gesis Leibniz Institute for the Social Sciences

Georeferenced Survey Data at the GESIS Data Archive

Stefan Müller | Wolfgang Zenk-Moeltgen | Stefan Schweers
GESIS – Leibniz Institute for the Social Sciences
Data Archive for the Social Sciences
{stefan.mueller|wolfgang.zenk-moeltgen|stefan.schweers}@gesis.org

EDDI15 - 7th Annual European DDI User Conference, December 2-3, 2015

Member of the Leibniz Association

EDDI

Content

1. Background
2. DDI and Georeferenced Survey Data
3. The Future? Archiving 'real' spatial data
4. Outlook


ONE DOES NOT
SIMPLY
GEOREFERENCE SURVEY DATA

gesis

High efforts

Conceptual

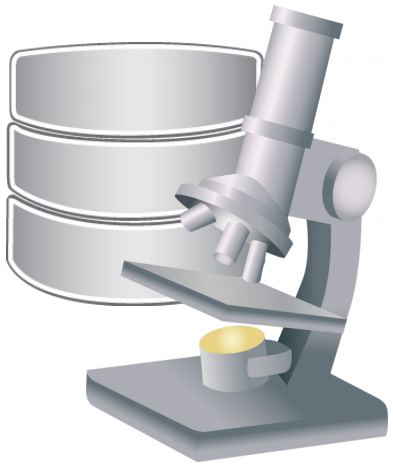
gesis

High efforts

Conceptual

gesis

High efforts



Conceptual

Technical

gesis

High efforts

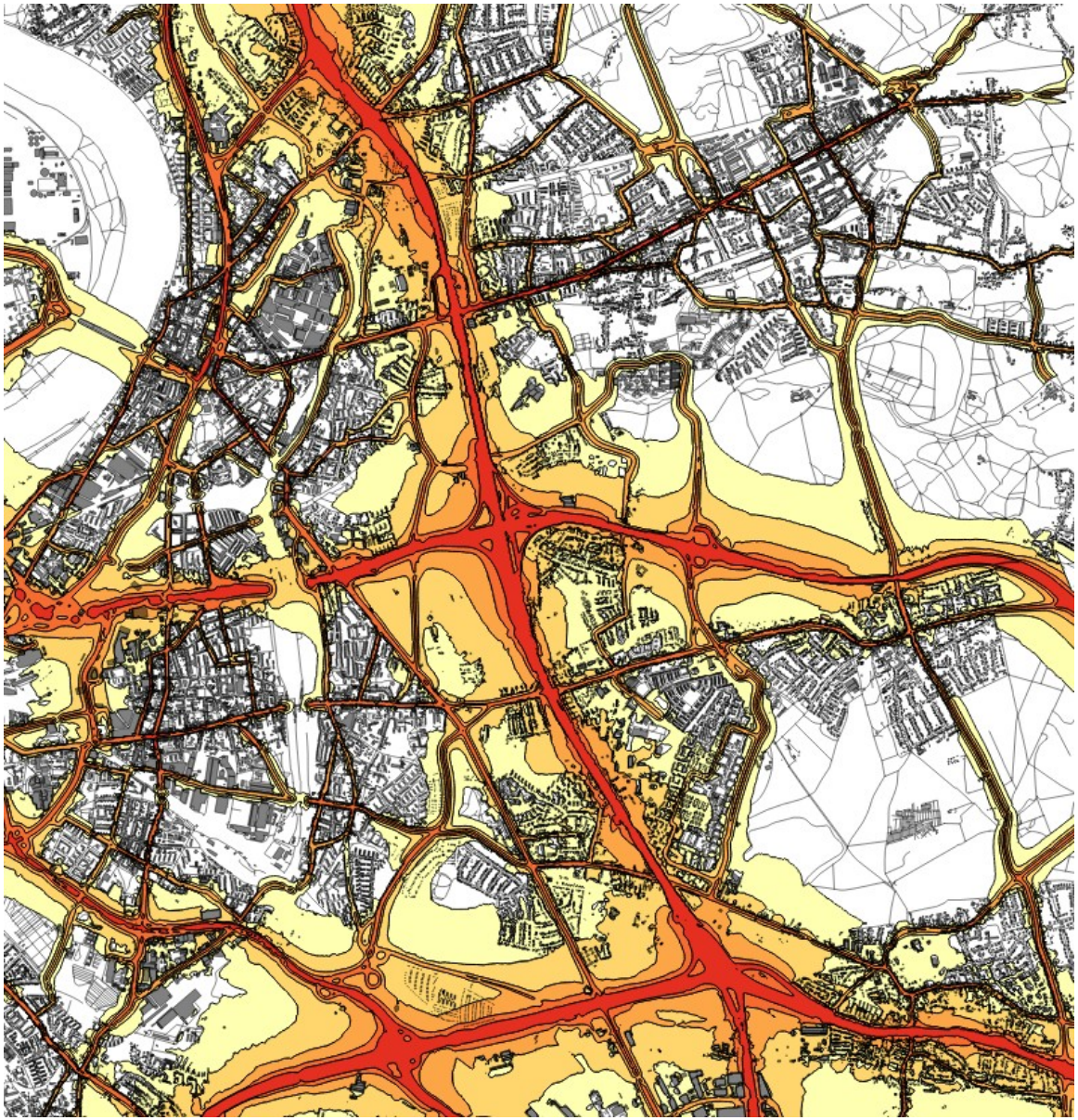
Conceptual

Technical

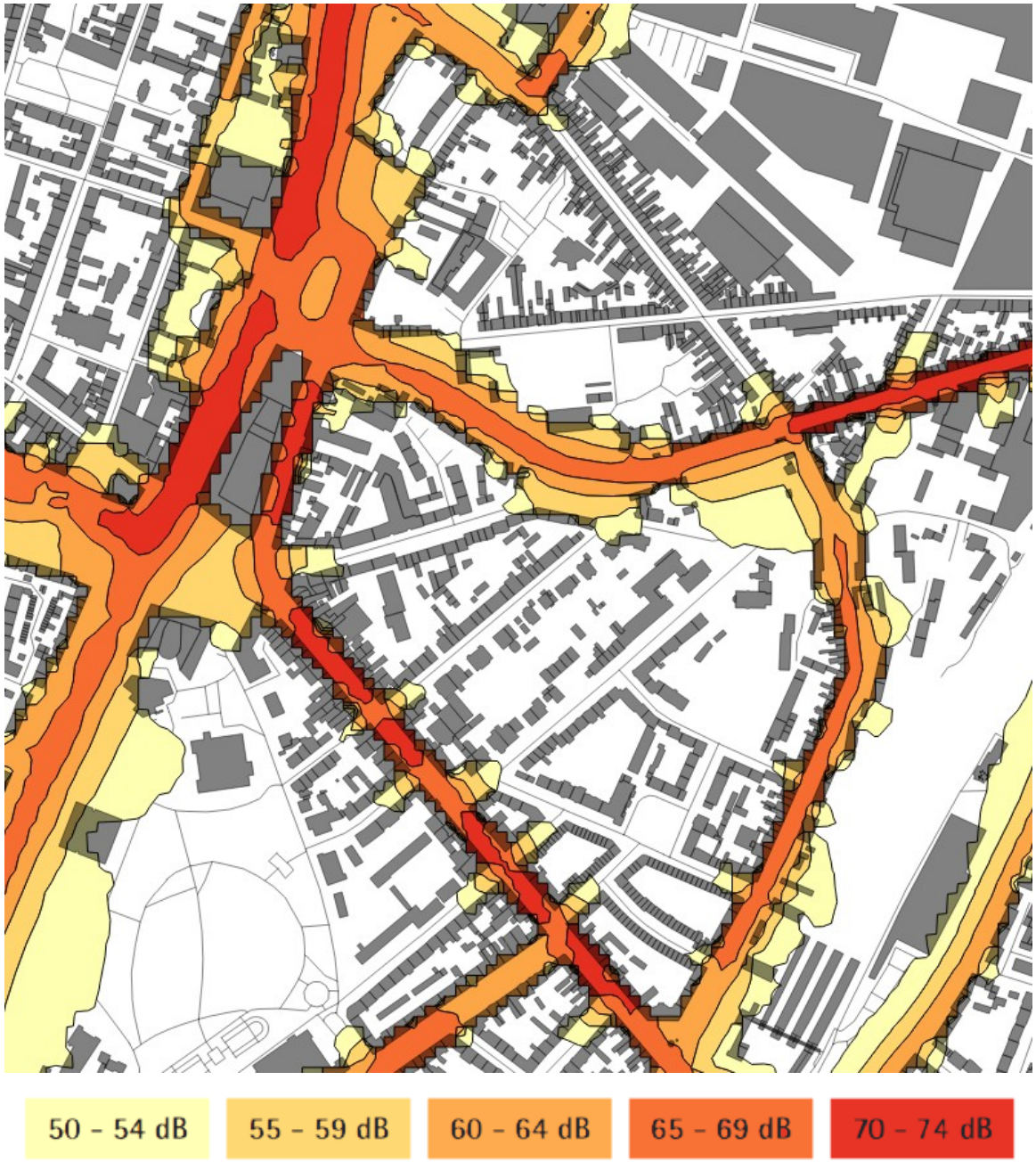
Documentary

The project: GeorefUm

- Geocoding and georeferencing
- Building a database of available spatial data
- Disseminating in a secure environment
- Consult and guide

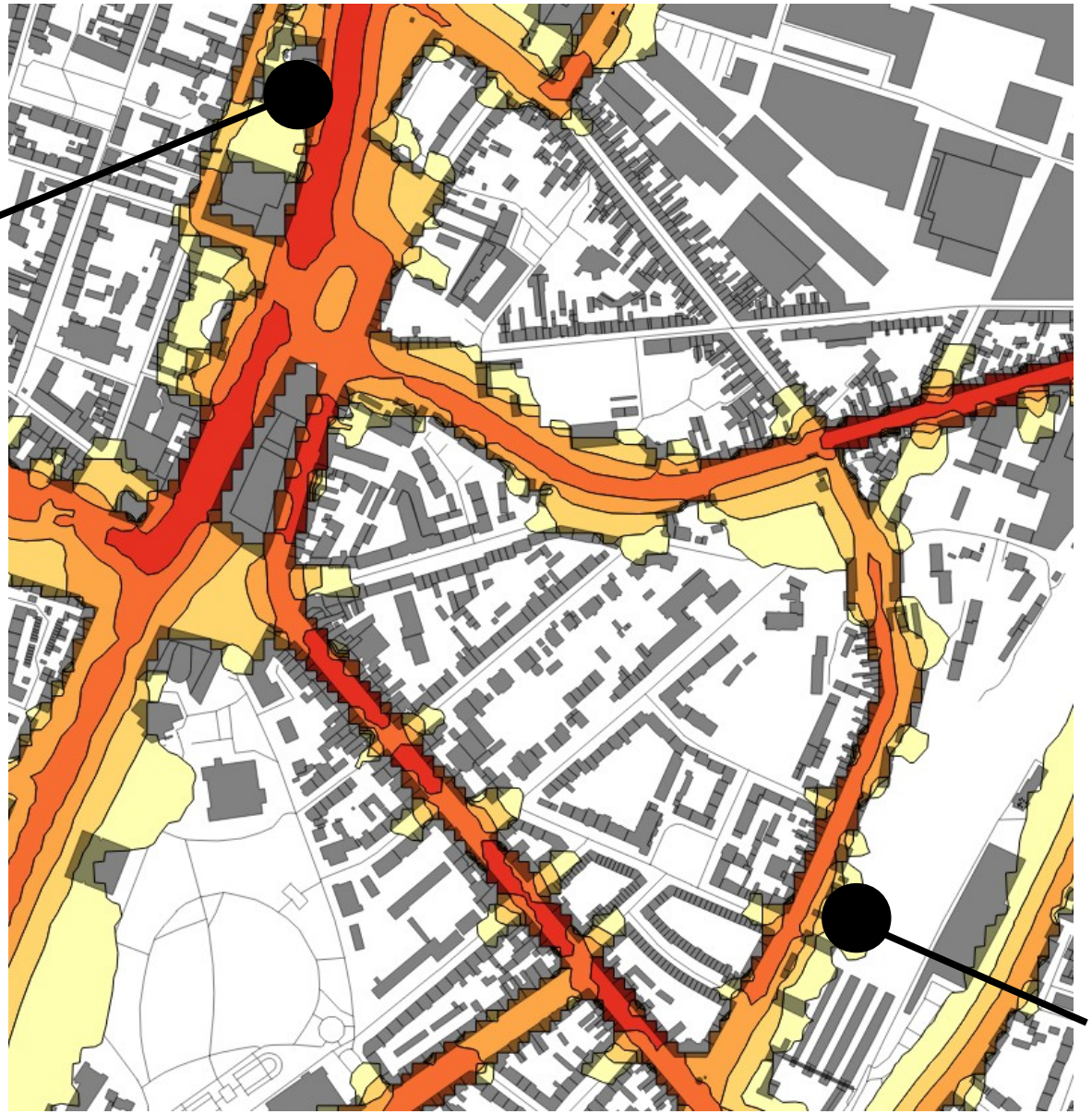


Road traffic noise in Cologne, Source: EIONET Central Data Repository (CDR) and OpenStreetMap



Road traffic noise in Cologne, Source: EIONET Central Data Repository (CDR) and OpenStreetMap

60 – 64 dB

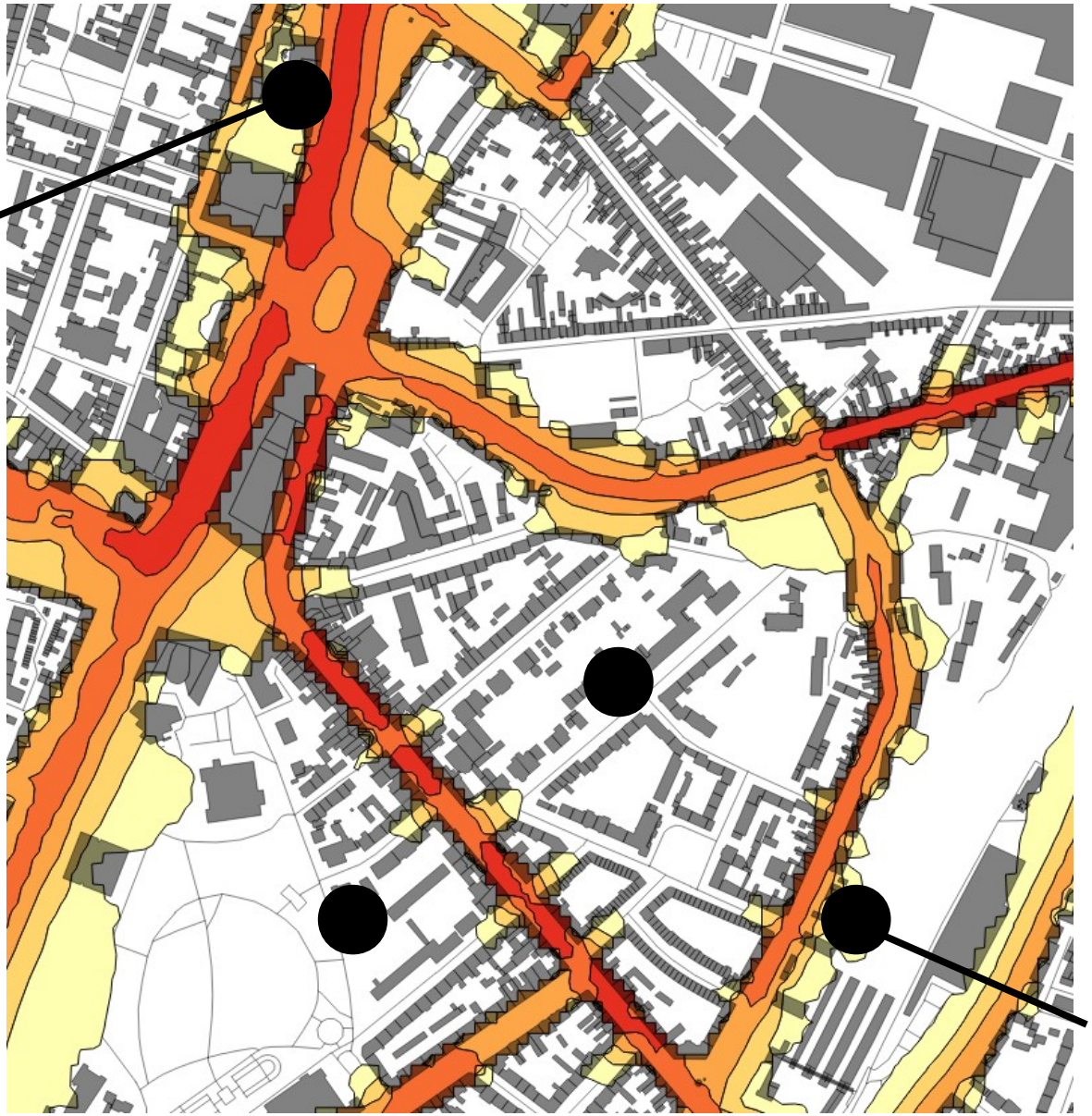


50 – 54 dB

50 - 54 dB | 55 - 59 dB | 60 - 64 dB | 65 - 69 dB | 70 - 74 dB

Road traffic noise in Cologne, Source: EIONET Central Data Repository (CDR) and OpenStreetMap

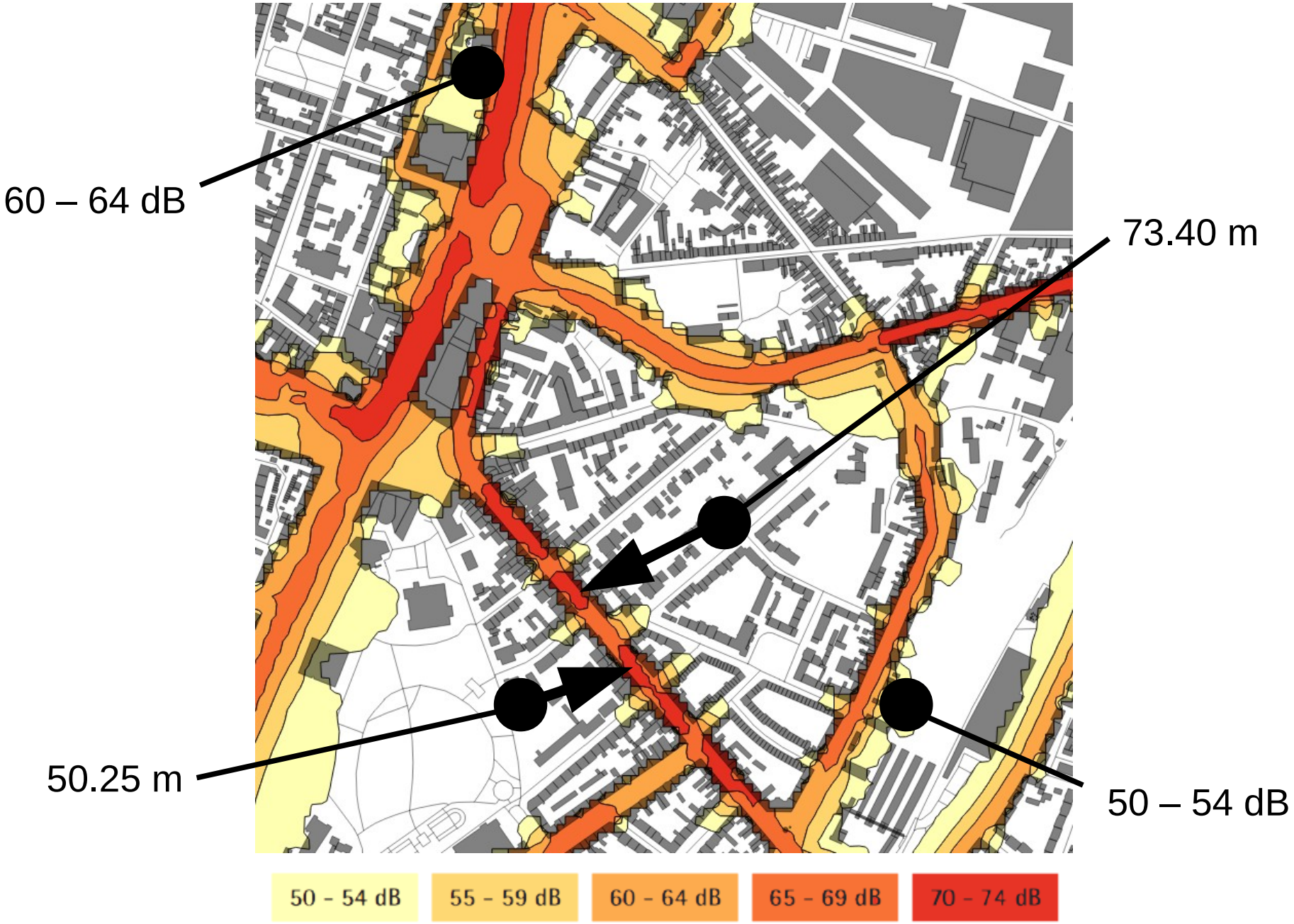
60 – 64 dB



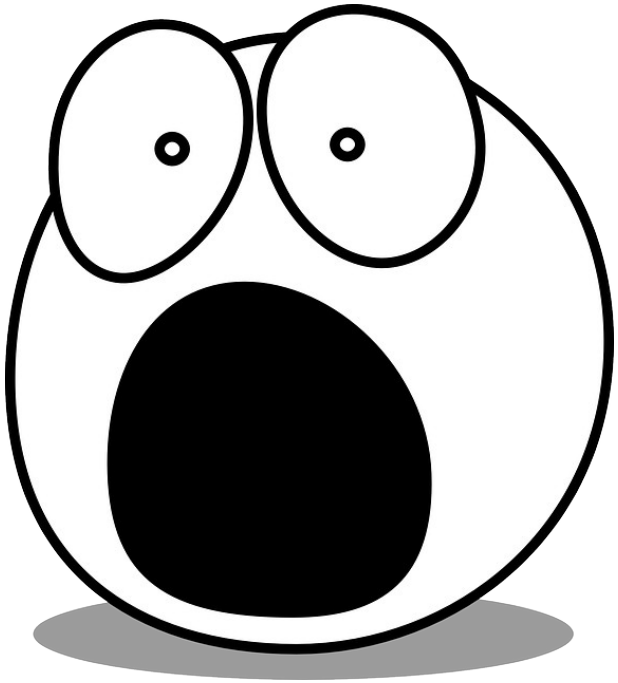
50 – 54 dB

50 - 54 dB | 55 - 59 dB | 60 - 64 dB | 65 - 69 dB | 70 - 74 dB

Road traffic noise in Cologne, Source: EIONET Central Data Repository (CDR) and OpenStreetMap



Road traffic noise in Cologne, Source: EIONET Central Data Repository (CDR) and OpenStreetMap



Source: pixabay.com

New data types, formats, contents etc.

...new documentary demands!

gesis

Actually, not that different data

ID	...	ROAD_DEN	ROAD_N	DIST_ROAD_DEN	DIST_ROAD_N
1	...	55	50	24.56	0
2	...	75	60	0	0
...	...	...	...	...	...
3146	...	0	0	6.23	10.76

Still, we need new metadata

Different standards or frameworks:

- ISO 19118, 19115, 3166 and INSPIRE

Different purposes:

- cataloging
- usage
- structure and meaning
- preservation

Source: pixabay.com

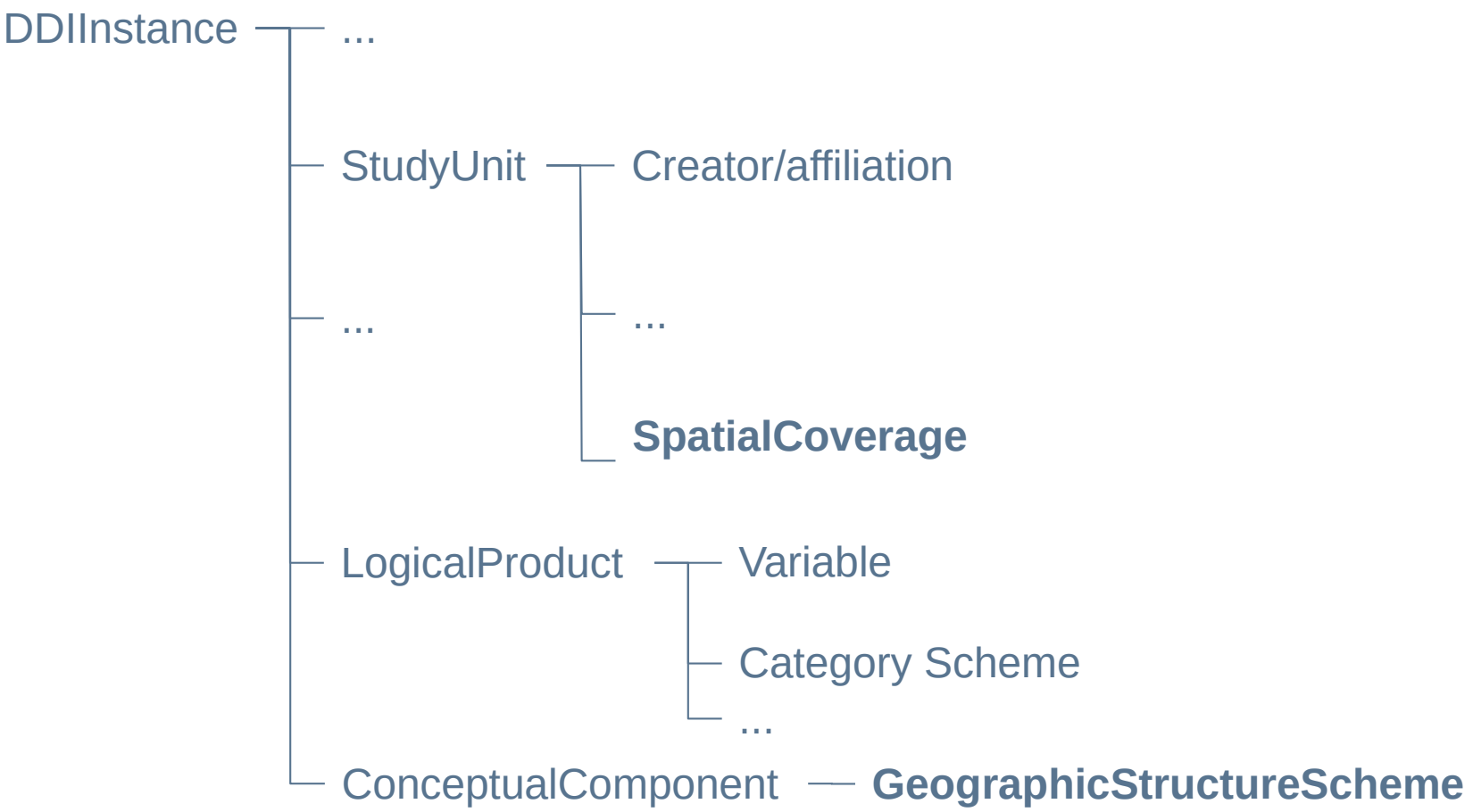
gesis

ISO 19115 wishlist

- Abstract
- CitedResponsibleParty
- Extent
  - spatial
  - temporal
- SpatialRepresentationType

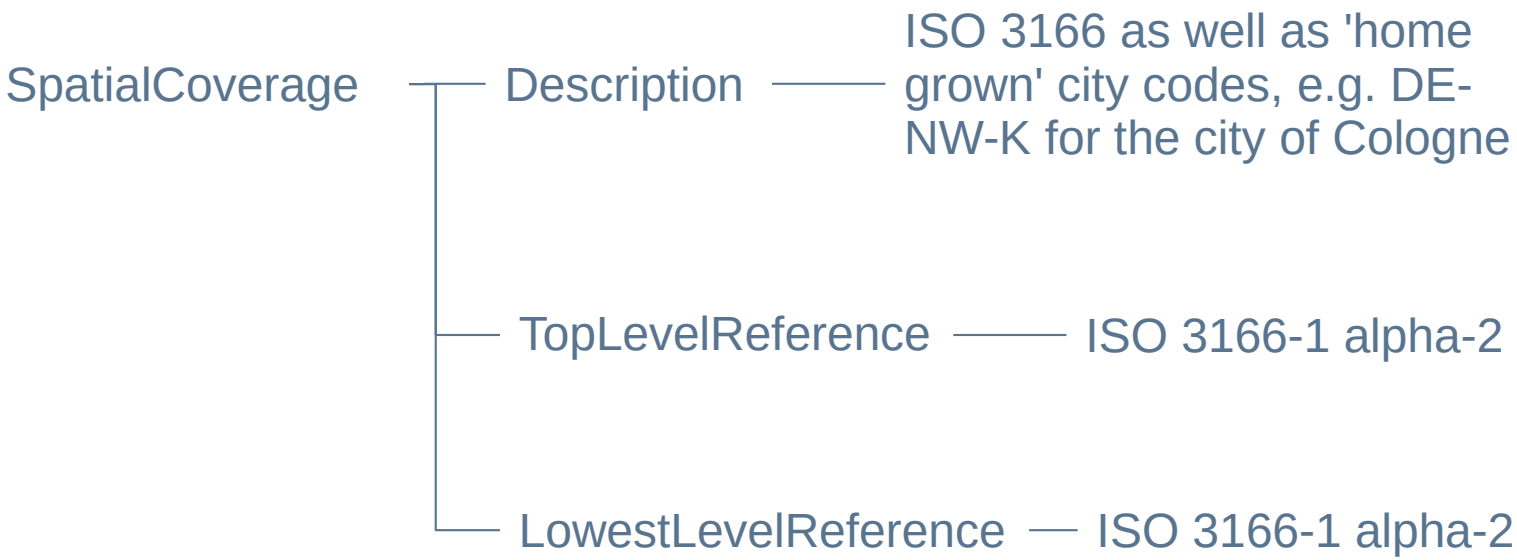
gesis

GESIS DDI Lifecycle Snippet



gesis

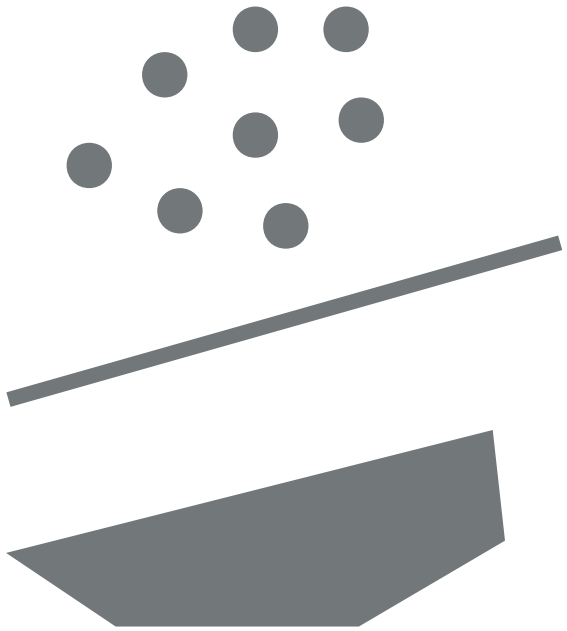
Study Level (as is)



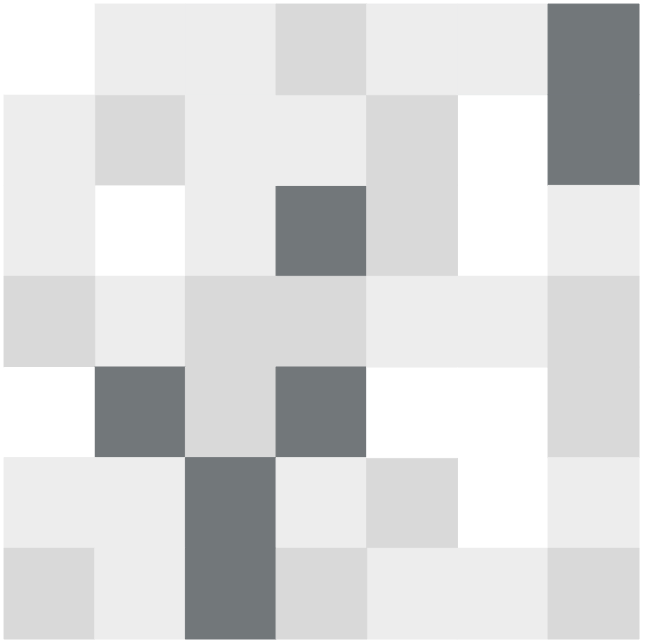
gesis

Study Level (as could be)

Structure of spatial data



Vector data



Raster data

gesis

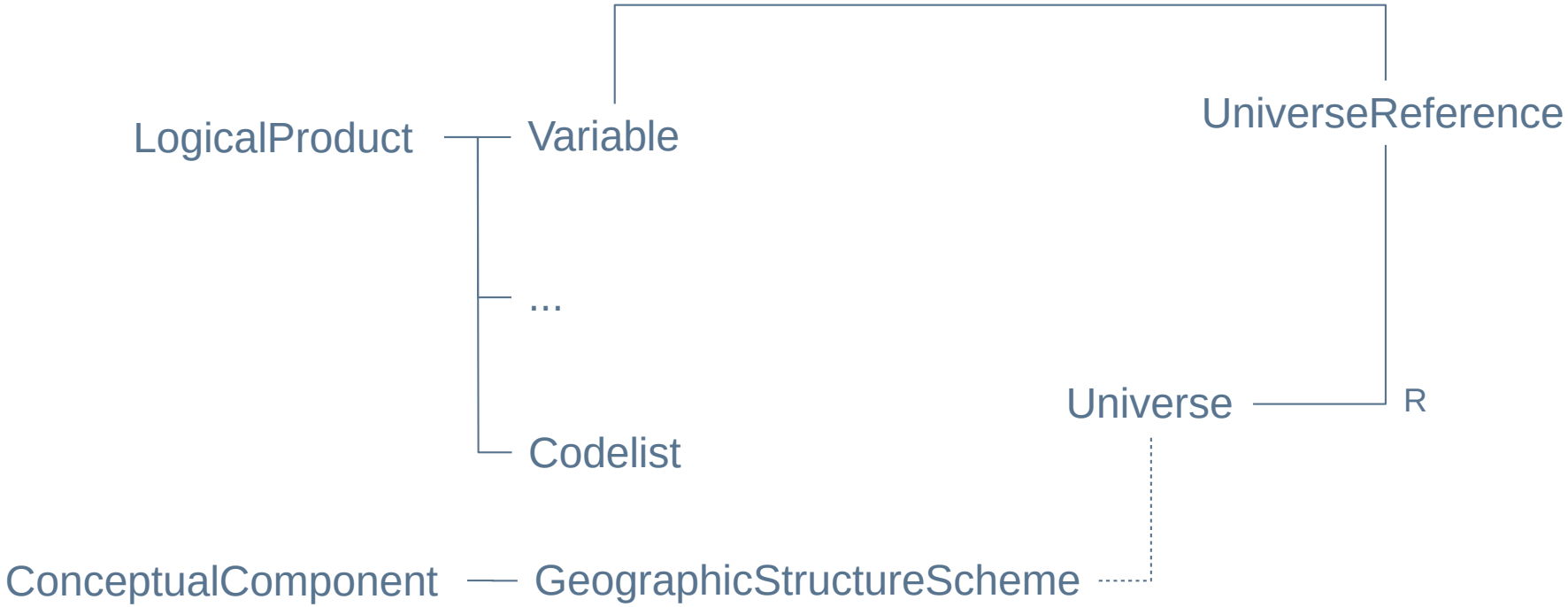
Structure: raster data

GeographicStructure

Hello. I am raster data, based on 1km INSPIRE compliant grids. This is important as I am not serving straight to the point information. Use me when georeferenced variables were produced with data from the same grid structure.

gesis

Variable Level ('whole' story I)



gesis

Structure as Universe?

Universe

Hello. I am raster data, based on 1km INSPIRE compliant grids. This is important as I am not serving straight to the point information. Use me when georeferenced variables were produced with data from the same grid structure.

gesis

A shortened example

```
<l:Variable>
    ...
    <l:VariableName>
        <r:String>road_lden</r:String>
    </l:VariableName>
    <r:Label>
        <r:Content>Road Traffic Noise (DEN) in decibels</r:Content>
    </r:Label>
    </r:Description>
        <r:Content>Road traffic noise that was measured at the
respondent's dwelling in context of EG (2002). </r:Content>
    </r:Description>
    <r:UniverseReference>
        <r:ID>INSPIRE_grid_1km</r:ID>
    </r:UniverseReference>
    ...
</l:Variable>
```

'Whole' story II

- Create one study unit for each geographic structure
  - e.g. one for every 1km grid structure

SU I:

'Ordinary' survey variables

SU II:

Variables extracted from rasters

SU III:

Variables extracted from polygons

What's left?

- Actual implementation
- Final mapping to ISO 19115
- Using DDI-RDF and controlled vocabulary

- What about 'real' spatial data?

gesis

'Real' spatial data: base data

Cologne's 'Veedel', Source: City of Cologne

Different understandings of these data

- Administrative borders (e.g. Switzerland)
- Shapes of cadasters, buildings, roads as well as administrative borders (e.g. Germany)

Archiving 'real' spatial data

Great, because:

- Important to document data collection process or to visualize results
- Makes data linking easier

However:

- Pitfall of redundant archiving
- DDI uncommonly used format

Outlook

- Georeferencing survey data is costly
  - This is why we do it
- New documentary demands
  - e.g. ISO 19115 and INSPIRE compliance
  - Indeed, we use DDI for that
    - Variable level is tricky
- 'Real' spatial data
  - Future is uncertain, depends on demand

Thanks to:

Wolfgang Zenk-Möltgen and Stefan Schweers as well as to Julia Klinger, Katharina Kinder-Kurlanda and Pascal Siegers