Evolving
the
OSM Data Model

Jochen Topf

What I mean when saying:
OSM Data Model

Nodes, Ways, Relations
Coordinates
Tags
Versions
Changesets

OSM Data Model

What I do not mean is details of tagging…

Change History

API 0.4 - May 2007
API 0.5 - October 2007
API 0.6 - April 2009

API 0.7 ?

OSM Data Model

Is it really so perfect
that we don't want to change anything
in more than a decade?

OSM Data Model

We have been talking about this forever…

API 0.7 wiki page created December 2009

OSM Data Model

I have been talking about this forever…

SotM 2013: Areas
SotM 2018: Data Model

OSM Data Model

There are a lot of issues.

OSM Data Model

OSM has been growing → We always have something important to do right now.

So we keep pushing the big picture stuff away.

OSM Data Model

It high time we tackle this!

The Study

In spring of 2022
the EWG
(Engineering Working Group of OSMF)
commissioned me to
write a study.

The Study

Write down issues that have come up.

Document ideas and state of discussion.

Figure out what's possible.

Develop a plan how to approach this.

The Study

Written for a technical audience
familiar with OSM internals.

The Study

Most of what we are talking about here
the average mapper
will never see or be interested in.

The Study

This study should be the
start of an organized discussion.

The Study

1. Introduction
2. Background
3. Issues with the Current OSM Data Model
4. Discussions and Possible Improvements
5. Next Steps
6. Software and Systems
7. References

Issues

- References Between Objects
- Missing Area Data Type
- Coastline
- Fuzzy and Large Areas
- Building a Navigation Graph
- Mismatch Between OSM and Usual GIS Model
- Level of Detail
- Limited Coordinate Precision

- Coordinate Wraparound at Anti-Meridian
- Relations and Extracts
- Versioning in Extracts
- Stability of Ids
- File Types
- Tags
- Robustness and Security
- More Issues

Issues

I talked and listened to lots of people.

Of course still somewhat subjective.

More issues will come up in further discussions.

No Solutions

Some issues are out of scope.

For some we have no good solutions.

Some would change OSM fundamentally.

Solutions

Stick to issues where we have
good ideas for solutions.

Still discuss some others so we can
improve our understanding of them.

Issues not solved

Object Identity

Stable Ids

Requested very often.

Would need large changes to our thinking.
Fundamental difference to Wikidata model.

For details see my talk at SotM 2018.
→ Many cases can be solved with tags.

Tags

Structure for Tags

start_date
check_date, survey:date
opening_hours
generator:output:electricity
turn:lanes:both_ways

Structure for Tags

Complex issue.

The simplicity of the OSM data model
hides a lot of complexity.

We need to think about this a lot more.

See talk by Sarah Hoffmann and Richard
Fairhurst after the coffee break in this room.

Structure

Tags

Nodes, Ways, Relations, Ids, Coordinates, ...

Structure

Tags

Flexible

Nodes, Ways, Relations, Ids, Coordinates, ...
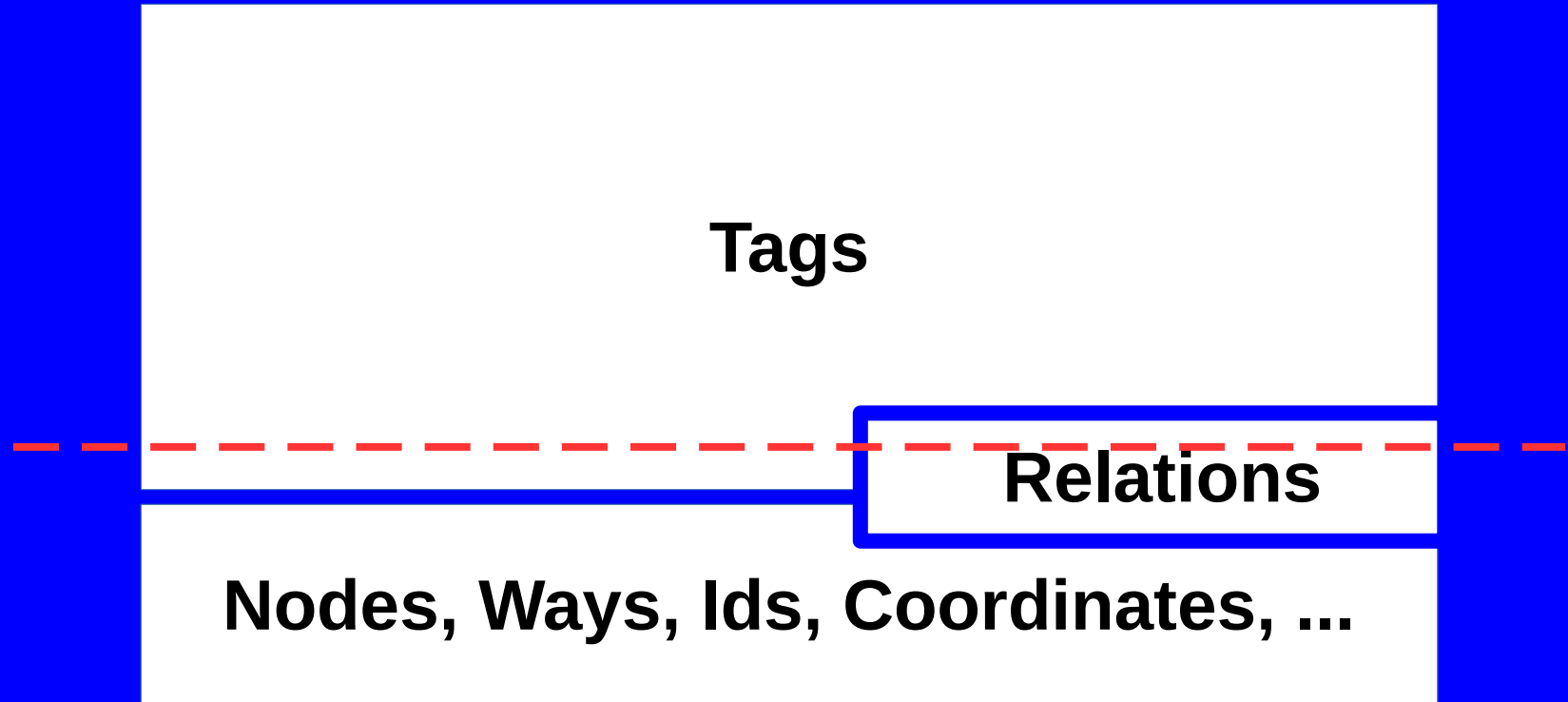
Structured

Structure

Flexible

Tags

Nodes, Ways, Relations, Ids, Coordinates, ...

Structured

Structure

Tags

Relations

Nodes, Ways, Ids, Coordinates, ...
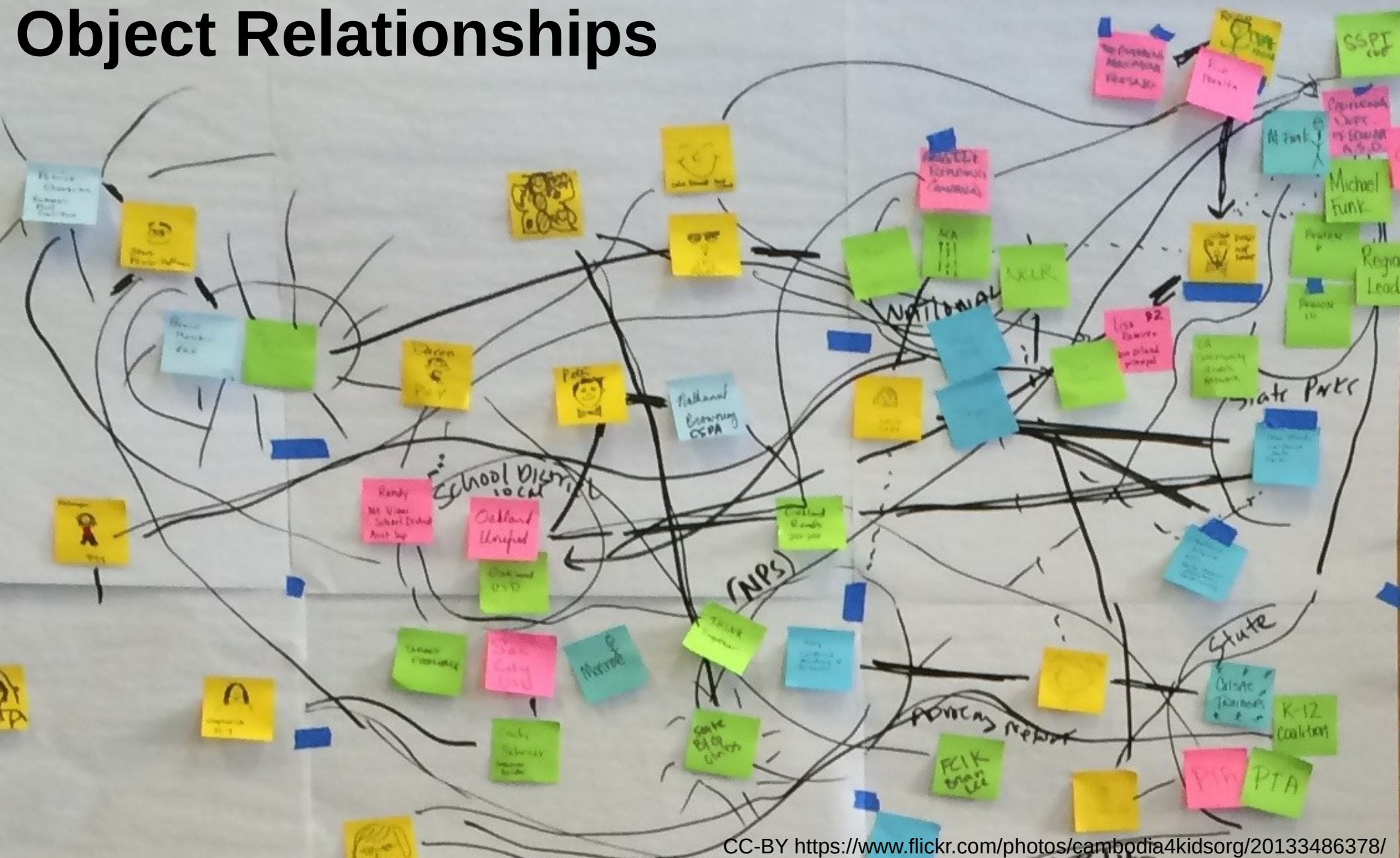
Issues (sort of) solved

Object Relationships

Relationships between objects

Objects are related by location of course.

But they are also related by id.

nodes ↔ ways
members ↔ relations

Extremely expensive to process.

Remove untagged nodes

Removes >95% of nodes.

Ways will have their own location.

Massively reduces memory/disk need when working with OSM data.

Processing of changes becomes much simpler.

Area Datatype

New area data type for small areas.

Paper suggests three options.

Partial solution for large multipolygons.

Large change that nobody will see

These are large, but internal changes.

Normal users will see not much difference.

If anything, things will work more like people expect, especially around changes.

Oh no, the world is going to end!

Without explicit connections from nodes to ways everything will be a huge mess.

It might be, but not because of this

Editors can handle this.

Editors download a complete area anyway.

Snapping to other objects still possible.

Next Steps

EWG has to decide on all of this...

Next Steps

Phase I: Community Consultation and Decision
Phase II: Detailed Planning and Preparation
Phase III: Implementation and Changeover
Phase IV: Follow-Up

Phase I

Talk to all stakeholders.

Make sure we have a solid plan.

Decide on the scope.

Phase II

Detailed planning.

Define APIs, formats, etc.

Start implementing tools, libraries...

Phase III

Finish implementation.

Help community changing their software and systems.

Changeover.

Phase IV

Follow-up.

Solve any issues that came up.

This is a Huge Project!

Expect this to take
two or three years!

Funding Needed

Project Management
Technical Lead
Community Relations
Servers
Supporting Projects

Very Next Step

BoF Session
Tomorrow, Saturday 11:30
Room 101

Next Step

Evolution not Revolution

There will be future changes,
don't try to do everything at once.

Thank you!

https://github.com/osmlab/osm-data-model

blog.jochentopf.com

Jochen Topf
jochentopf.com

jochen@topf.org