



Cleaning Up the **plan.at** Import

The Next Step

Wolfgang Schreiter
(wolfgang@schreiter.at)

A Data Import with After-Effects...

- In 2008/2009, publishing house Compass donated its aging map database of Austria to osm.
- The import provided much-needed data, but also resulted in massive cleanup efforts which are still in progress.
- Among others, more than 360.000 superfluous tags are causing confusion and needless editing.
- It is therefore proposed to delete these tags semi-automatically within the next few weeks.

The Size of the Problem

Tags	Remarks	Volume ca.
plan_at:acad_id=*, :acad_id=*, acad_id=*	Reference to Compass database no longer needed or usable	61.723 1.188
at:maxspeed=Ortsgebiet	Used regardless of city limits, speed limits or highway type	22.504
uploaded_by=Wolfgang W. [...] downloaded_by=Wolfgang W. [...]	Additionally, 600 occurrences from other mappers	89.331 94
note=imported from plan.at [...]	„No copyright“ notice; duplicates source-tag	82.418
addr:*=* on highway	Highways don't have addresses; imitated by other mappers as well	> 99.981
Total		> 357.239

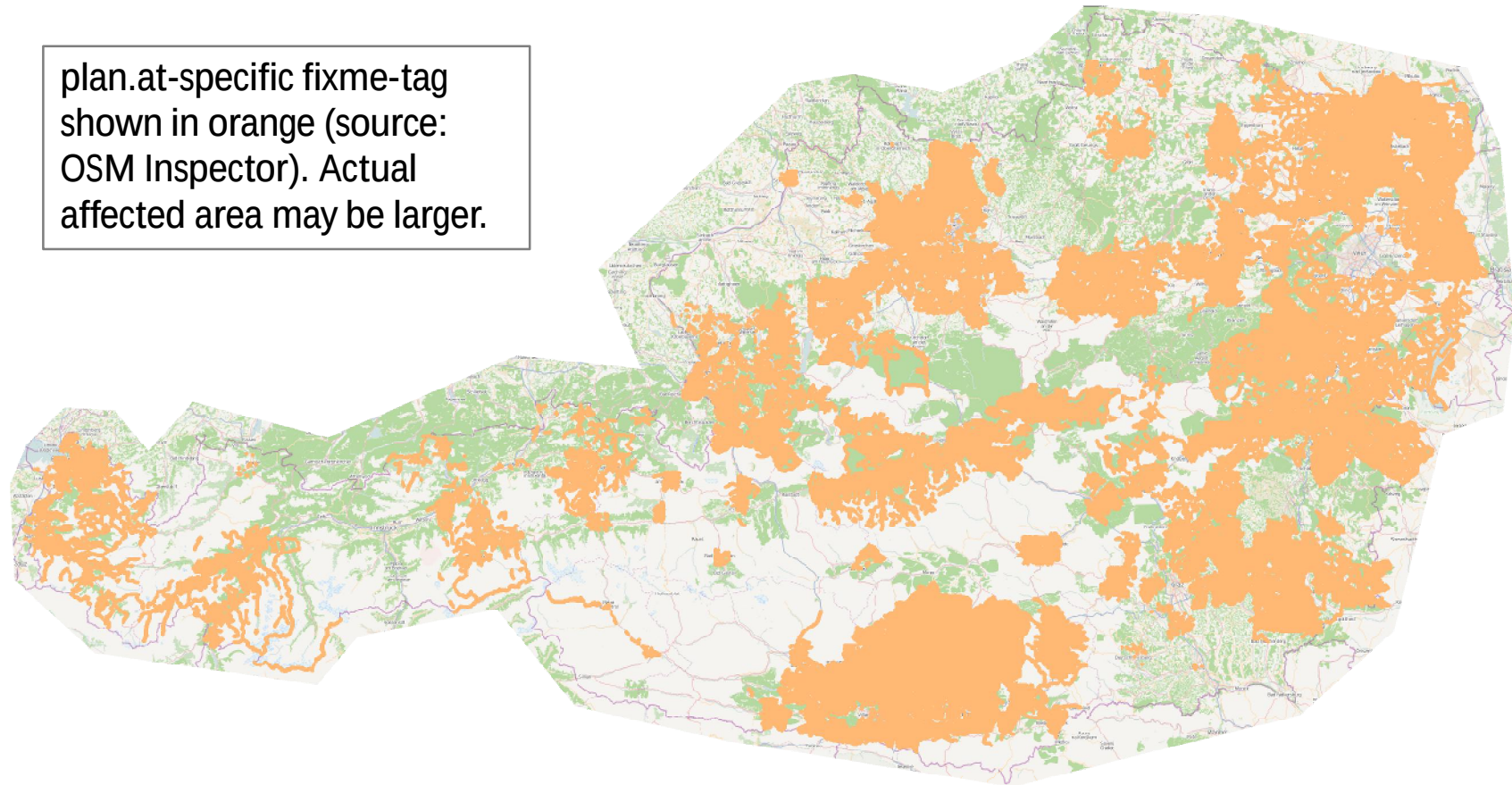
Data based on Austria data excerpt from 15.07.2011 by Geofabrik

Things to Watch Out For

- Limit change to Austria
- Ensure roll-back is possible shortly after the change
- Preserve reference to plan.at
(add plan.at to source-tag where missing)
- Keep addr:street on highway if name-tag is not set
(161 occurrences; rename tag e.g. to fixme:name)

Estimate of Most Affected Areas

plan.at-specific fixme-tag
shown in orange (source:
OSM Inspector). Actual
affected area may be larger.



Proposed Approach

- Announce change on Austrian mailing list
- Query objects to be changed from latest database excerpt and analyse for anomalies
- Process only reasonable sized batches of objects at a time
- Perform actual change either by means of a proven bot (e.g. xybot) or manually for special cases
- Verify change in database

Anything else we should think about?