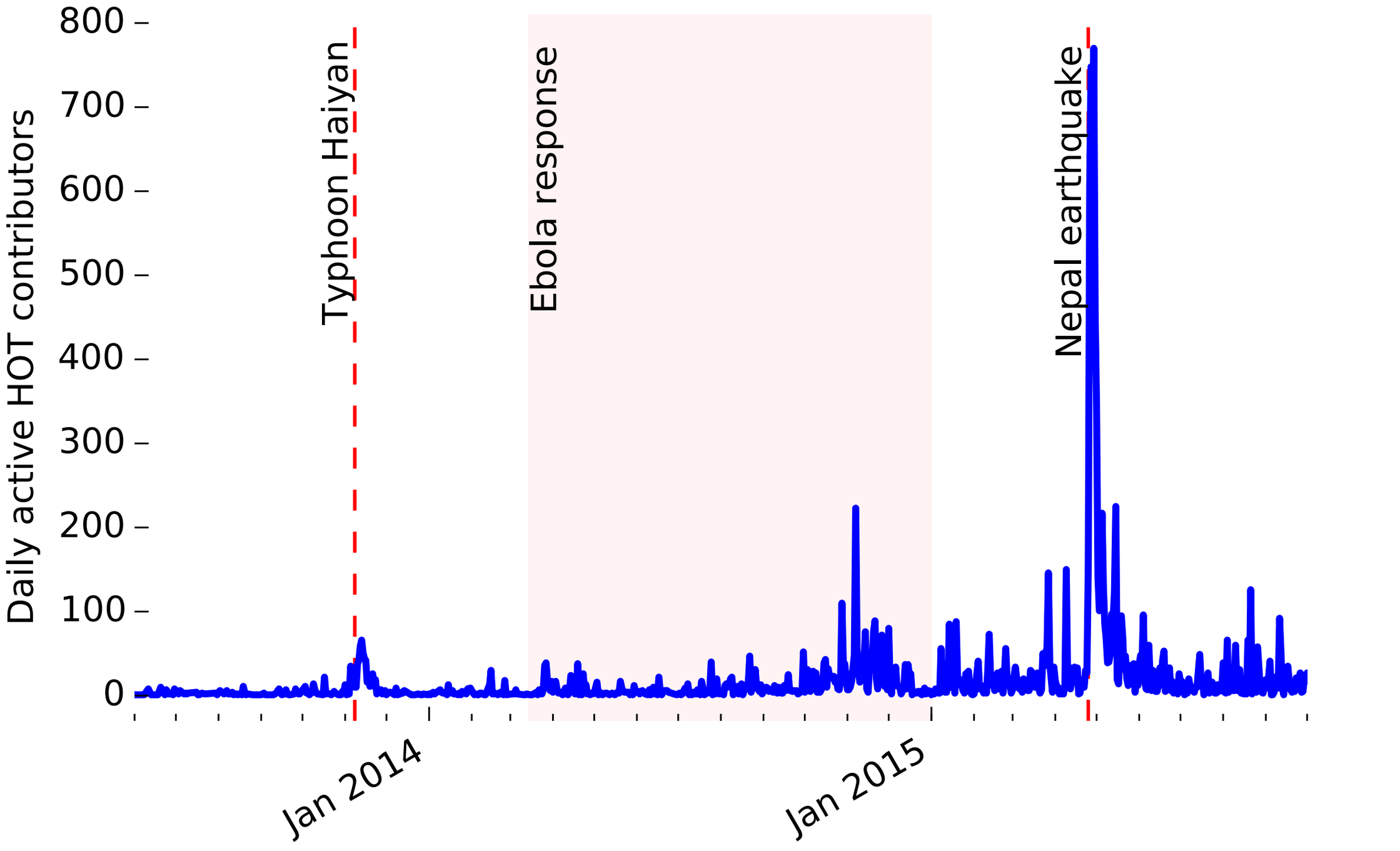
Humanitarian OpenStreetMap Team

MISSING MAPS

ICRI sustainable connected cities

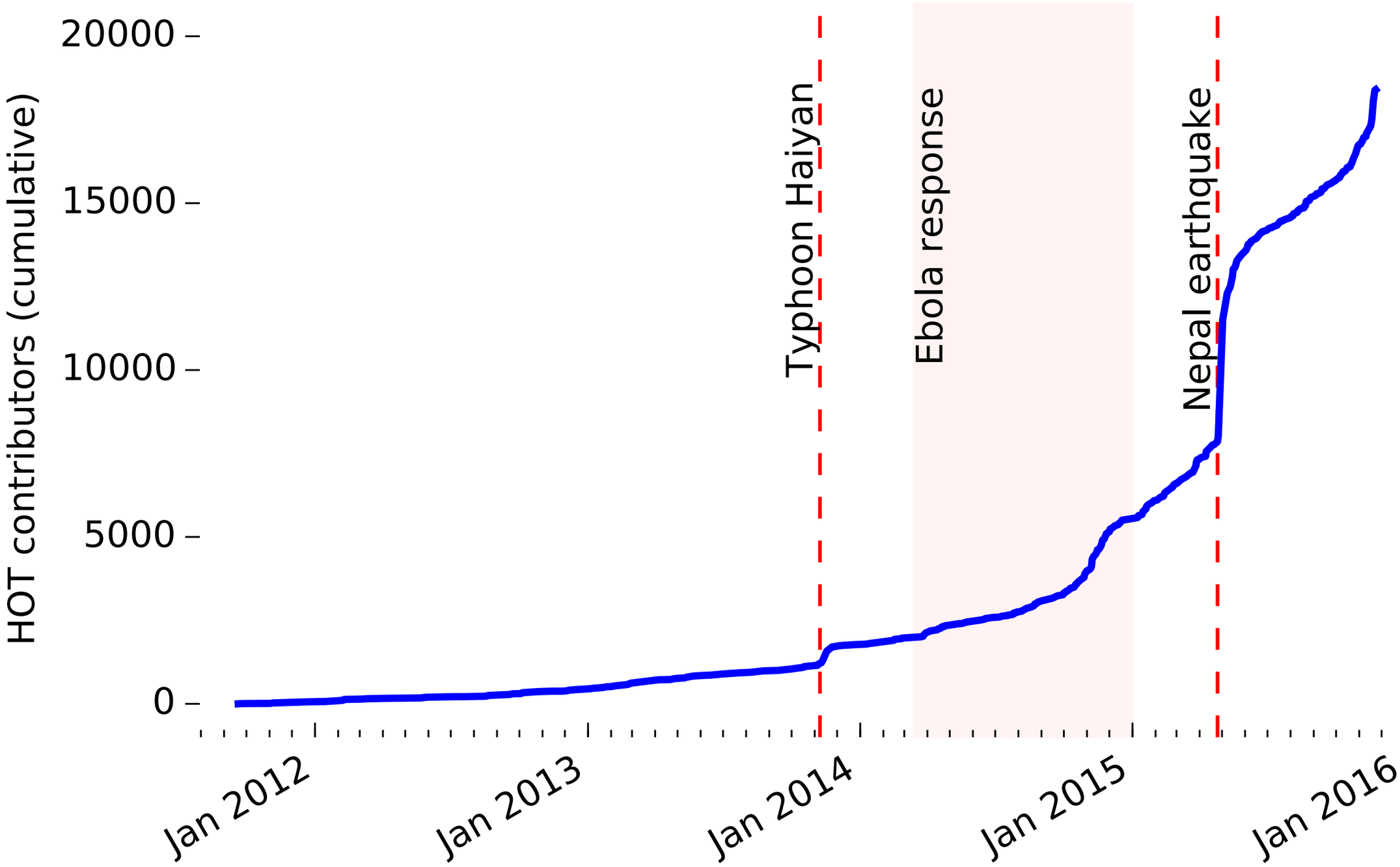
Martin Dittus · ICRI Cities, UCL · @dekstop

Activity peaks during disaster events.

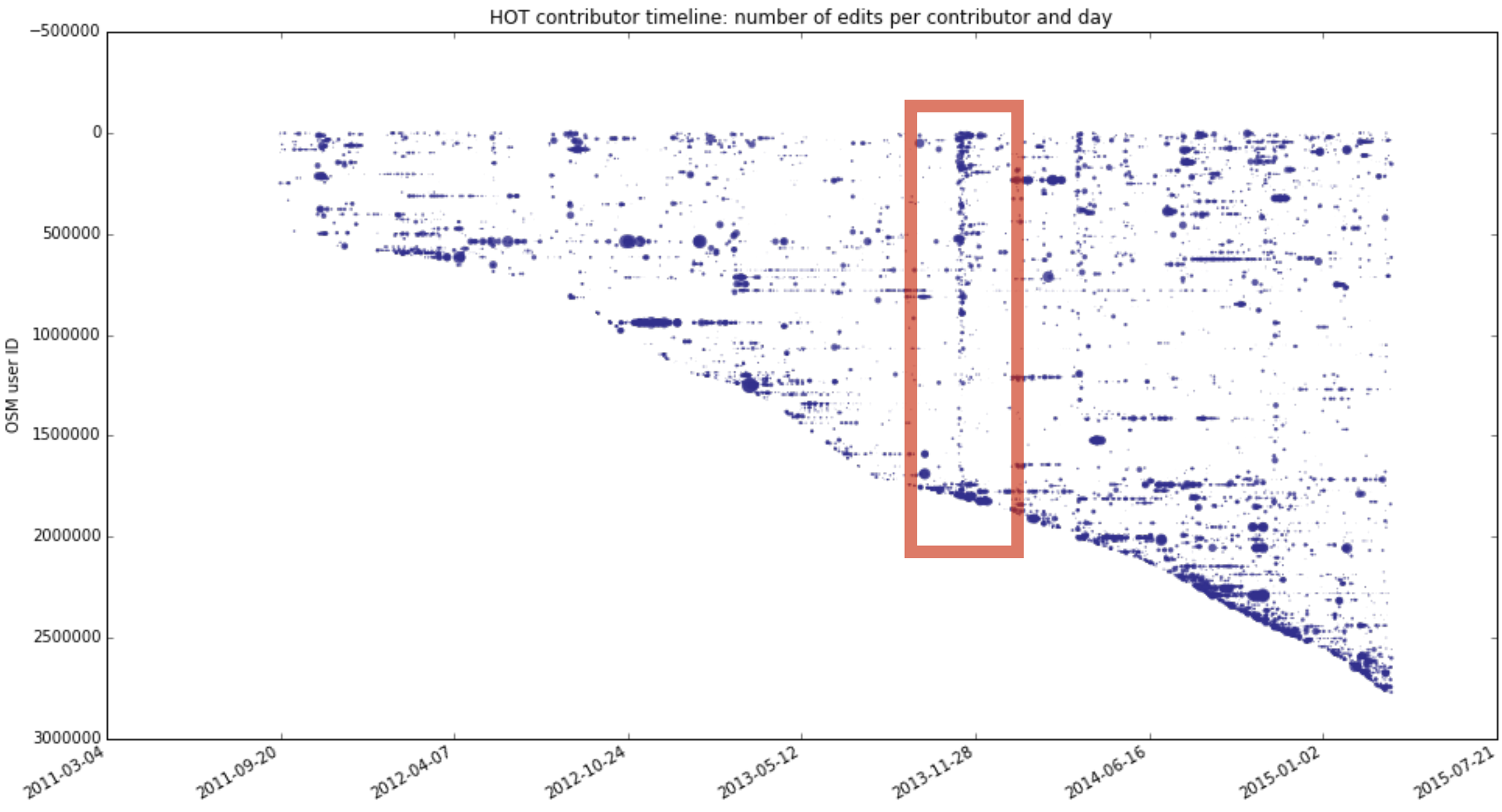


Martin Dittus · ICRI Cities, UCL · @dekstop

These also have a recruiting function.

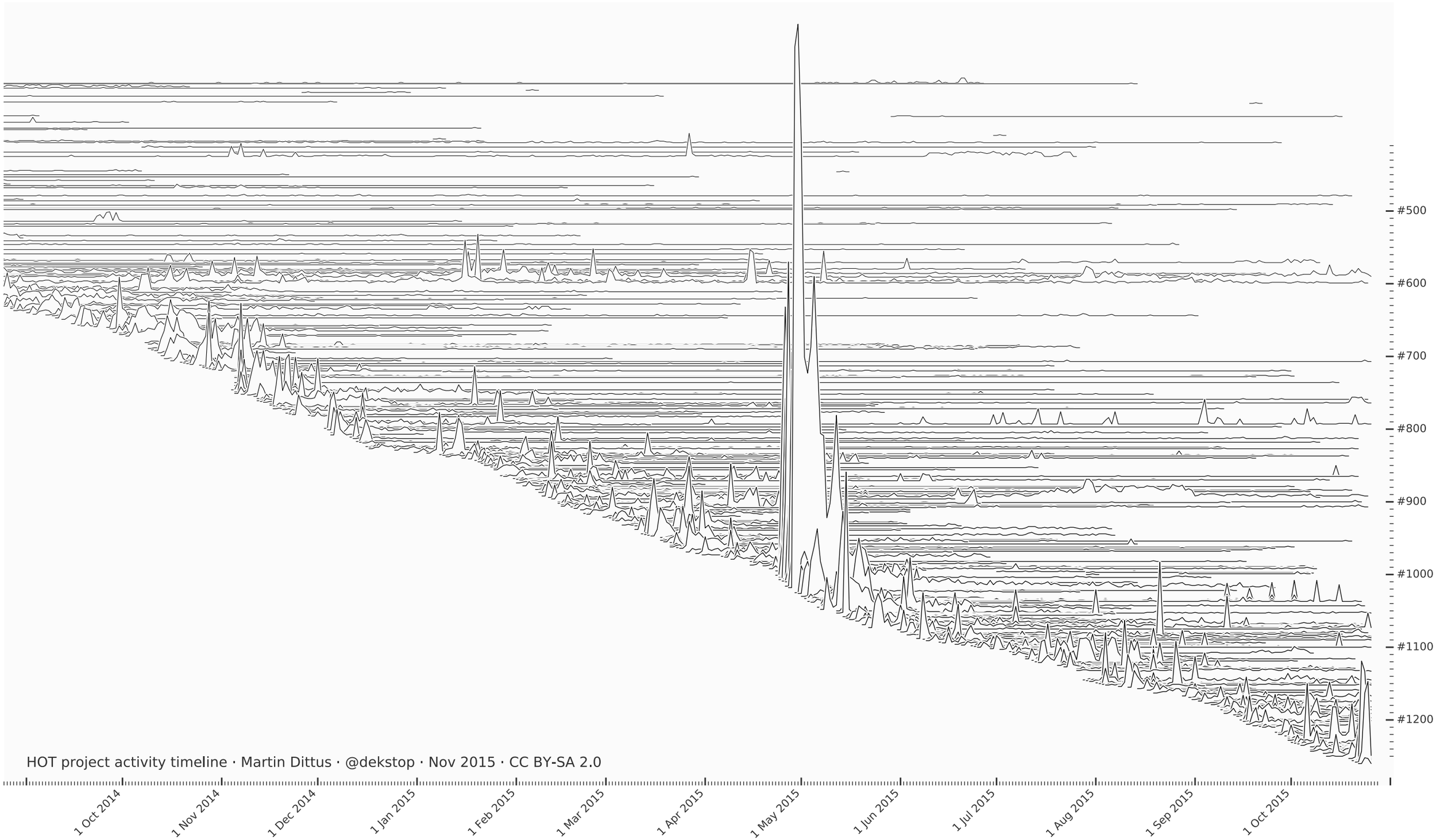


Other means of capturing collective activity…



Humanitarian OpenStreetMap Team · MISSING MAPS · ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop

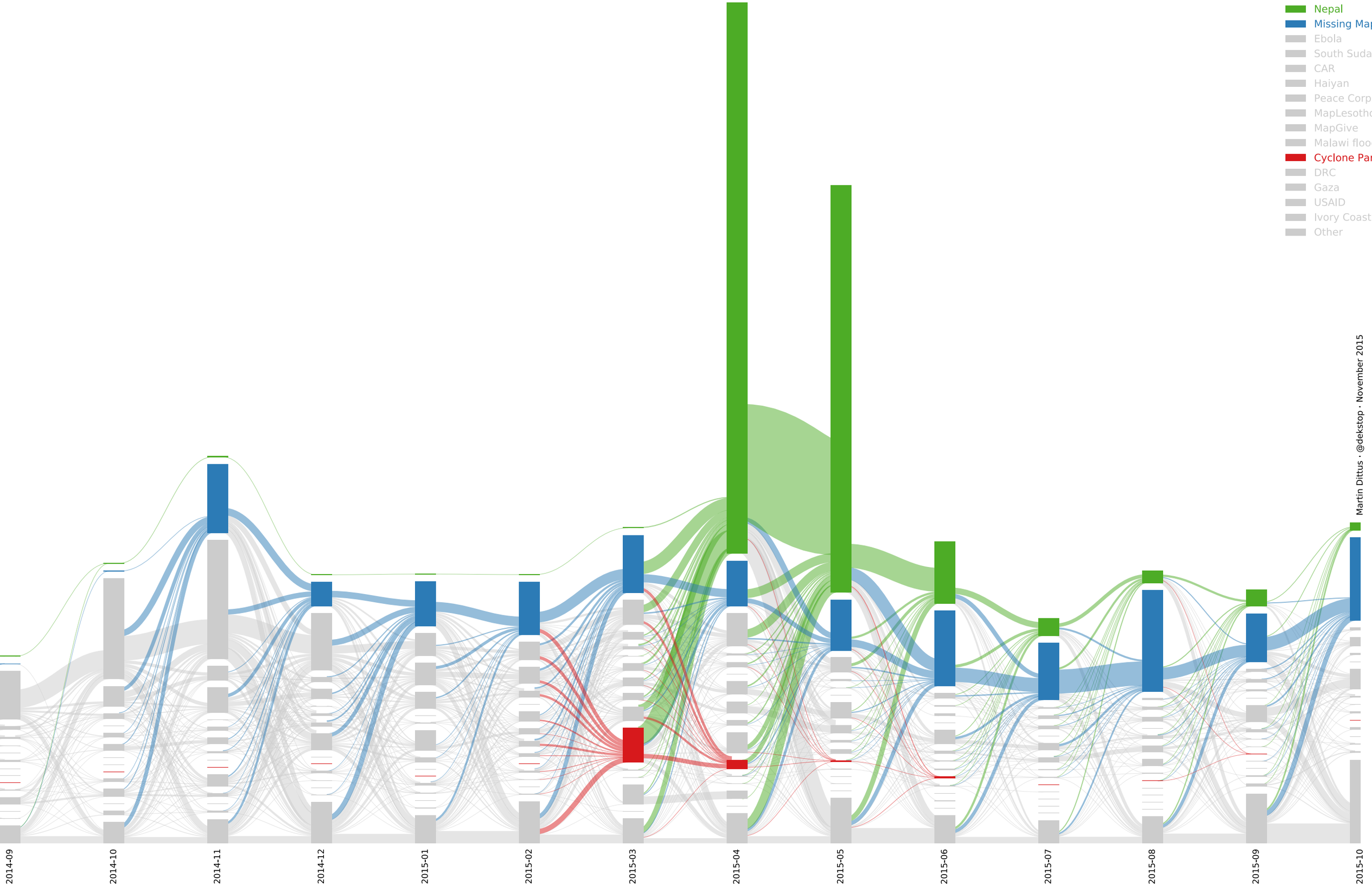


Humanitarian OpenStreetMap Team

MISSING MAPS

ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop


Nepal
Missing Maps
Ebola
South Sudan
CAR
Haiyan
Peace Corps
MapLesotho
MapGive
Malawi floods
Cyclone Pam
DRC
Gaza
USAID
Ivory Coast
Other
2014-09
2014-10
2014-11
2014-12
2015-01
2015-02
2015-03
2015-04
2015-05
2015-06
2015-07
2015-08
2015-09
2015-10
Martin Dittus · @dekstop · November 2015

My research: contributor engagement

(Developing a better understanding of what factors are involved in sustained contributor engagement.)

More specifically…

The Missing Maps challenge

According to organisers:

How to **grow sustained capacity** for proactive work, in the absence of urgent causes?

We would add…

How to **increase retention**?

So that more participants are already prepared when they join a new mapping effort.

Humanitarian OpenStreetMap Team · MISSING MAPS · ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop

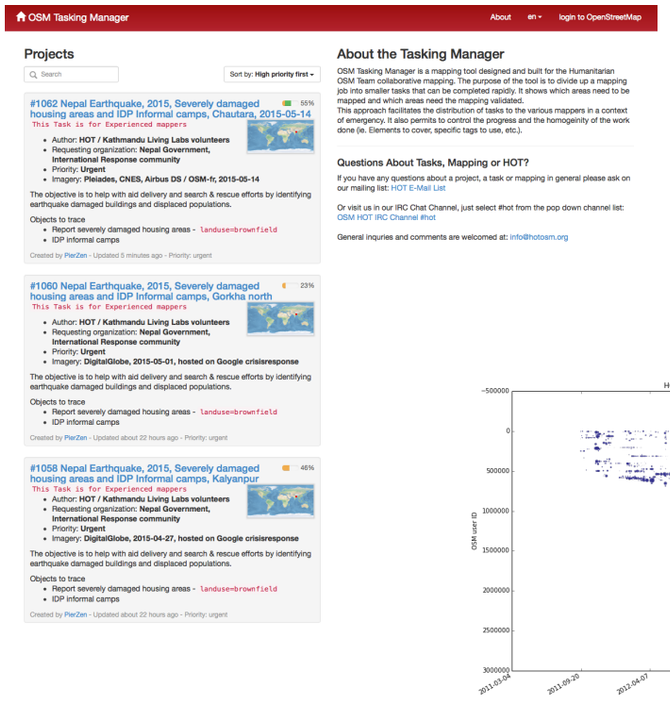
Research approach: large-scale quantitative (forensic/obs.) studies

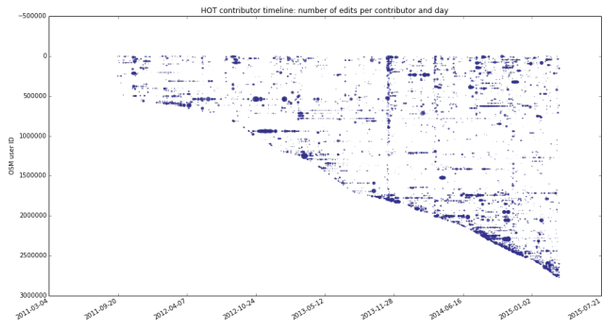
Computational methods at the core:

- **Observe contribution history** across a wide range of settings.
- **See what works** well, identify barriers to engagement.
- Infer how this may affect **long-term engagement.**

A key constraint: we cannot personally observe most of the participants.

Instead we seek to infer aspects of their experience from their actions, their contribution history.





Humanitarian OpenStreetMap Team · MISSING MAPS · ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop

Augmented with theory and domain knowledge:

Theoretical models rooted in social science.

Volunteer motivations, task complexity, self-efficacy, …

Spending time with the community.

Participant observer, mapathon attendance, many discussions online and in person, a public research diary, …

dekstop's diary

A researcher's scrapbook: understanding contributor engagement in humanitarian mapping

Posted by dekstop on 7 June 2015 in English (English)

Hallo! My name is Martin Dittus, and I'm a PhD student at the ICRI Cities at University College London. I research community engagement in the Humanitarian OpenStreetMap Team (HOT), a volunteer initiative with thousands of contributors. At its core this is quantitative work, and my main outputs are statistics and data visualisations. I also spend a lot of time with the HOT community, am a contributor myself, and have spent much of the last decade with a range of similar community organisations.

I like that my job allows me to combine my experience in large-scale data analysis with my personal interest in community organisations. I spend a lot of time exploring data sets, producing things like this:

Although I was surprised by this, this is not actually entirely unexpected. JOSM use tends to be associated with higher engagement: the most active mappers are often JOSM users.

However this does not necessarily mean that JOSM is the key trigger. It might simply reflect that the JOSM mappers at our events are a great bunch of people, fun to hang out with, and many of them know each other quite well; whereas the people at our iD tables are typically newcomers who are not yet as well-connected to the community. So maybe the difference is in the people, not the editor.

In closing I would say that we need many more observations across different kinds of settings to make these statistics meaningful. At the moment this is little more than anecdotal evidence. There's definitely space for further experiments!

Comment from joost schouppe on 7 December 2015 at 19:24
Hi Martin,

This is a slightly more elaborate answer than what I expected :)

Here's some thoughts: - the longer term retention is worth keeping an eye on as numbers increase, keeping in mind selection and social effects of course - are you sure the difference in labour hours at the events is real and not an artifact of the way JOSM and iD save information to the database? For example, if you take the difference between first and last object saved as labour time, that might affect JOSM negatively. Did you (or could you try) look at the last save during an event minus the start time of the event? Probably difficult on a larger scale, but might be worth a check at a single event. - you filtered away the people who used both editors. It might be interesting to see if they started off using JOSM and gave up to fall back to iD or if they took the other way around. If it looks like the former, one can imagine more people giving up before having anything worth saving.

Thanks again. We'll think about how we can experiment more at our events in Belgium.

Comment from dekstop on 7 December 2015 at 19:29
Ah you're right -- it is possible that JOSM captures changeset timestamps differently. In past explorations I've seen JOSM preserve timestamps for individual edits within a changeset, but I don't know enough about the editor to understand what exactly is going on.

Only 6 people used both editors at their first event, so I felt it's better to simply ignore these :)

Comment from Warin61 on 7 December 2015 at 20:22
What is the coverage factor in the confidence interval? Around 2 is usual for metrologists.

I don't think difference of the initial contribution should be a deciding factor. People learn differently. Most would 'test' at a later time, so sampling as , say 2 months later for contributions, and the number of nodes added/modified/deleted? With possible follow up at 1 and 2 years?

Three studies at complementary scales of observation…

1. Individual experience. (Completed.)

Contribution mechanics of individual projects.

2. Group experience. (Finishing.)

Social contribution environments.

3. Collective experience. (Soon)

Overall community activity.

Humanitarian OpenStreetMap Team · MISSING MAPS · ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop

Aspects we have considered to date:

Impact of **task design?**

E.g. can task complexity be a barrier? Can better guidance help?

Impact of **project purpose** and **coordination practices?**

E.g. participation in one-off disaster aid vs sustained proactive work.

Impact of **social contribution environments (mapathons)?**

E.g. corporate one-offs vs regular monthly events.

Impact of **tool use?**

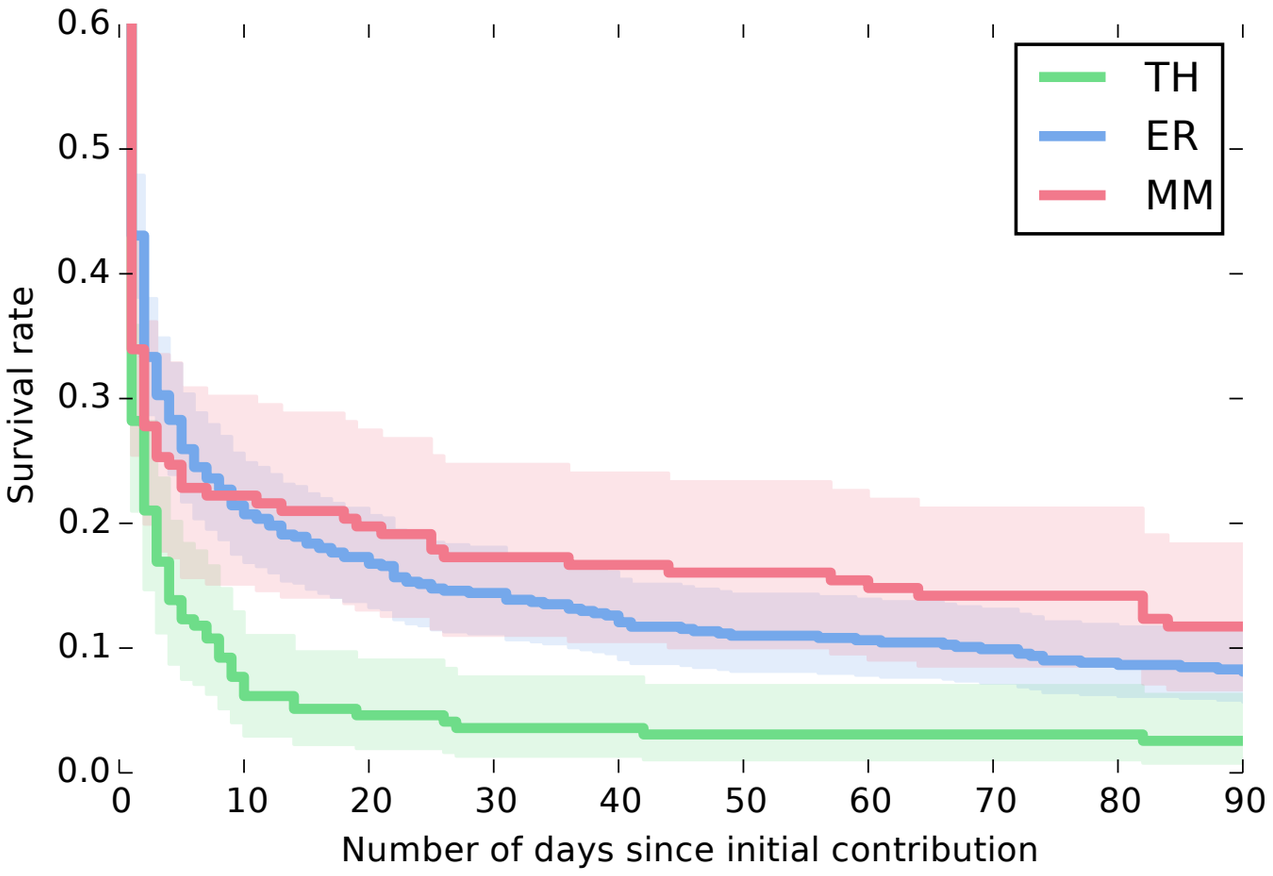
Does a more complex but powerful tool present as a barrier to entry?

Can **initial contributor performance** predict engagement?

E.g. length of session, contribution rate, task completion rate.

Martin Dittus · ICRI Cities, UCL · @dekstop

Retention of newcomers w/o prior OSM experience: Missing Maps is doing very well.



Disaster response project Haiyan: low retention.

Ebola, MM: high retention.

Possibly related to coordination practice? Mapathons, activities over longer periods, email notifications, …

Dittus, Quattrone, Capra (2016): "Analysing Volunteer Engagement in Humanitarian Mapping: Building Contributor Communities at Large Scale", CSCW 2016.

Martin Dittus · ICRI Cities, UCL · @dekstop

Impact of task design on contributor labour hours.
Do complex tasks turn people away?

We assessed task complexity for 100 projects across 3 different HOT initiatives:

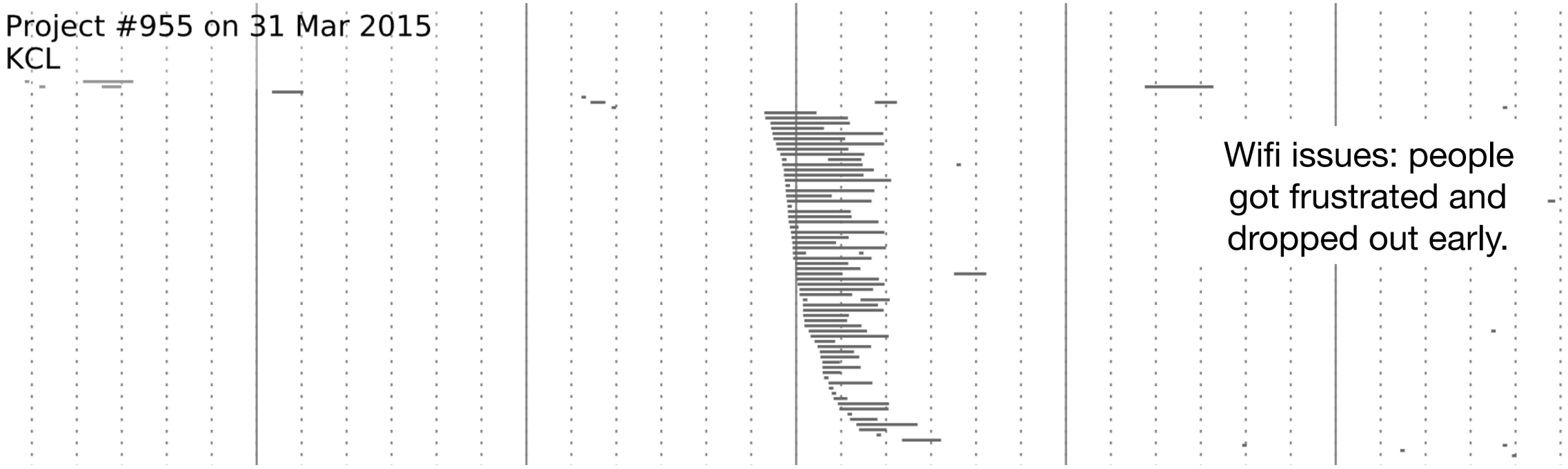
Aspect	Variable	Description
Motivation	*has_context*	Does the project description state an explicit purpose?
Visual complexity	*urban_density*	Is the mapped region rural (simple), mixed, or urban (complex)?
Task complexity	*num_concepts*	How many different types of map objects are to be mapped?
Task complexity	*building_trace*	Are buildings to be mapped as points (simple) or polygons (complex)?
Guidance	*num_cues*	Number of information cues provided in the documentation?
Guidance	*num_tag_ex*	Number of tag examples listed?

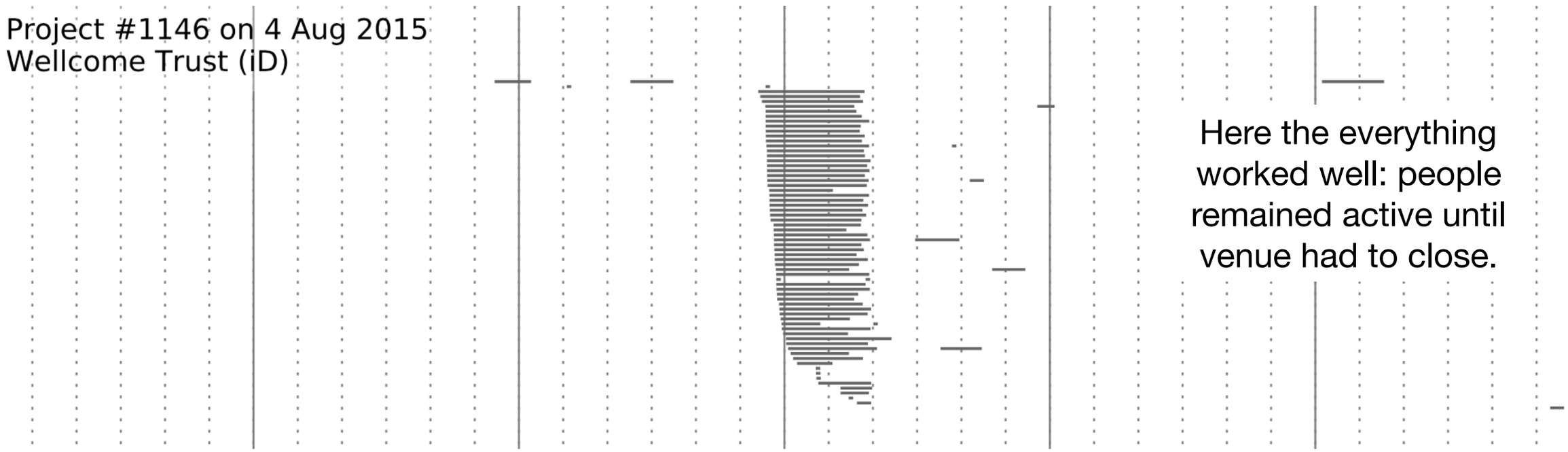
Table 3. Task design feature vector produced by our task analysis.

On Missing Maps projects, **people remained active for longer on tasks that involved the mapping of less than three distinct map features**. No other factors had a consistent impact.

Dittus, Quattrone, Capra (2016): "Analysing Volunteer Engagement in Humanitarian Mapping: Building Contributor Communities at Large Scale", CSCW 2016.

Currently assessing impact of mapathon settings.





Humanitarian OpenStreetMap Team · MISSING MAPS · ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop

Mapathon attendee mix, tech issues, food breaks, …

According to our observations, the specific contribution setting has little bearing on subsequent newcomer engagement.

Food breaks (opportunities to socialise/bond): no observable impact on newcomer retention.

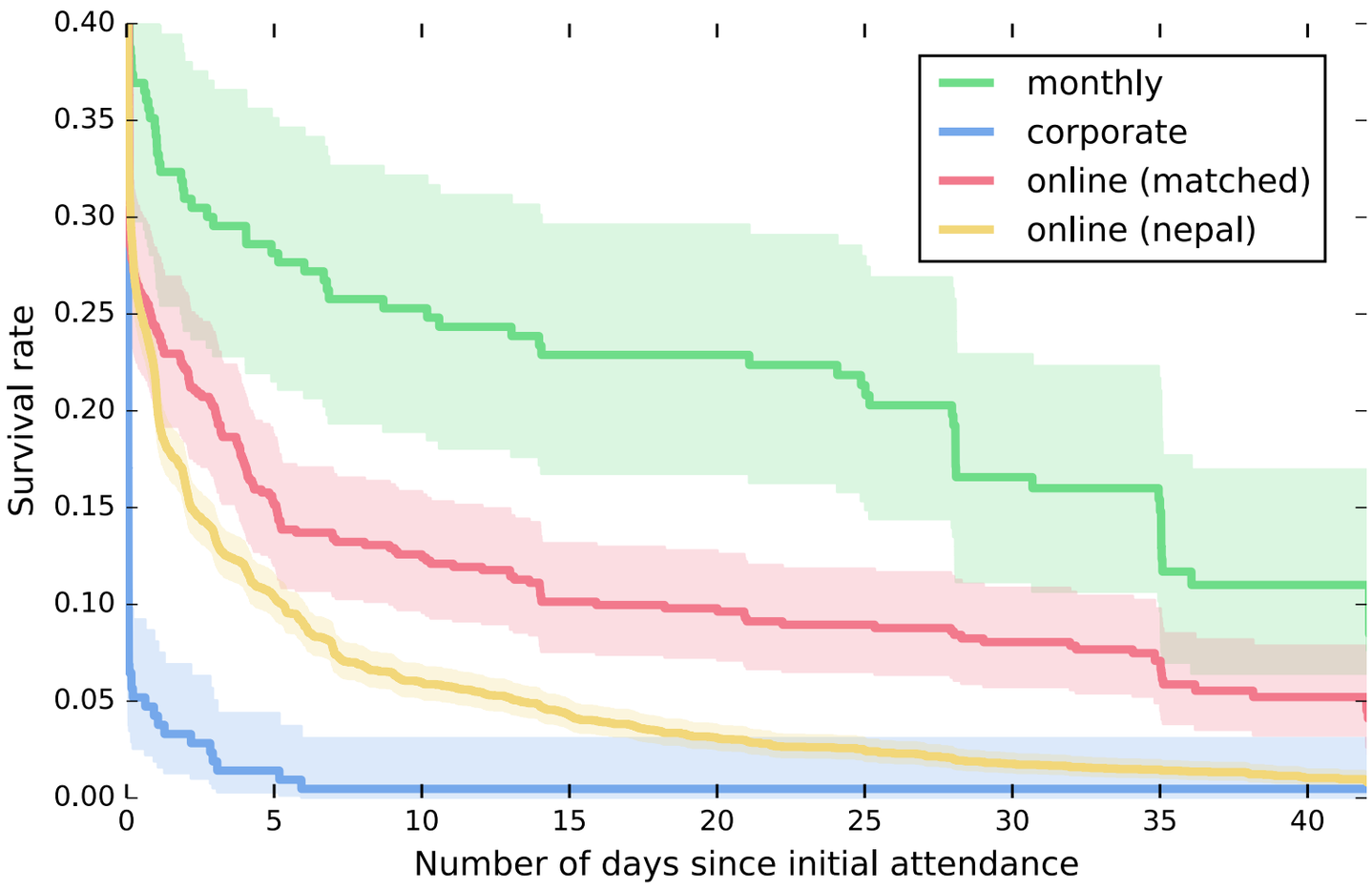
Aspect	Variable	Outcome	ρ_S	p
Cohort	*is_monthly*	*activation_rate*	0.62	< 0.02
	is_corporate	*activation_rate*	-0.62	< 0.02
Attendees	*hot_mappers*	*activation_rate*	0.52	< 0.05
	home_mappers	*activation_rate*	0.56	< 0.02
	prev_attendees	*activation_rate*	0.58	< 0.02
Setting	*tech_issues*	*completion_rate*	-0.54	< 0.05

Table 8: Significant correlations between mapathon features and measures of engagement.

Tech issues may reduce event outcomes, but don’t significantly harm retention.

Attendee mix: presence of experienced mappers is associated with increased newcomer retention, but causality is unclear. Access to peer support? The motivating presence of a community of practice? Etc.

Mapathons can yield significantly higher retention compared to online cohorts doing the same work.



Monthly mapathons in London: high retention.
20% newcomers still active after 28 days.

Comparable online cohort: 10% after 28 days.

Corporate events: virtually no retention. That's 0%.

However the **participation context really matters**!

Humanitarian OpenStreetMap Team · MISSING MAPS · ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop

Findings to date (1/3)

Many **highly engaged contributors can be identified early**.

Increased activity in first session, activity in the following days.

(Highly engaged contributors are likely to remain engaged.)

Complex task designs can be a deterrent for certain contributor groups.

However most task designs we observed yielded comparable engagement outcomes.

Findings to date (2/3)

At mapathons, **technical problems** can impact activity during the event, but they have no significant longer-term effects.

Other observed differences in settings had no significant effect.

The **presence of more experienced attendees** at mapathons can improve newcomer activation.

The specific causal factors are unclear. Presence of a community of practice, shared experience, peer support?

Findings to date (3/3)

Different HOT initiatives are associated with different engagement profiles.

Disaster responses attract new contributors, but have low retention.

More sustained engagement: Missing Maps, Ebola: long-term initiatives with regular mapathons, email notifications.

Different participation contexts are associated with different engagement profiles.

Monthly public Missing Maps mapathons in London have much higher newcomer retention than the corporate mapathons by the same team.

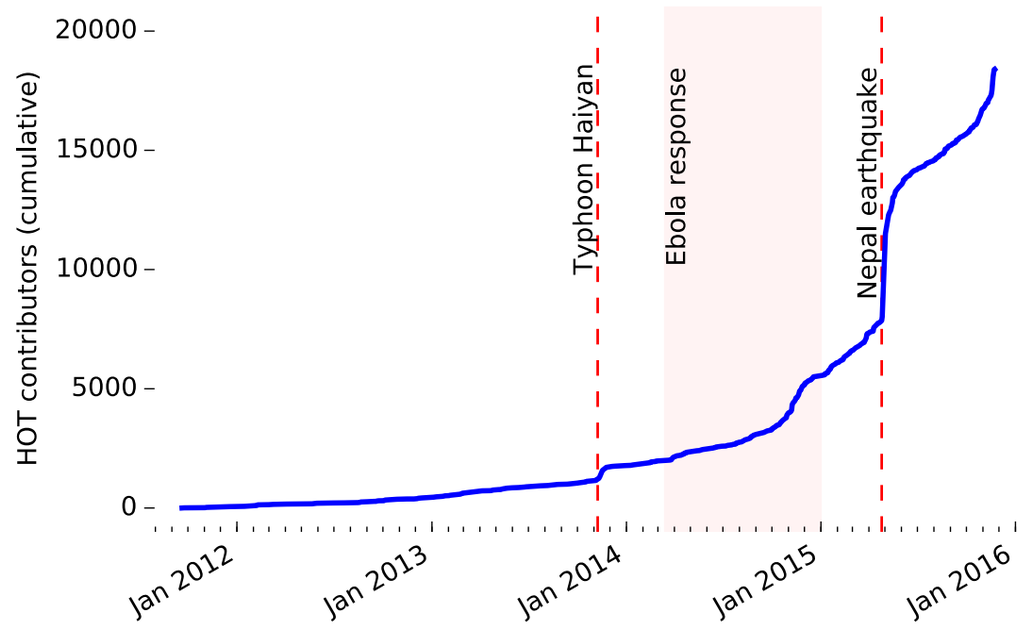
Newcomers doing the same work online: twice as likely to leave early.

Humanitarian OpenStreetMap Team · MISSING MAPS · ICRI sustainable connected cities

Martin Dittus · ICRI Cities, UCL · @dekstop

Implications (1/2)

Overall, HOT seems to be doing very well. It appears **relatively easy to get people interested** at first.



However it's **harder to get sustained engagement.**

That’s only partly about the technology, the process.
It’s likely also about the **human experience**!

- What do people get out of it?
  Personal motivations, social experiences, ...
- Will vary across people and participation contexts.

Implications (2/2)

Can we imagine **online support experiences that imitate the mapathon experience**? Expert guidance, peer support, the presence of a community of practice, …

If the aim is sustained engagement, then **indiscriminate growth may not be advisable.**

Instead:

- **Identify highly-engaged contributors early** and give them support.
- **Identify participation contexts** that are more likely to yield highly motivated contributors.

Impact of the work to date

First paper published at CSCW 2016.

Research diary now widely referenced in the community.

https://www.openstreetmap.org/user/dekstop/diary

Working closely with organisers at HOT, MSF, BRC, ARC, …

A recurring conversation: "Can you help us understand X?"

HOT groups have already picked up many of my insights.

E.g. Missing Maps organisers have made changes to task design, email notifications for mapathon attendees, and other aspects of their coordination practice.

Humanitarian OpenStreetMap Team · MISSING MAPS · ICRI sustainable connected cities

Thank you.

Martin Dittus · @dekstop

https://openstreetmap.org/user/dekstop/diary

Humanitarian OpenStreetMap Team

MISSING MAPS

ICRI sustainable connected cities