

OpenStreetMap Foundation, Licensing Working Group **Agenda & Minutes**

Thursday February 8th 2018, 20:00 - 21:00 UTC

final

Present: Simon, Kathleen, Christina, Paul

Guests:

Apologies:

Minutes by: Dorothea

1. Adoption of Minutes of last meeting

2018-01-11:

<https://docs.google.com/document/d/1H2SmTUZbCAugLjLLjIZ0sckQkHHzLOxzeovqCrHIYd8w/edit>

Proposed: Simon

Seconded: Kathleen

- 2016-05-06 **Dermot** licence question wrt Ordnance Survey Northern Ireland Open Data
- 2016-05-06 **Simon** arrange procedure with board for signing of NDA <https://docs.google.com/document/d/106JVPBsmFVKIPZw5bGLIvADLM8NbvCvqdgS-cH8F2IY/> some still missing pls send them to Simon for forwarding.
 - Anyone who hasn't signed yet, contact Paul by email
- 2017-03-02 **Simon** determine existing obligations towards sources listed on the copyright page
- 2017-05-04 **Simon** follow up with lawdit on JP TM
- 2017-05-04 **All/Simon** review import guidelines wrt licence "approval"
- 2017-09-05 **Simon** ask Lawdit for quotes for registering for the additional class suggested
- 2017-11-02 **Paul** draft doc on GDPR related changes to planet distribution
- 2017-11-02 **Simon** draft recommendation doc on GDPR consequences
- 2018-01-11 **Simon** fill in numbers for 2017 expenditures
- 2018-01-11 ~~**Simon** send the current 2018 Budget draft to board for consideration~~
- 2018-01-11 ~~**Simon** modify the trademark policy documents on the OSMF wiki, to reflect that they are now official and in force~~
- 2018-01-11 **Kathleen** create a revised draft of the DMCA takedown procedure
- ~~2018-01-11 **Kathleen** update the calendar to reflect the new LWG meeting dates~~
- 2018-02-08 **Simon** to widen the audience of the GDPR draft document to the Advisory Board and other involved groups
- 2018-02-08 **Simon** to find out what specific pieces of metadata the companies are using and for what reasons
- 2018-02-08 **All** to report to back in two weeks

2. Reportage

DMCA takedown procedure

- Kathleen started drafting a new policy, framed as copyright policy with some DMCA related procedures to match US law # *Discussion notes on the AOB section.*

3. GDPR Document

https://docs.google.com/document/d/1EjccQNm3awl7eQk1jGYyoGJVavJG_bEFX8iCMEuC9U/edit?usp=sharing

(some points have been reordered)

GDPR

Applies to any entity, including natural persons, regardless of whether they are located in Europe or elsewhere (Enforcement is an independent issue).

Wikimedia and personal info

- They have personally identifiable information too and while it might be harder to analyse, you can scope down to who a person is.

Wikimedia & GDPR

Paul had contact with WMF person that will deal with GDPR.

- WMF main concerns: operational.
- They do not intend to be removing contributed crowd-sourced data based on GDPR
- Probably will not remove the ability of the community to do the kind of analysis and bots that they currently do.
- Probably not intending to rely on the position that they are outside of EU restriction.

One of our issues: generating profiles from the contributor metadata (such as editor used, time, device)

Not enough guidance regarding metadata, we could take a conservative approach removing all of that but not sure that the law actually requires that.

Suggestions

- Update privacy policy making it more explicit, so that community members can take this data and know what they are allowed to do with it
- [Planet dump](#): provide metadata-free dump with some barriers.
 - Right to be forgotten problem: We have an account removal request every day.
 - If you take the data from planet.osm.org and process it in a way that you generate personal data that can be attributed to one person, you would be subject to GDPR in any case regardless if the OSMF says something or not.
- Historic planet dumps: could turn off public access & offer research access.
- API: add a login to access metadata and state that you can't use the API for reverse engineering personal data.
 - Special cases: services that produce data which is associated with

- individuals (e.g. HDYC, OSMCha) # *see discussion below*.
- API access via login has been done in the past for different reasons (CGI map).

Community projects

If we say that community projects that process metadata are actually data processors in GDPR terminology, long agreements would need to be signed, which might scare people and would make it unattractive to provide such a service.

Question

How can we still enable our community projects with as small overhead as possible and still apply GDPR?

Potential problems

1. with the few services which produce data which is associated with individuals (e.g. [HDYC](#)), as well as
2. adding entry barriers to any new similar efforts

HDYC

- Need to login to access.
- Currently users that have asked to be deleted from our database can be found there, for technical reasons.

OSMCha

- Does not require login in general (just for some functions)

Underlying assumption 1

We assume we would treat OSMF and community projects that are processing data, as separate entities.

Suggestion to treat the whole OSM movement as one entity/project

Treat the whole movement as one entity/project. There's no exact equivalent in common law countries but essentially it would be an unincorporated entity where everybody shares liability.

- Advantage: It would simplify relationship with HDYC, and OSMCha, as we could potentially generate a legal ledger where we are treating all processing of this data inside the OSM movement as one controller working on this data.
- Seem to have massive downsides
- In US you can have an incorporated entity, where members are part of an entity and they are not personally liable, but then you would have to have some kind of agent agreement with individuals who are working on project, a kind of employment agreement but without any money.

Suggestion to provide a feed of deleted user IDs

Provide a feed of deleted user IDs, so that such community services can decide whether they want to remove the data from public view or to keep it internally in some pseudo-anonymised form

- probably not a problem for services that reprocess the planet every week.
- probably not big addition for services feeding from live data

Users' feed

- We currently have no good way to follow renames (we don't publish renames, we only publish when someone edits with the new name).
- Deleted users can be a special case of renames

Distributed analysis software

- Analysis software need usernames or user IDs (example: [ChangesetMD](#)).
- They usually provide aggregated statistics, so at least the result of processing cannot be associated with individuals

Companies analysing OSM data

- Some companies are internally processing OSM data involving OSM usernames, for reasons such as vandalism detection (Mapbox ([OSMCha](#)), Apple, probably Facebook).
- We should have agreements with companies that want to use OSM metadata.
 - Probably won't be a problem for them.
- We would have to provide the deleted users feed to remove stuff from users that have withdrawn consent.

Underlying assumption 2

We treat this issue in binary mode: provide data with metadata or without.

There's clearly some metadata which we could supply, which allow to link contributions but not other info (userid but no timestamps, no changeset metainfo).

Action items

- Simon to widen the audience of the GDPR draft document to the Advisory Board and other involved groups
- Simon to find out what specific pieces of metadata the companies are using and for what reasons.

GDPR risk scenarios

- European authorities going after big companies first, so we can learn from them.
- An unhappy OSM contributor complaining to their national data protection officer and starting trouble - irrelevant whether we are big or not
 - Addition of parameter of limited resources. When GDPR goes into effect, national authorities might get many complaints, and due to limited resources they might have to focus on bigger companies.

Formalising internal agreements

We need to make sure we have all agreements in place (e.g. between us and sysadmins) formalising that personal information won't be shared.

Wikimedia

- won't be limiting anything to requiring a login
- already have agreements with their employees

Suggestion

Invite the Wikimedia person that will deal with GDPR to our next meeting.

Action item: all to report back in two weeks

4. AOB

DMCA takedown procedure

Background

- a lot of US case law on takedowns
- we don't have as far reaching usage experience in other countries, as in US
- difficult to have multiple policies for different countries

On data redactions

- it's not easy to remove stuff, especially if it has been added a long time ago
- what takes longest is identifying what needs to be removed

Suggestions

1. Have one overarching policy based on the DMCA principles and apply it to all notifications worldwide
2. Perhaps internally tweak it if we know the takedown request is coming from a country where we might have additional difficulties
3. Have some documentation from the DWG regarding removals for various reasons, to back up what we're saying is reasonable
4. Explicitly state what we do with the historic planet dumps

Historic planet dumps

- Valuable from a data observation point
- Are also a convenience feature
- Have a handful of consumers that want a snapshot and don't want to go through the process of generating that snapshot out of the current full history dump, which wouldn't contain redacted data
- Are mirrored to some extent by libraries like arxiv.org
- We have individual planet dumps for the last 10 years or so

Issues

- data redacted from the planet file is generally included in the historic planet dumps
- it is difficult to remove data from the historic planet dumps

History

- we have once redacted data out of historic planet dumps
- [OWG](#) has done something for operational reasons (run out of space). Unknown whether they want to keep them

Suggestions

1. turn off public access and provide access for research purposes
 - GDPR related: archive purpose is mentioned as a lawful purpose
 - can't impose any restrictions above the ODbL
 - While the contributor terms state that we can only distribute under ODbL terms, they don't actually force us to distribute anything, so we can remove data. We could argue that we could add additional conditions if somebody wanted to access for research purposes to data that we normally wouldn't distribute. The alternative is to not distribute them at all.
2. discuss with the OWG what works for them before we make the recommendation

Next Meeting:

March 8th 2018 20:00 UTC [on Mumble](#)