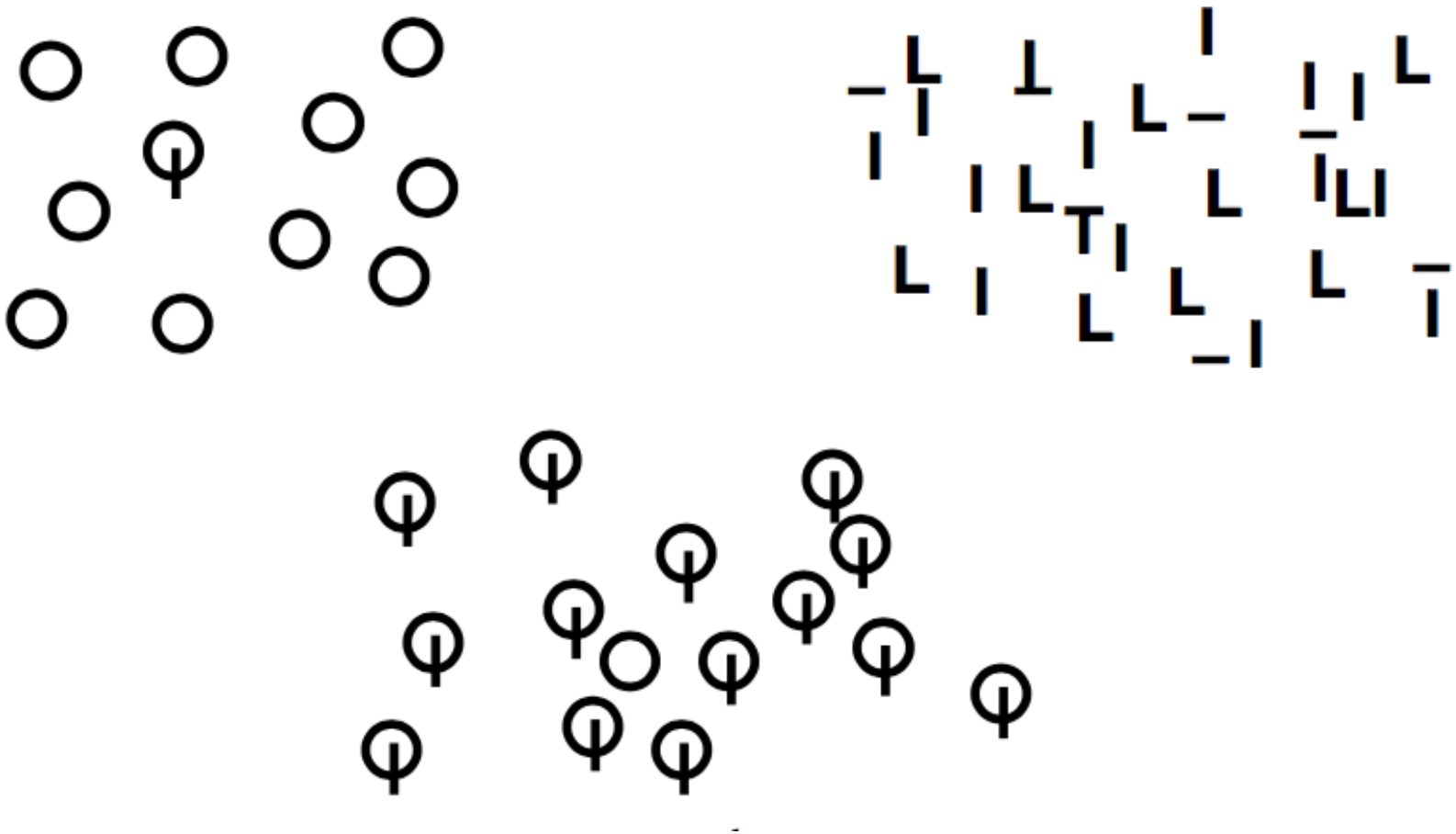
Aspects of visual(ized) dialectology in the DMW-Project

Kai-Uwe Carstensen
University of Siegen, Germany
www.kai-uwe-carstensen.de

Tagung „Visualisierungsprozesse in den Humanities:
Linguistische Perspektiven auf Prägungen, Praktiken, Positionen“
Zürich, 19.7.2017

Motivation: the „Pop-out“-Phenomenon in Vision

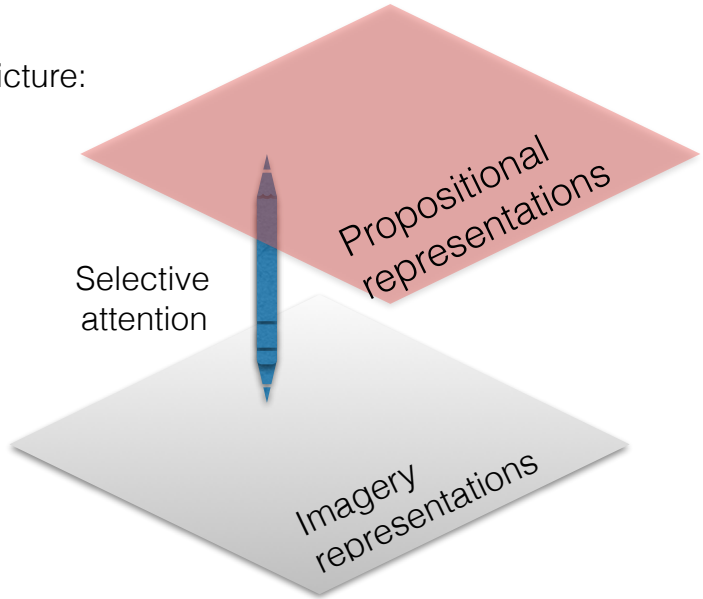


Carstensen (2001): *Sprache, Raum und* Aufmerksamkeit (Tübingen: Niemeyer), p.101.

The role of attention

- Until the 1970s: Hype of propositional representations
- From 1970s on: Hype of imagery (Paivio, Lakoff etc.)
- ~1980s: Rise of (selective) attention research
  - selectivity of cognitive processing
  - link between imagery and propositional concepts
  - both bottom-up and top-down control

simplified picture:

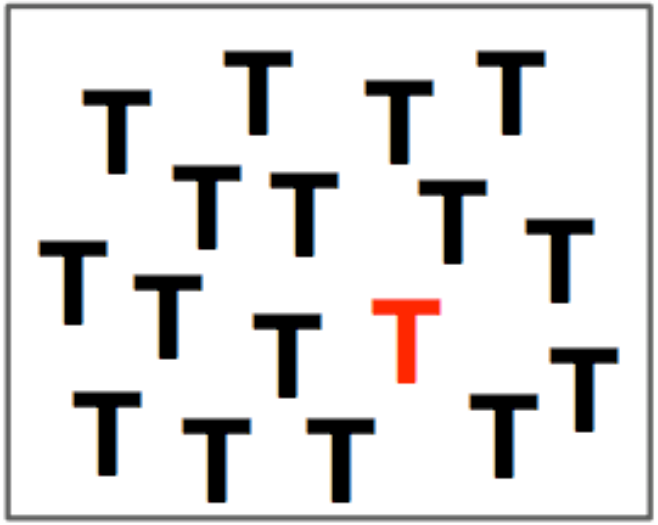


Der Begriff der Aufmerksamkeit war in den Kognitionswissenschaften des 20. Jahrhunderts aufgrund seiner Unschärfe lange Zeit verpönt, in den letzten Jahren kommt er jedoch in der Hirnforschung, der kognitiven Psychologie, der Linguistik und der Künstlichen Intelligenz, der Kulturwissenschaft und nunmehr auch in der Kunstgeschichte wieder zu Ehren.
(Apel 2005: 165)

Gegenüber den früheren sensualistischen Ansätzen verschiebt sich das Interesse von der Verarbeitung von Sinnesdaten **auf die Analyse der an kulturelle Fähigkeiten gebundenen, in beträchtlichem Maße individuellen Seh- bzw. Wahrnehmungserfahrungen**.
(ibid.: 168)

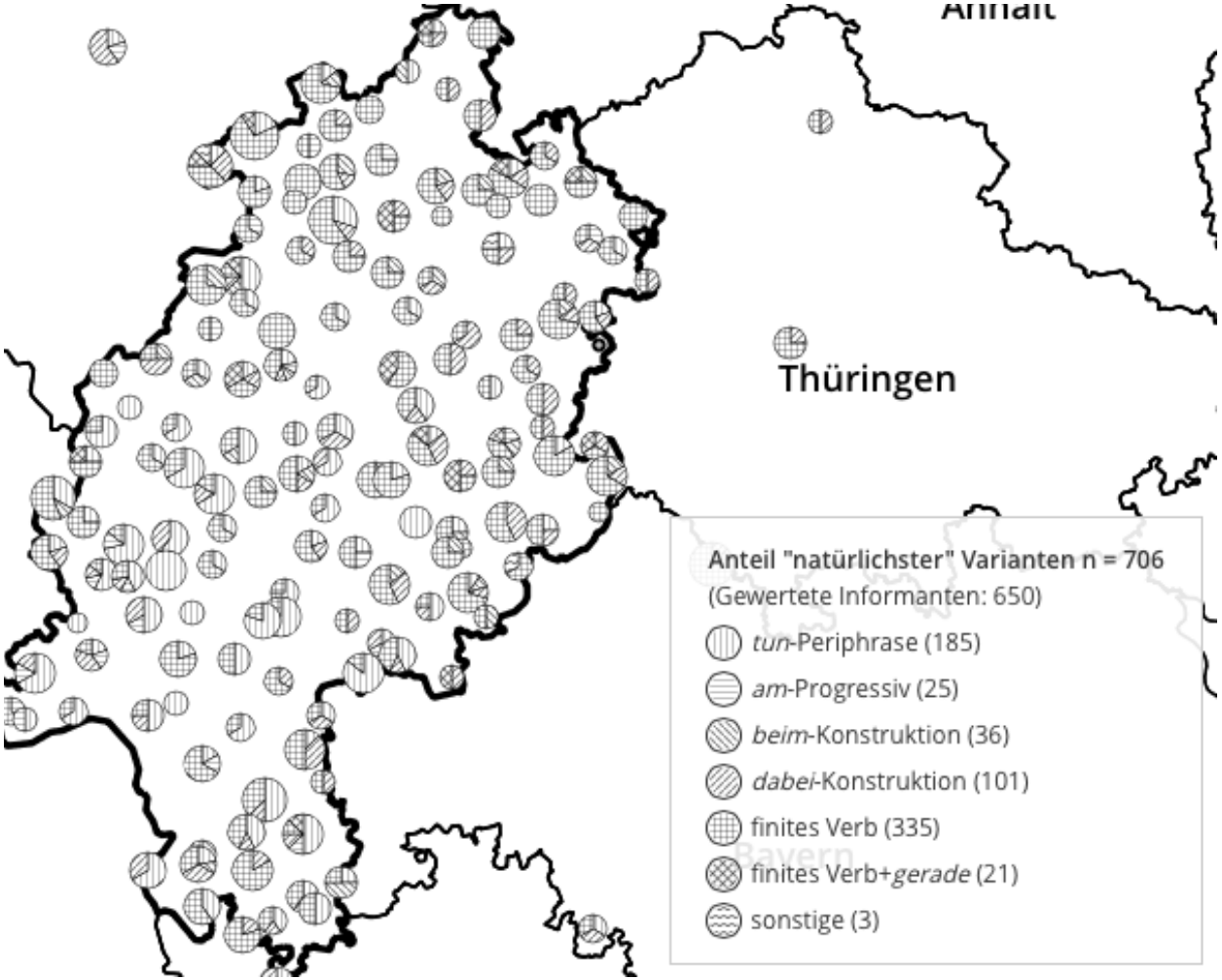
Friedmar Apel (2005): „Attention please! Dugald Stewarts Theorie der Aufmerksamkeit“. In: *Anblick/Augenblick. Ein interdisziplinäres Symposion*. Würzburg: Königshausen & Neumann. 165-172.

Motivation: the „Pop-out“-Phenomenon respected in Visual Analytics



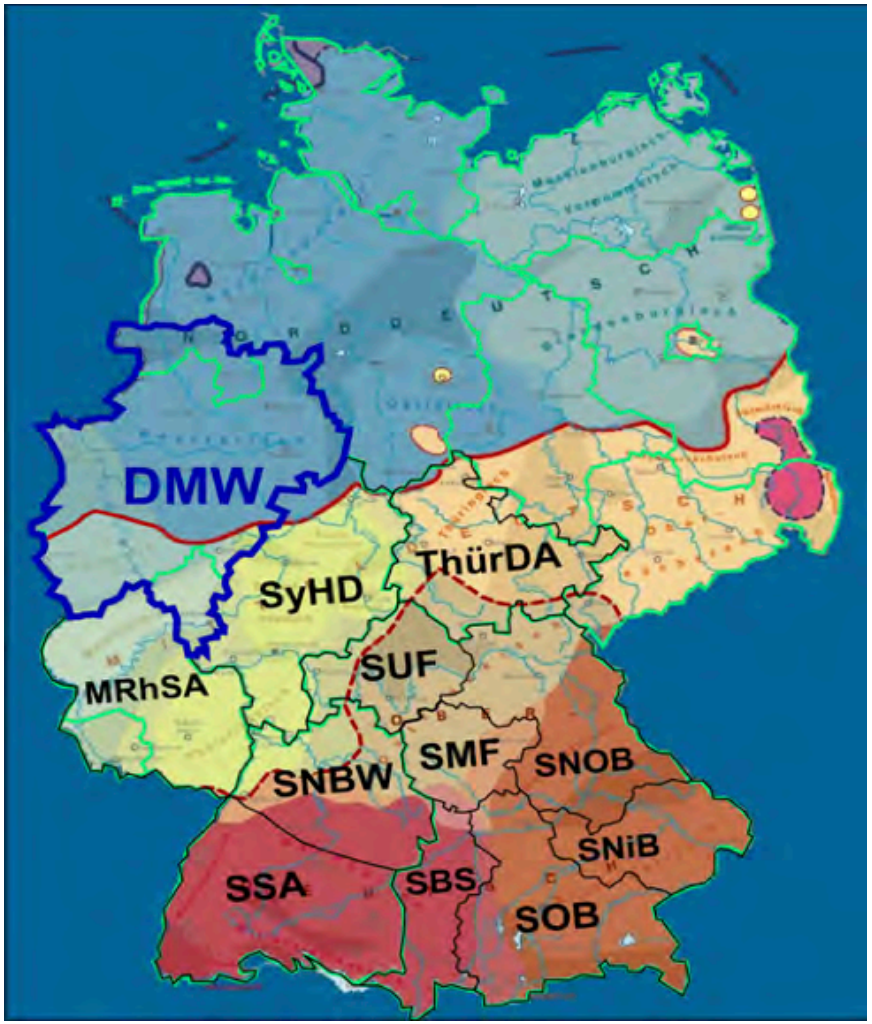
(Keim/Kohlhammer/Ellis/Mansmann (2010):
Mastering the Information Age with Visual Analytics. S. 119)

Motivation: the „Pop-out“-Phenomenon in Dialectology



SyHD

Dialectology research: the DMW-region



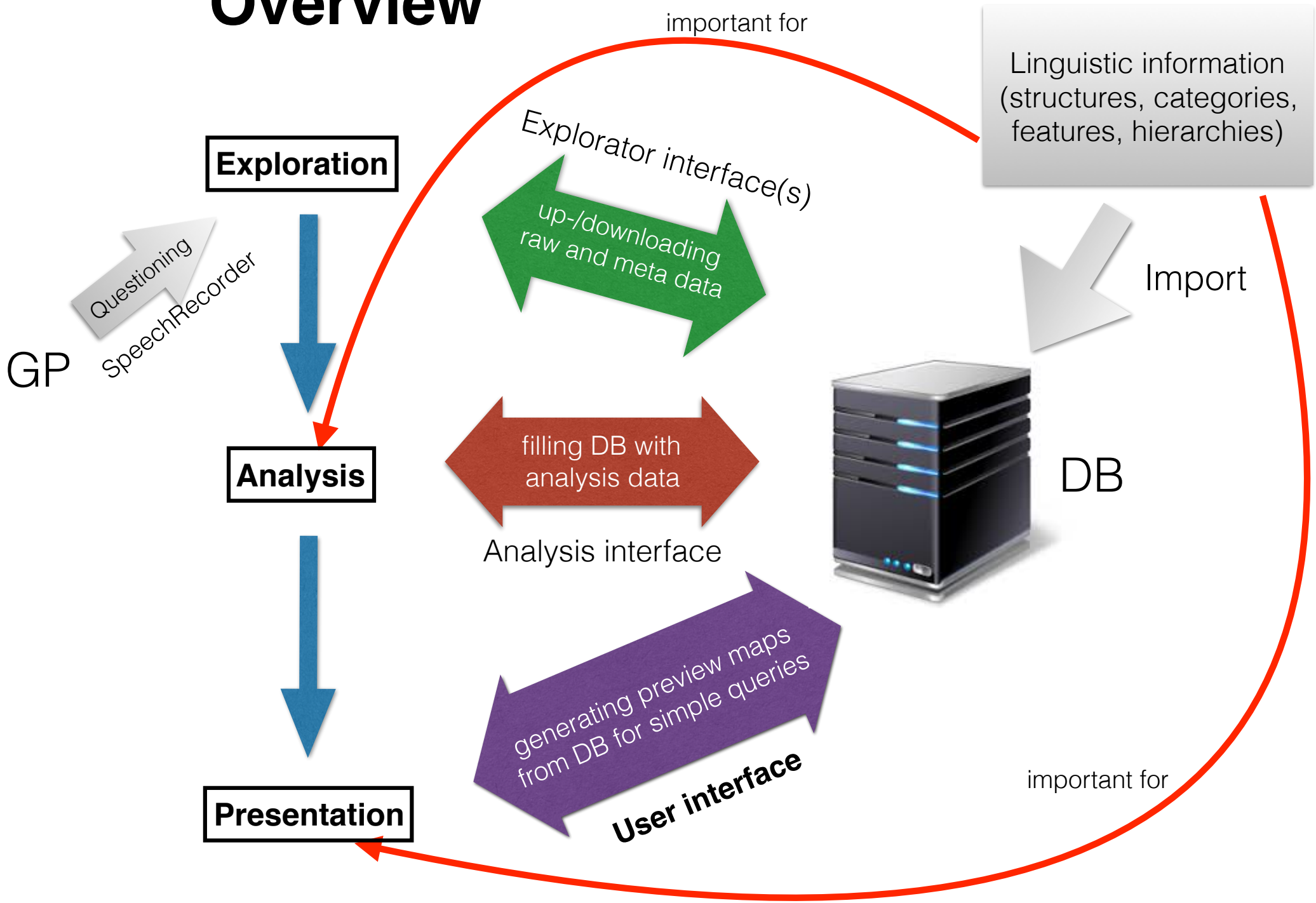
http://dmw-projekt.de

Overview of the DMW-project

- Project partners
  - Univ. of Bonn (Claudia Wich-Reif)
  - Univ. of Münster (Helmut Spiekermann)
  - Univ. of Paderborn (Doris Tophinke)
  - Univ. of Siegen (Petra M. Vogel)
    - Coordination and informatics site
- Goal
  - digital dynamic „speaking“ dialect atlas; role model: SiSAL
  - deliberately non-expert level (at least for now)
- Size of investigation
  - ~800-1000 locations
  - each 2-4 informants
  - Questionnaire
    - 600-800 items (still not fixed) per informant
    - Phon., lex., morph., synt. Items/tasks;
      + „Wenkersätze“ and „mental maps“ (perceptual dialectology)
    - typically more than one phenomenon per item

http://www.mundart.sisal.uni-siegen.de

Overview

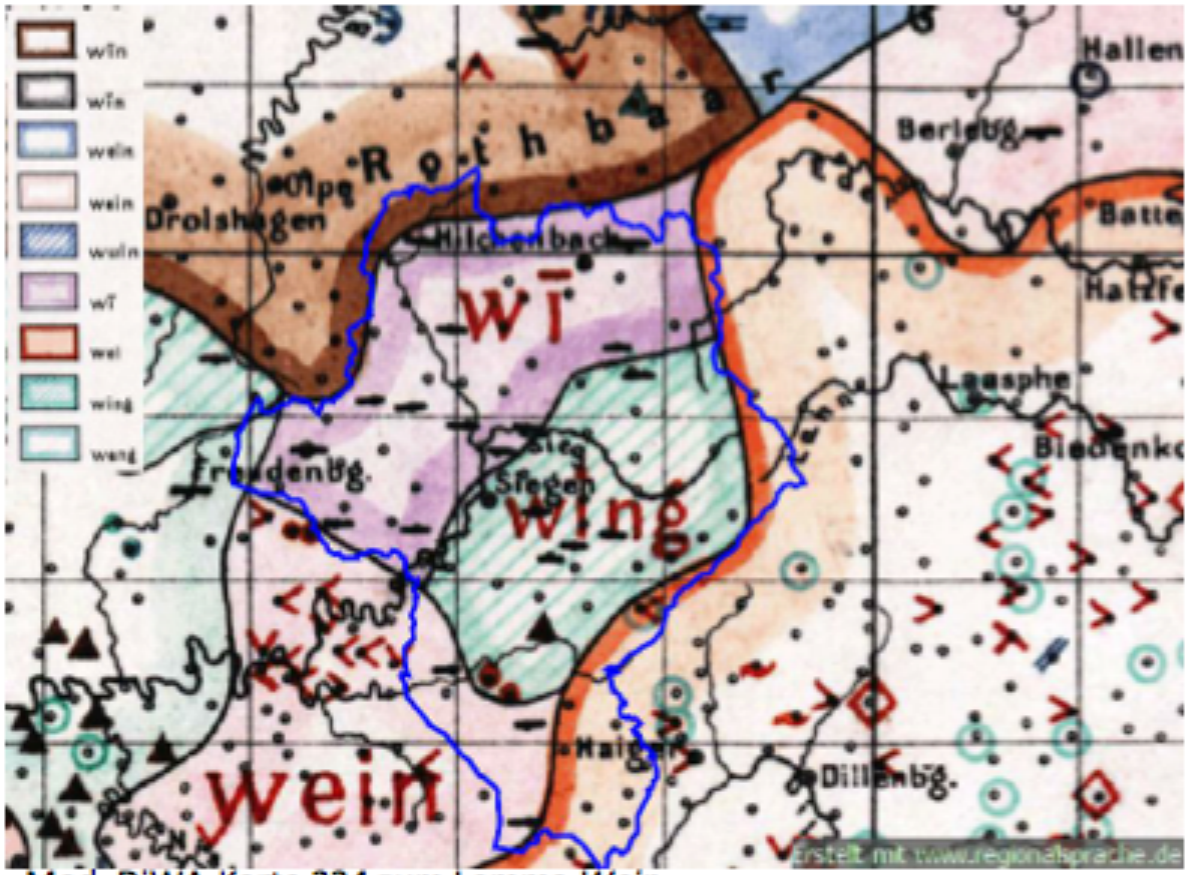

important for
Linguistic information (structures, categories, features, hierarchies)
Exploration
Explorator interface(s)
up-/downloading raw and meta data
Questioning
SpeechRecorder
GP
Import
Analysis
filling DB with analysis data
DB
Analysis interface
generating preview maps from DB for simple queries
User interface
important for
Presentation

Disclaimer

- the following is
  - not (yet) representative for the project
  - preliminary, prospective
- I'm in Ertl's foraging loop for interesting concepts
- comments and feedback are welcome, I'm here to learn!

Classical reference point of our maps

Digitaler Wenkeratlas (DiWA) maps

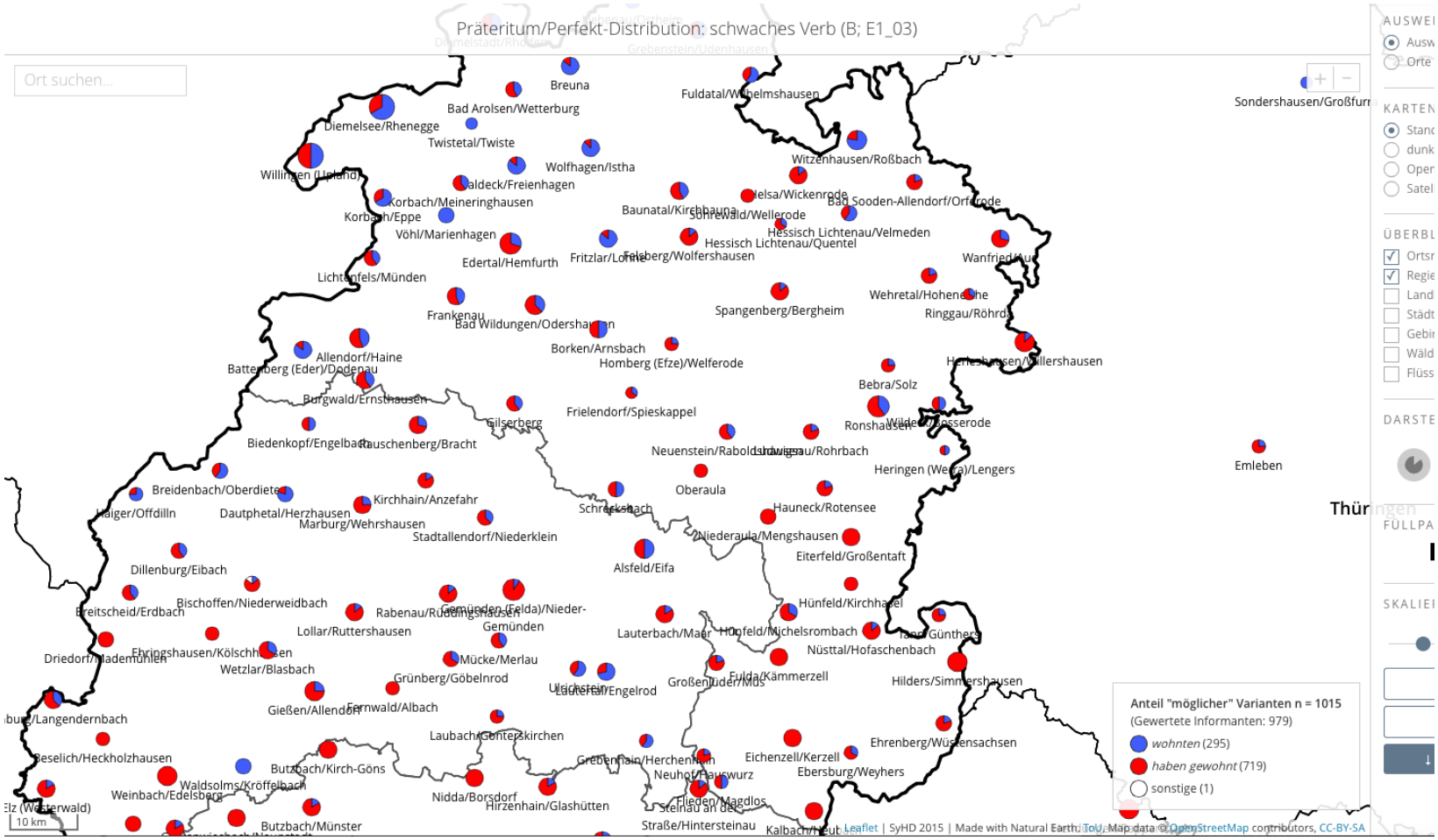


Mod. DiWA-Karte 224 zum Lemma *Wein*

(from: Solau-Riebel Diss)

But we can do without such **information overload** in dynamically generated maps!

What we probably do not want to offer



SyHD

http://www.syhd.info

Dysfunctional presentations of graphical elements (e.g., size of symbols, names)

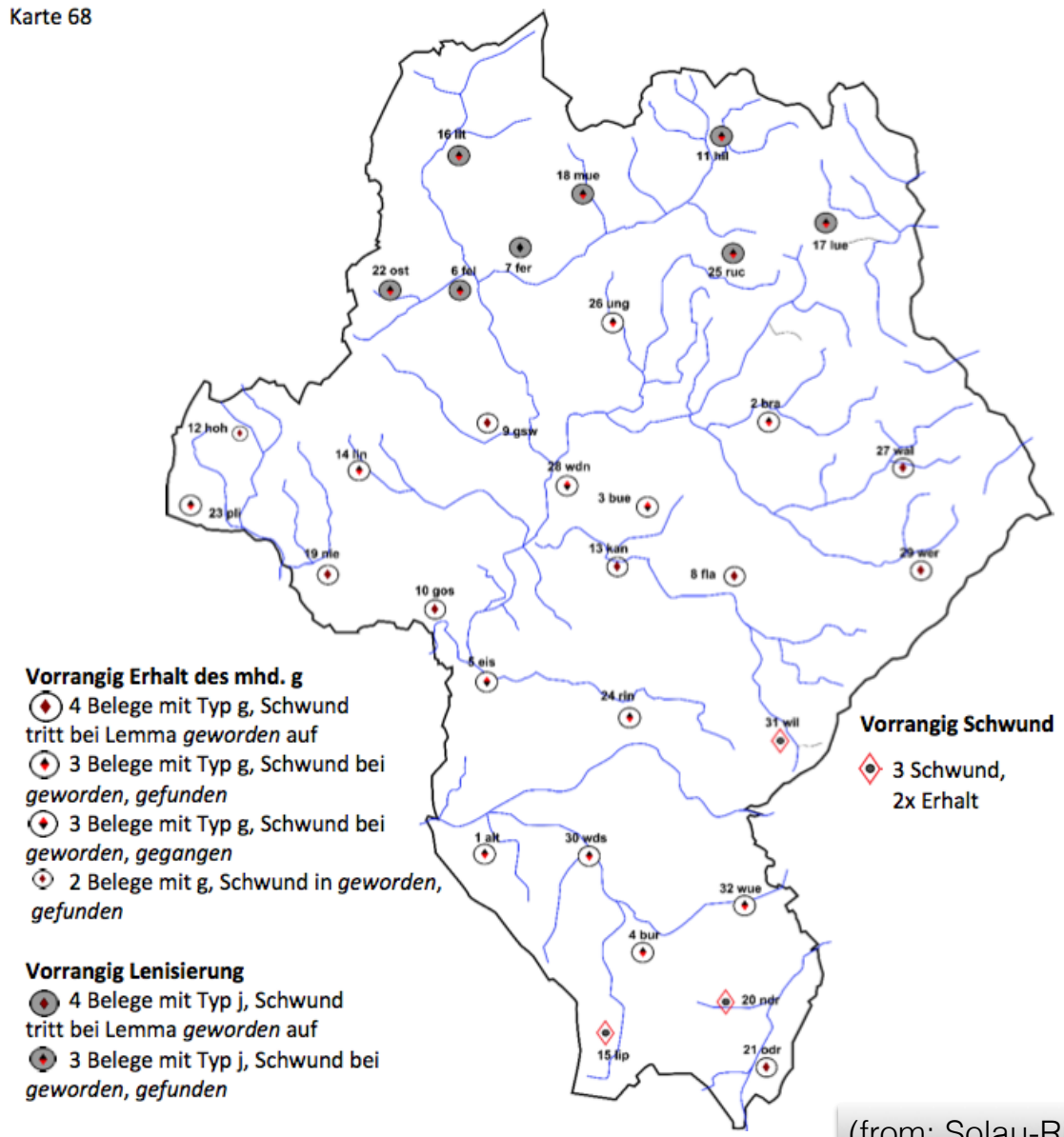
DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

What we probably do not want

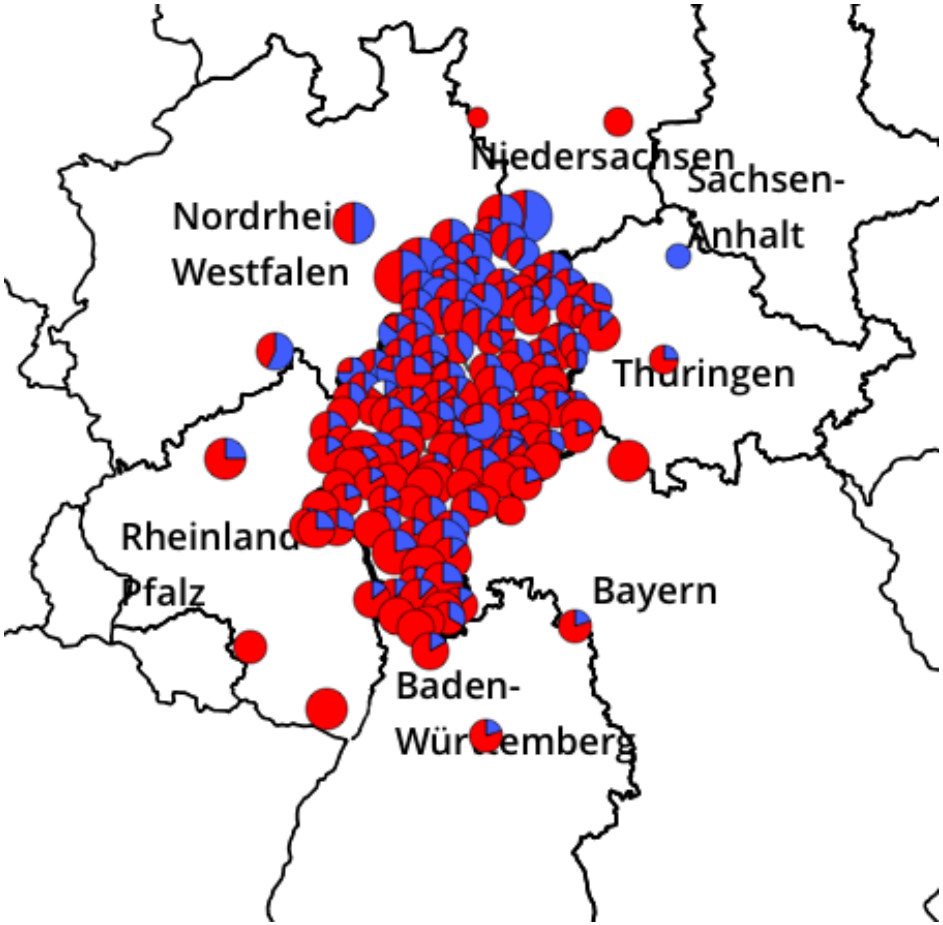
Dysfunctional presentations
(scarcely visible symbol types, numbers, names of information elements and their graphical attributes)



(from: Solau-Riebel Diss)

What we certainly do not want

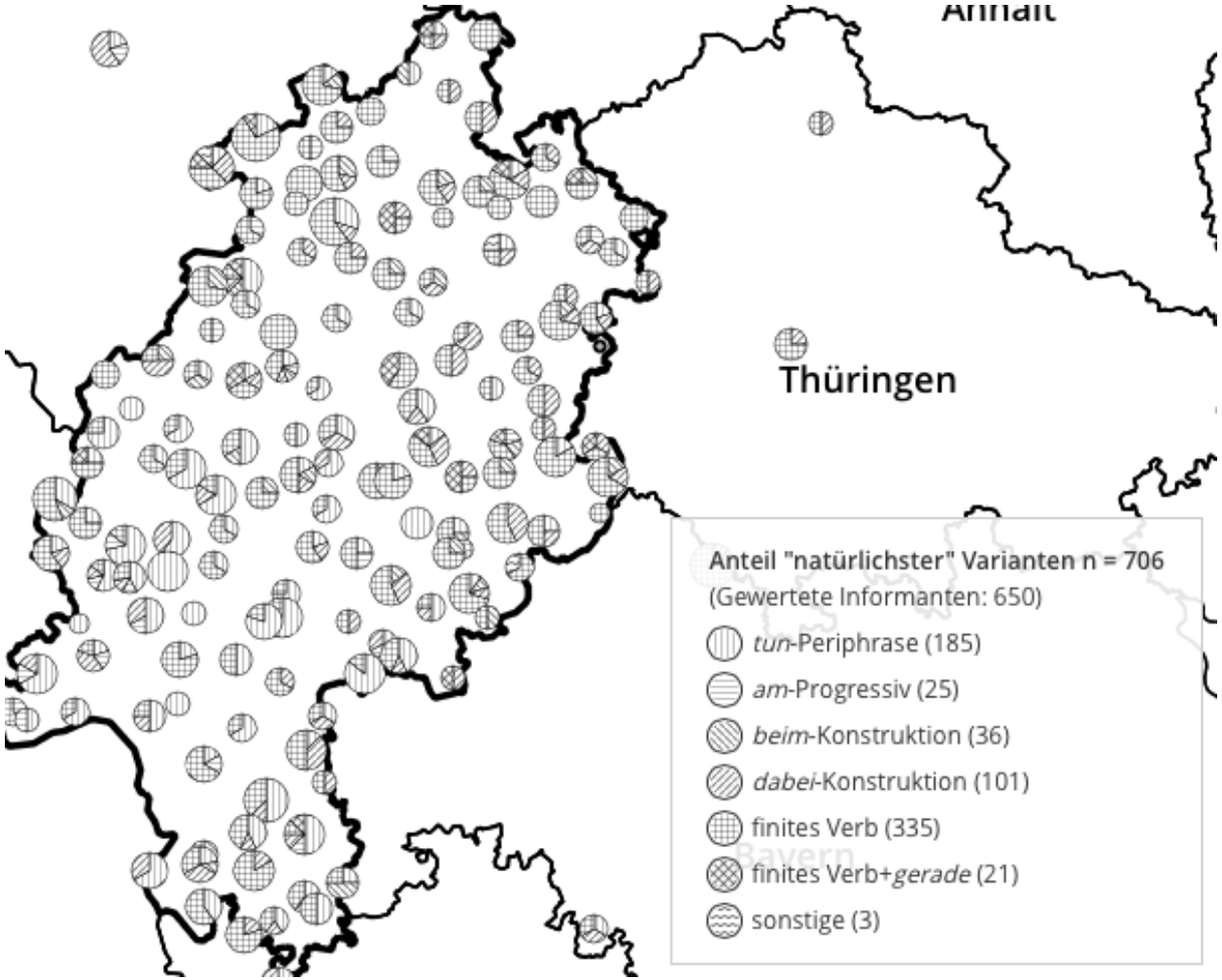
too densely packed/ superimposed information elements when zoomed out



SyHD after Zoomout

What we certainly do not want

visual **confusion**
(no popout!)

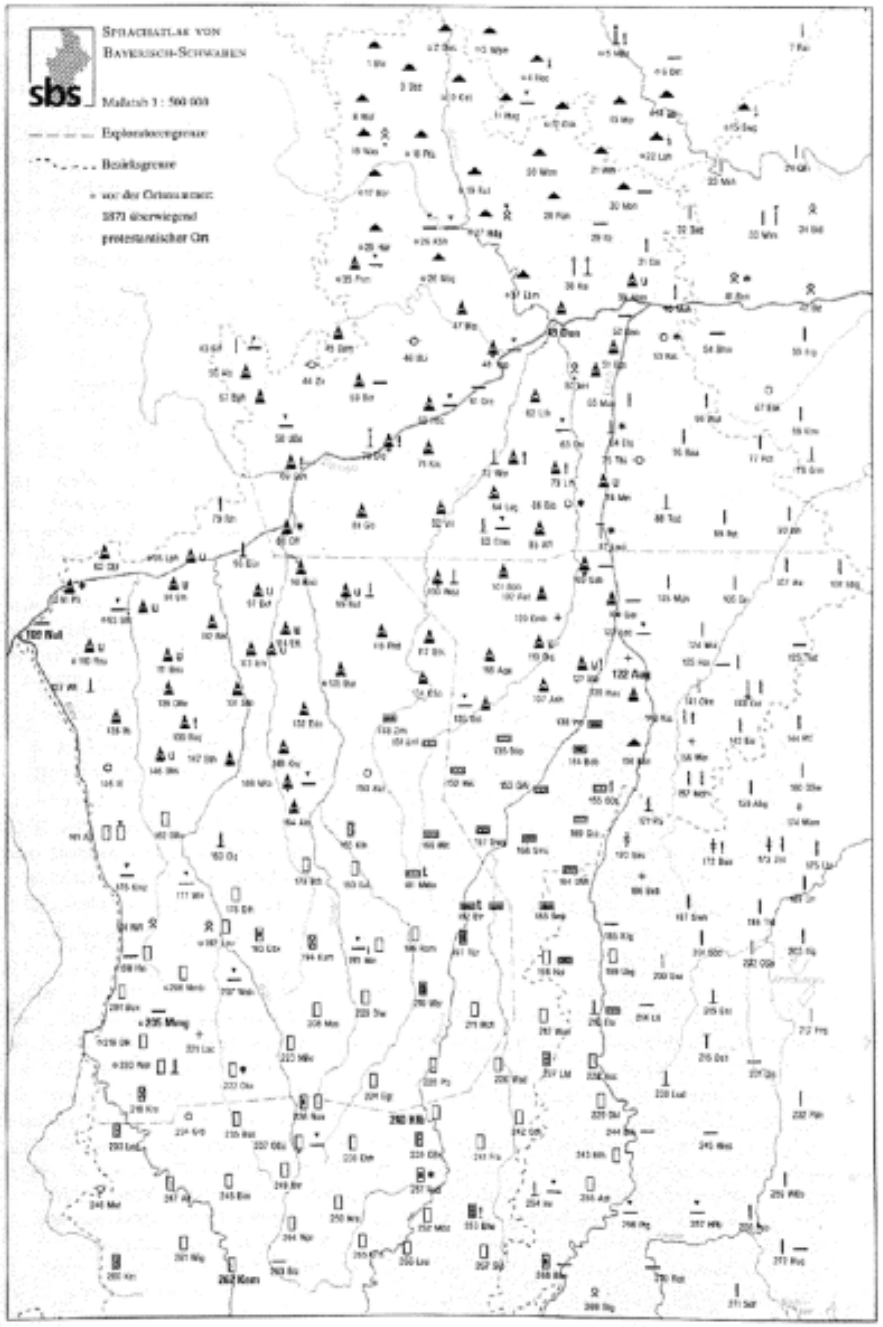


SyHD

What we certainly do not want either

visual **confusion** (no popout!)

don’t forget your magnifying glass!

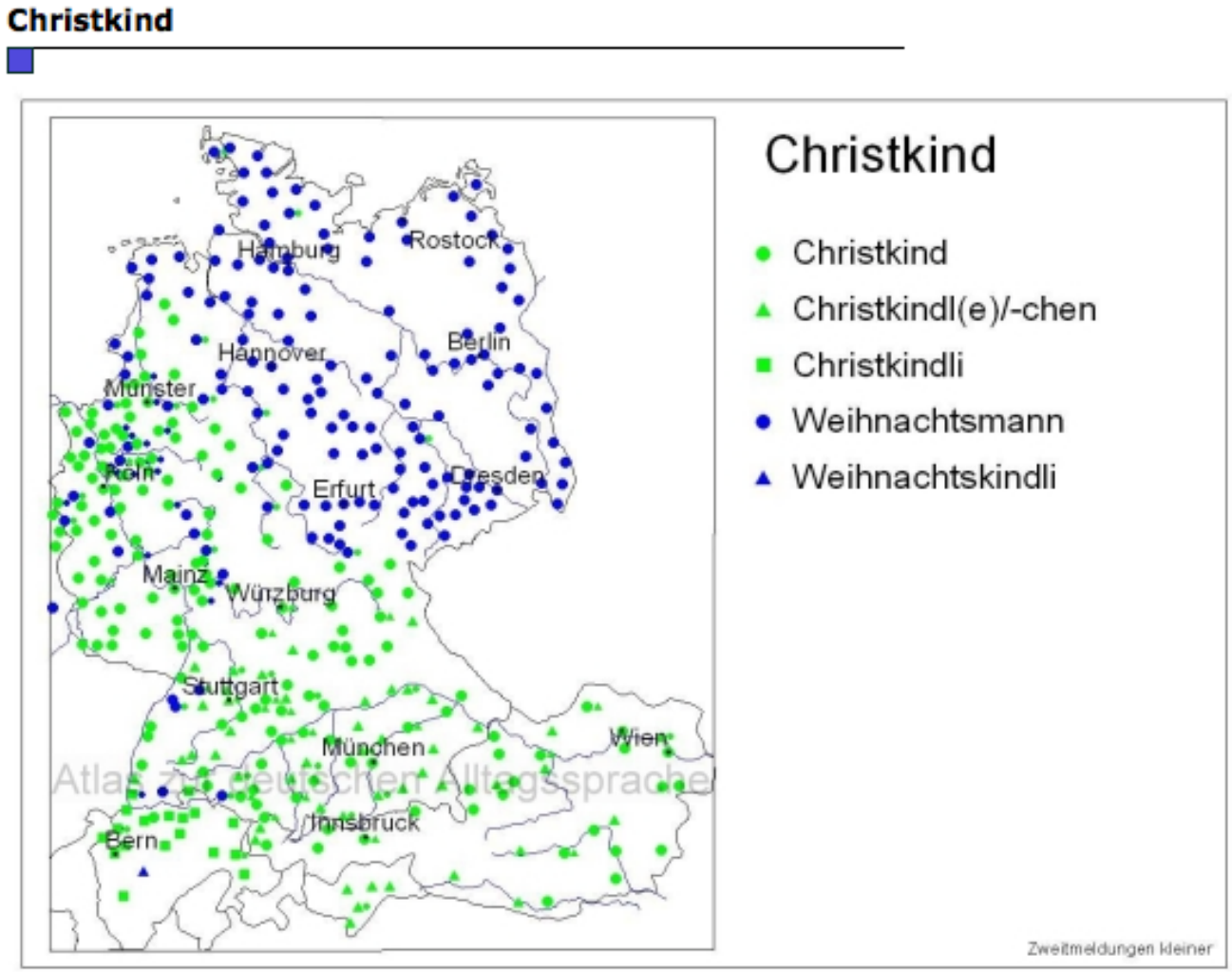


This seems ok

clearly represented
distribution of the data

But: often there is **more than one value/item per location**, i.e., this is just one of many possible/required kinds of presentations

, and still visually **confusing (different symbols, same color)!**

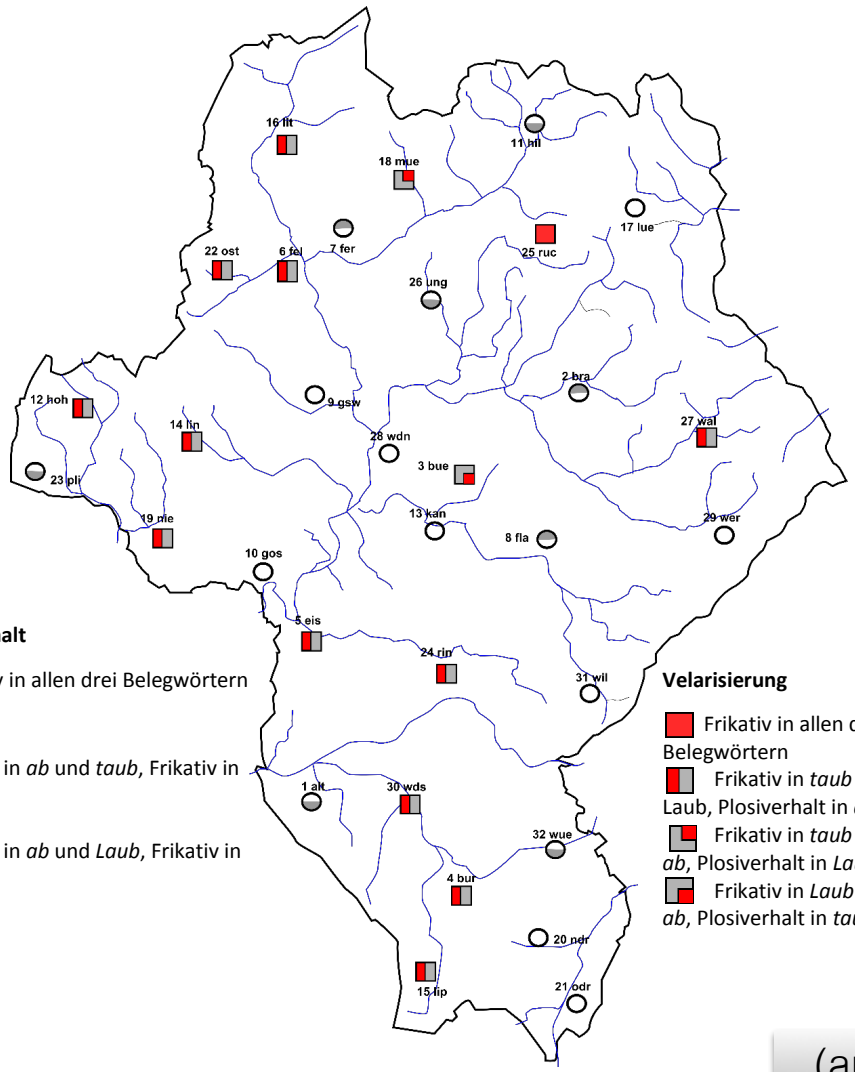


Atlas deutscher Alltagssprache

http://www.atlas-alltagssprache.de

Varying distribution per location

Karte 66



Plosiverhalt

- Plosiv in allen drei Belegwörtern erhalten
- Plosiv in *ab* und *taub*, Frikativ in *Laub*
- Plosiv in *ab* und *Laub*, Frikativ in *taub*

Velarisierung

- Frikativ in allen drei Belegwörtern
- Frikativ in *taub* und Laub, Plosiverhalt in *ab*
- Frikativ in *taub* und *ab*, Plosiverhalt in *Laub*
- Frikativ in *Laub* und *ab*, Plosiverhalt in *taub*

(aus: Solau-Riebel Diss)

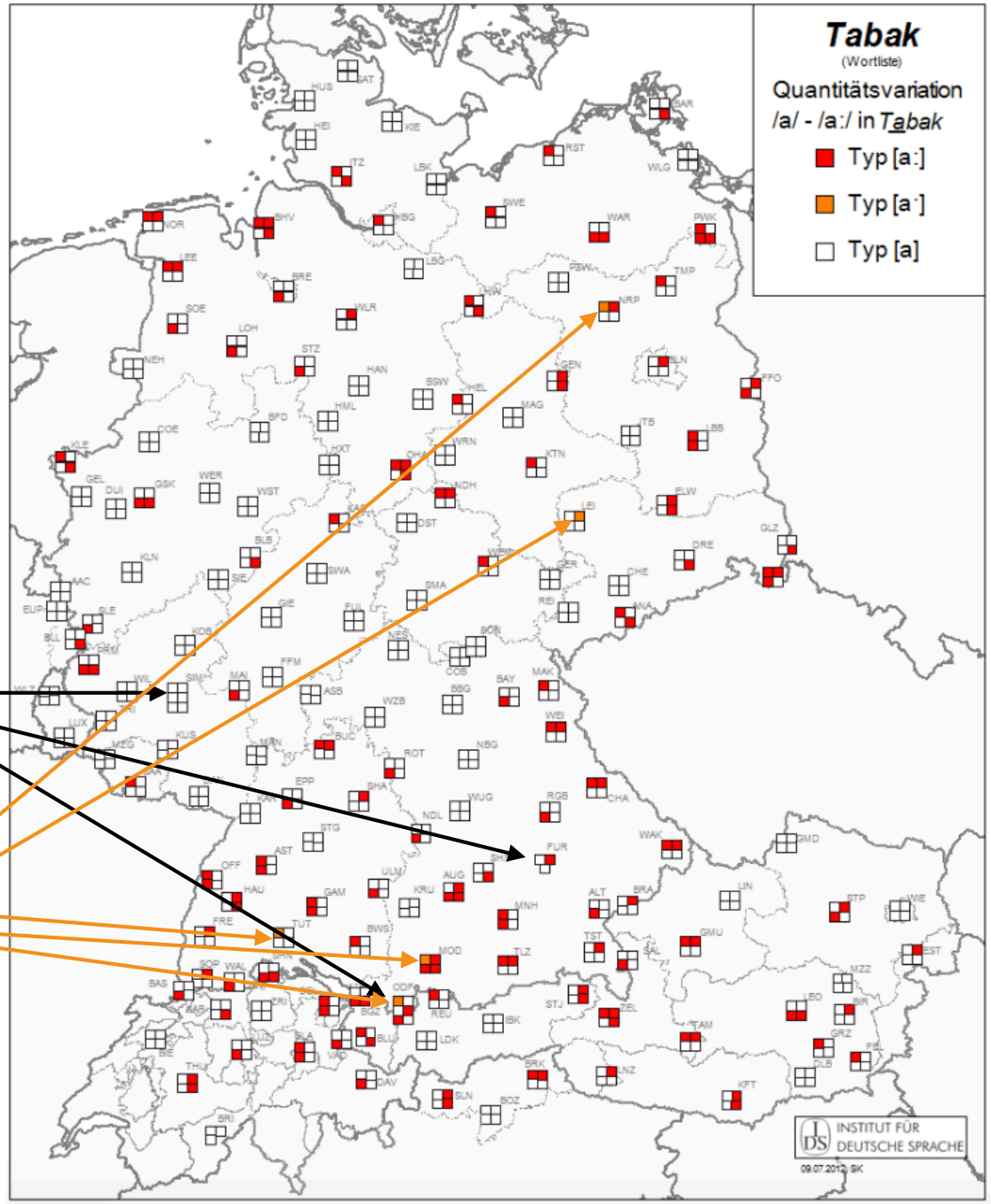
Varying distribution per location

Folie neu, nicht im Vortrag

visual **cluttering**

did you notice that?

or the occurrences of the second type?

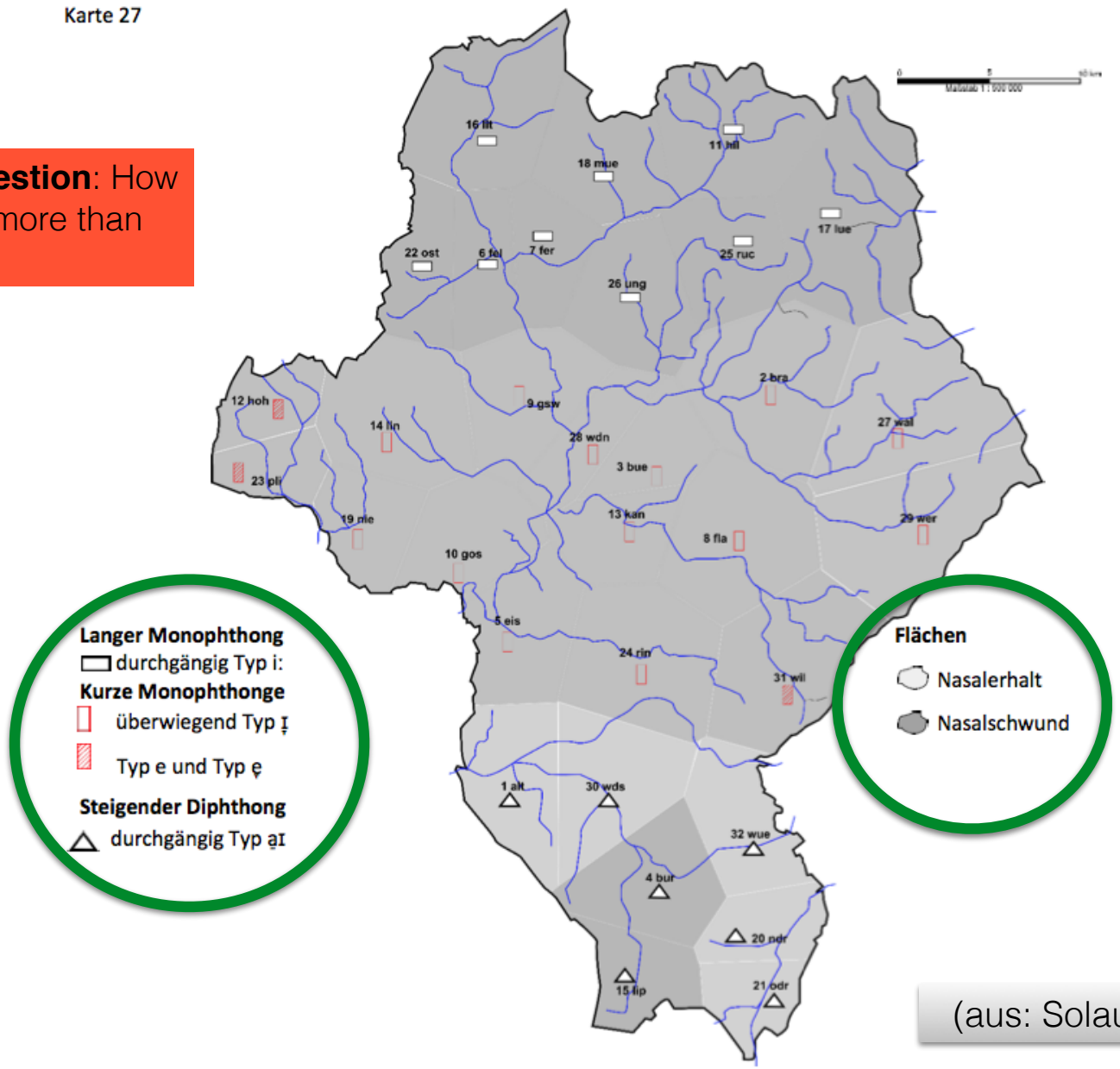


DMW Dialektatlas Mittleres Westdeutschland

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

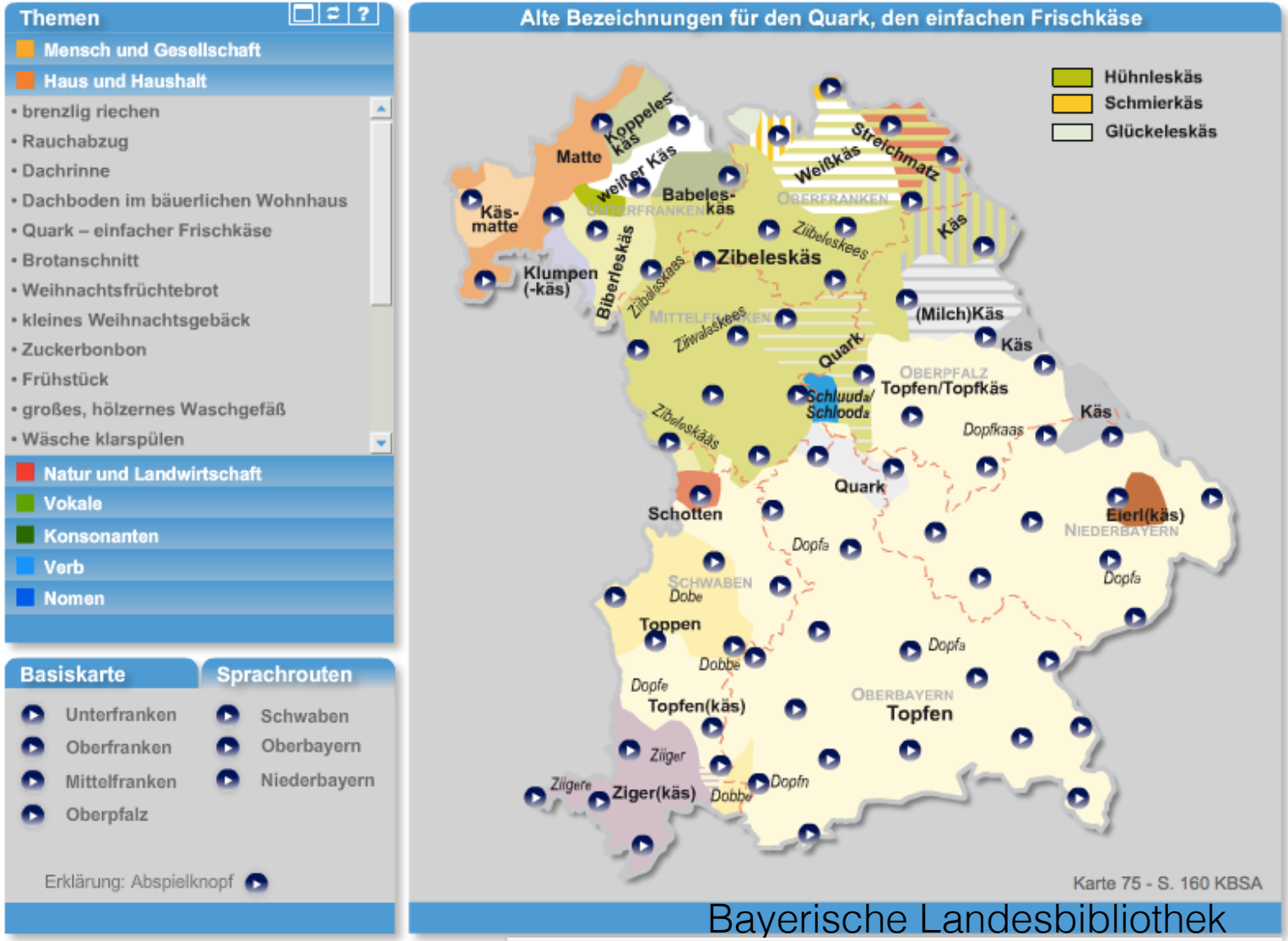
Probably needed: mix of symbols and coloured regions

Problem/Question: How to represent more than two aspects?



(aus: Solau-Riebel Diss)

What we should consider

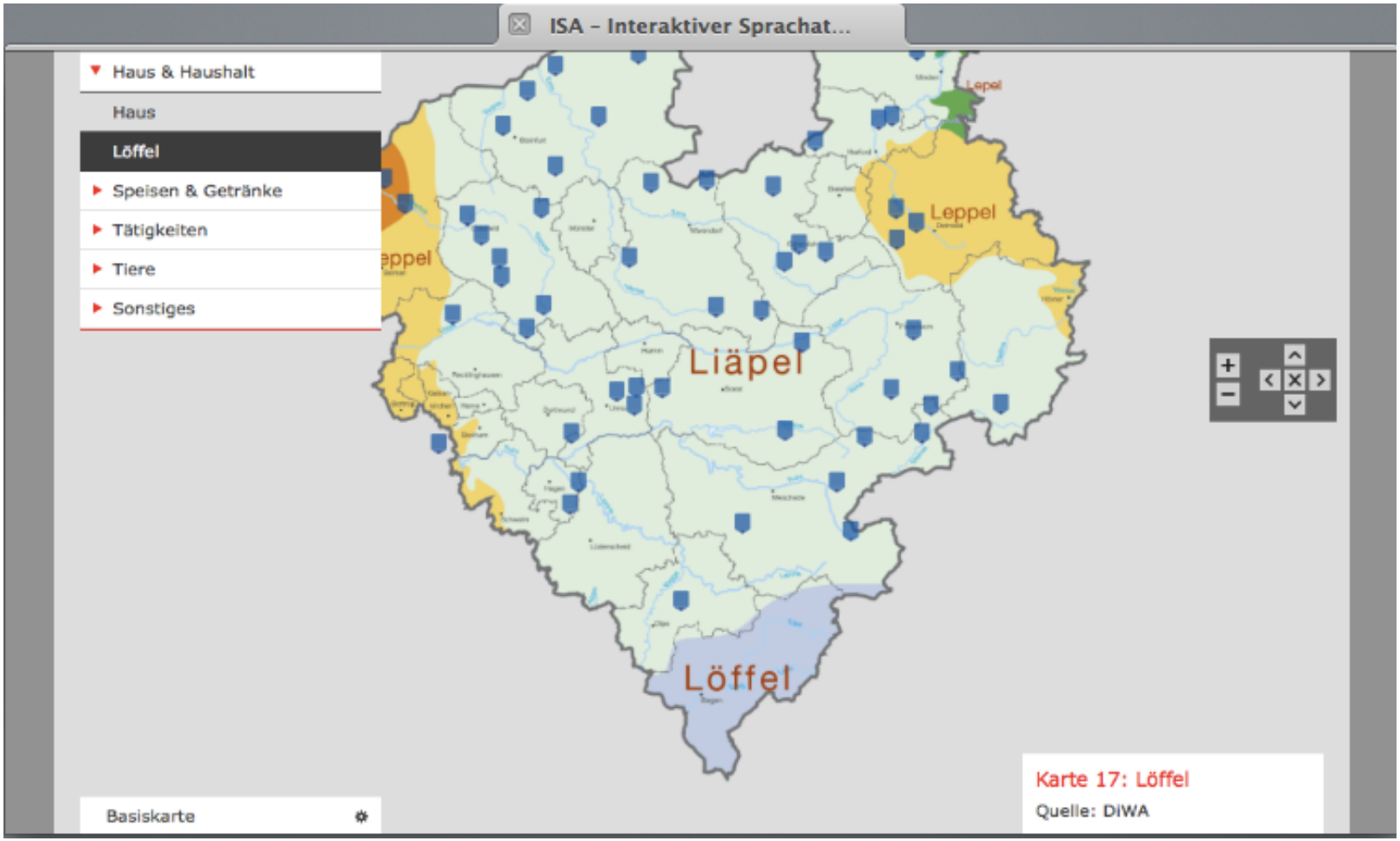


Bayerische Landesbibliothek

https://www.bayerische-landesbibliothek-online.de/sprachatlas

But: **which** symbols/colors/regions/labelings/interface aspects to use **when, how and where**?

What we should consider



Interaktiver Sprachatlas des westfälischen Platt (ISA)

http://www.lwl.org/LWL/Kultur/komuna/isa/#/10

But: **which** symbols/colors/regions/labelings/interface aspects to use **when, how and where**?

What we will probably want to offer:
Selective distribution maps à la Gabmap

Item: Georgia Select item

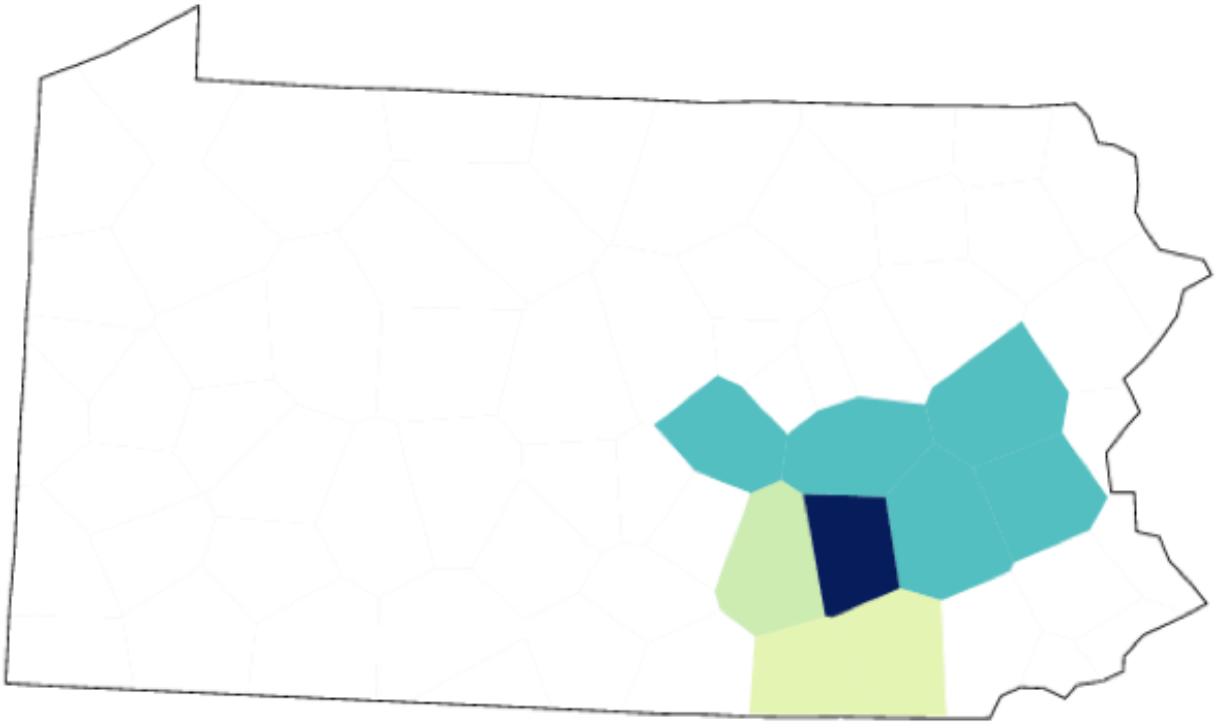
Variant(s):

- ḍʒɔ{ə}rḍʒə (1)
- ḍʒɔə·dʒə (1)
- ḍʒ̆ɔ{ə}ḍʒə (1)
- ʧ{ə‿ɹ}ʧə (1)
- ʧɔrʧə (4)
- ʧɔ{ə}{r}ʧə (1)
- ʧɔ{ə}{r‿ɹ}ʧə (1)
- ʧɔ·{r‿ɹ}ʧə (1)
- ʧɔə·ʧɪ-^ (1)
- ʧɔ^{ə}rʧə (1)

— or —

Regular expression: ?

Show distribution map



Role model for the DMW

- **SiSAL** (Siegerländer Sprachatlas)
  - distribution of phonetic types
  - on-click wav-playing
- **But**:
  - static map
  - small set of data (few locations in the Siegerland, only one informant)
  - types are computed after complete analysis
  - neat, complete covering of the area with coloured (Voronoi-)cells

http://www.mundart.sisal.uni-siegen.de

DMW Dialektatlas Mittleres Westdeutschland

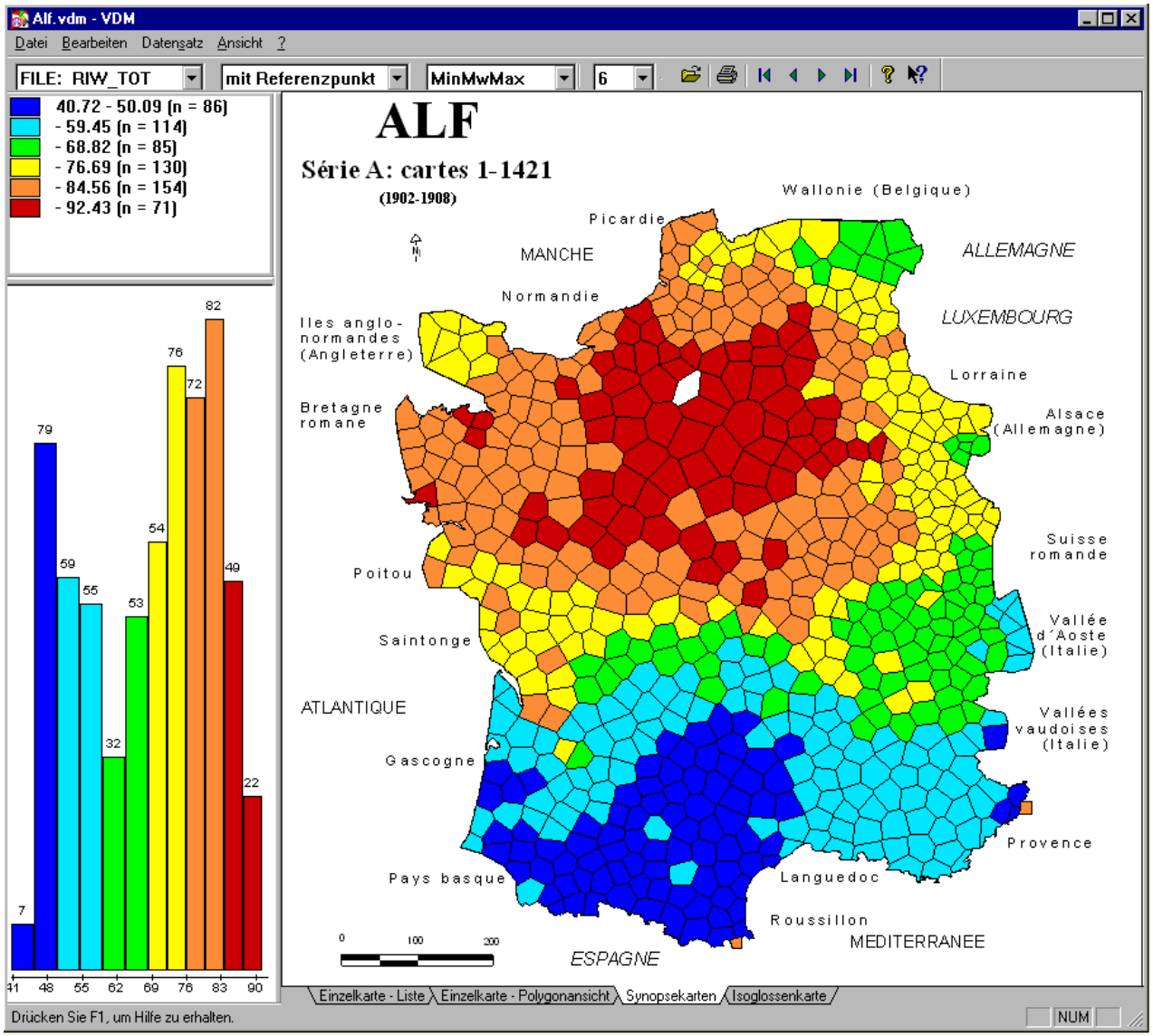
UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE DER WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

Dialectometry

- aka „quantitative dialectology“, enhances/extends functionality of dialect maps besides distribution aspects
  - + **comparison** (linguistic distances)
    - global (***difference maps***, identifying dialect areas)
    - local (***reference point maps***, comparison wrt. one location; identifying directions of language change)
  - + **interpolation by intensity estimation** (smoothing maps of discrete data)
  - + categorization/grouping/**clustering** (automatic, hierarchical grouping)
  - + **validation of distinctness** of clusters with multi-dimensional scaling (MDS)
  - + **discovery** of (fuzzy) dialect areas by factor analysis (Pickl 2016)
  - + aspects of interaction/**exploration**

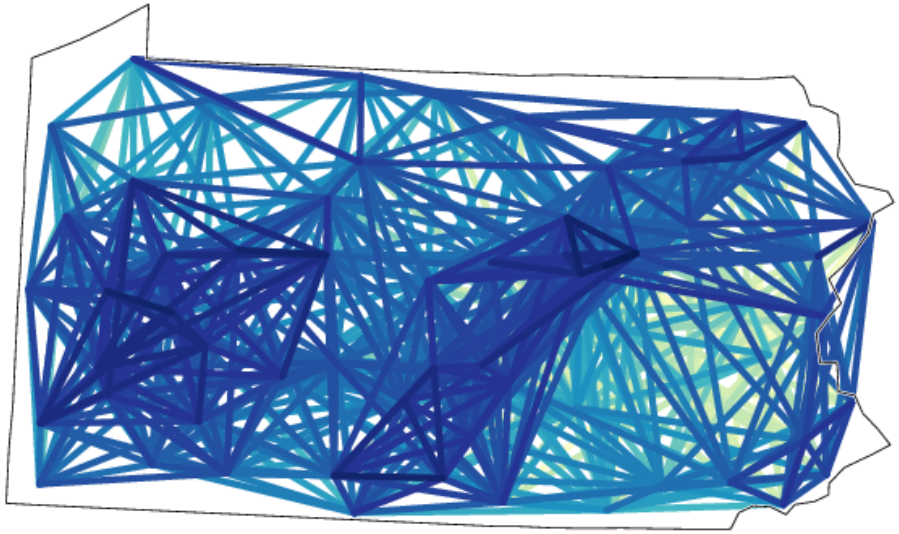
VDM reference point map

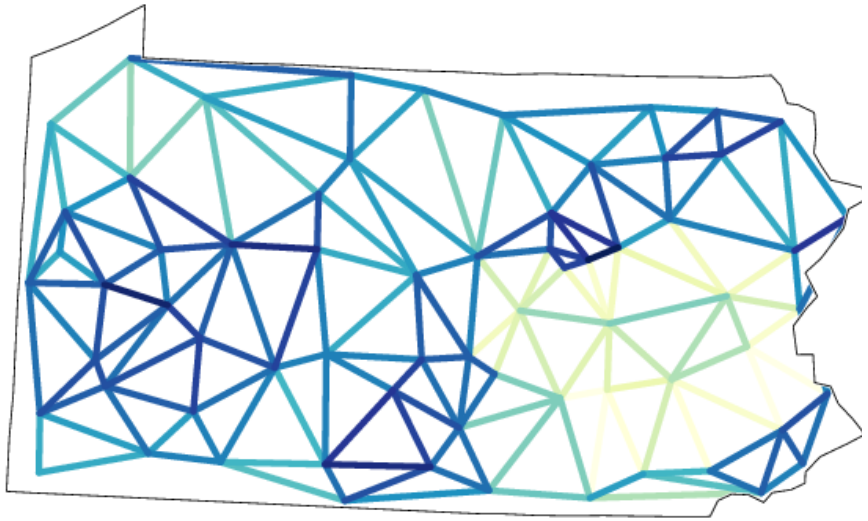


„Das Bild zeigt, wie durch Klick auf ein Polygon der Karte der Prüfsbezugspunkt schnell verändert werden kann. ***Dadurch wird es erstmals möglich, interaktiv*** Dialektgebiete zu erforschen“

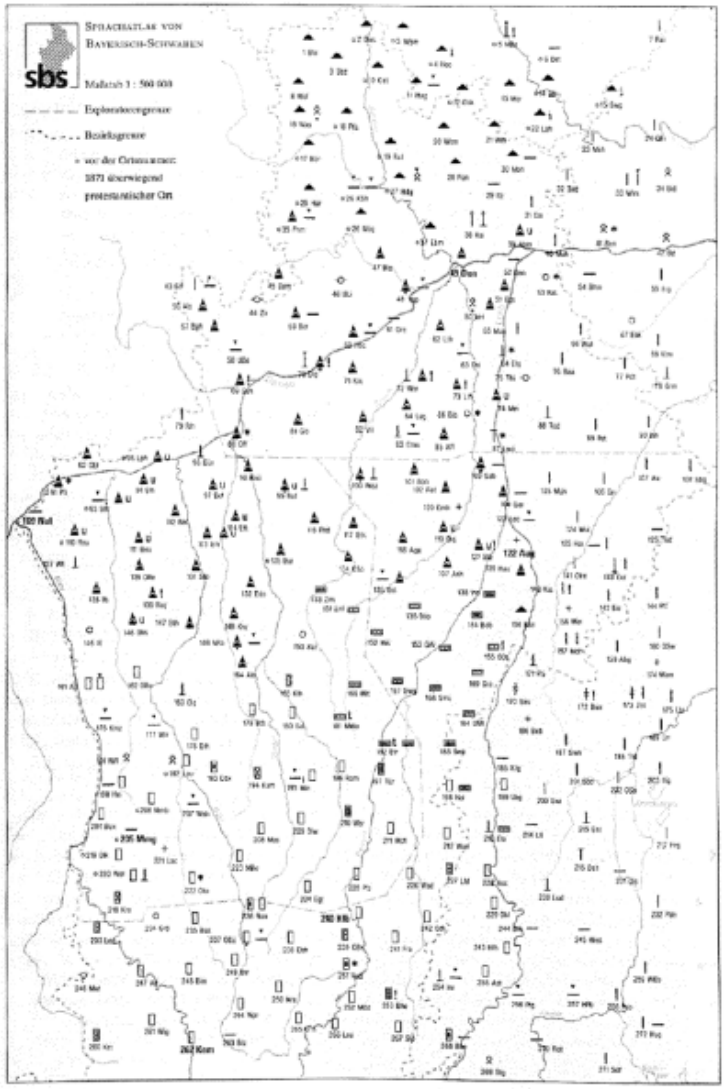
(http://www.dialectometry.com/dmdocs/, 1998+, Goebl/Haimerl, Uni Salzburg)

Gabmap difference maps

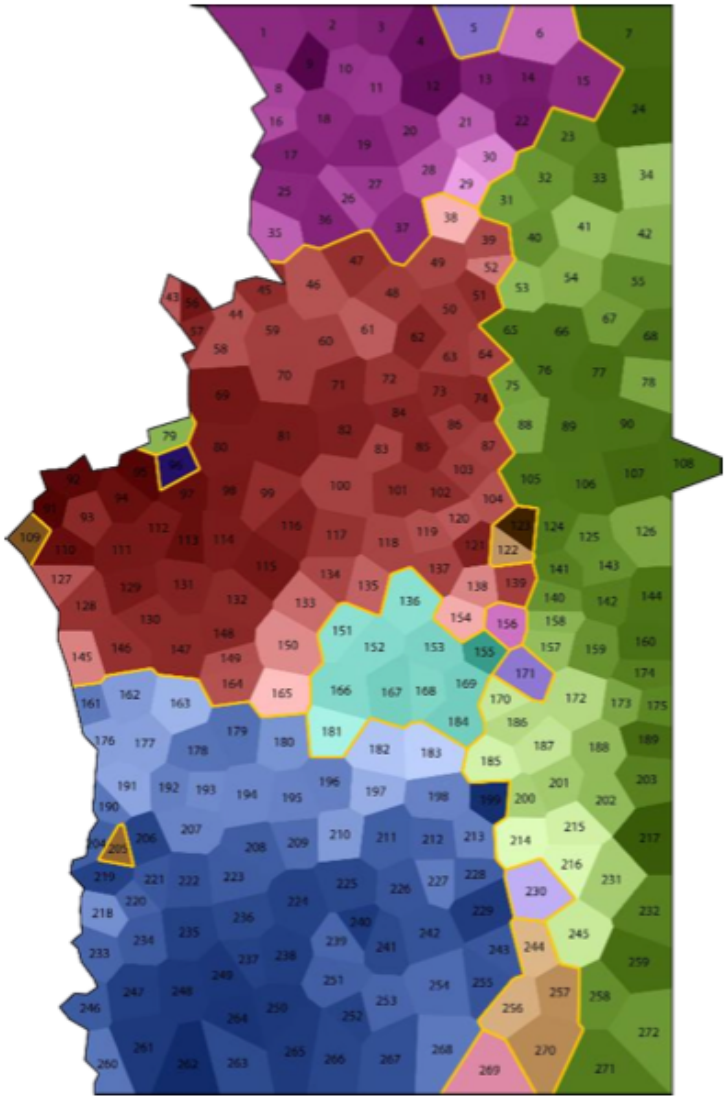




Intensity estimations à la Geoling



„Kellerassel“

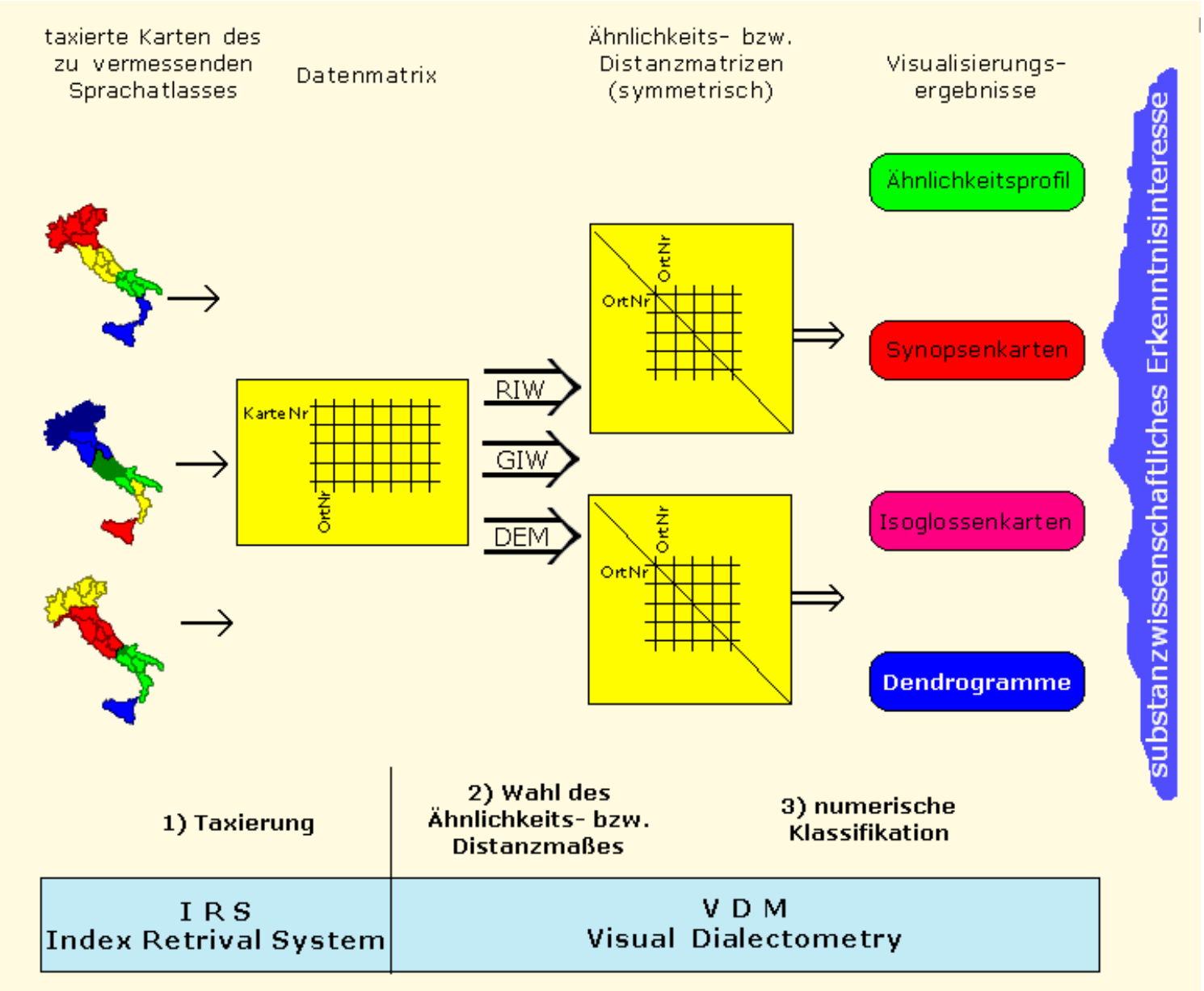


Intensity estimations based on distances

Dialectometry

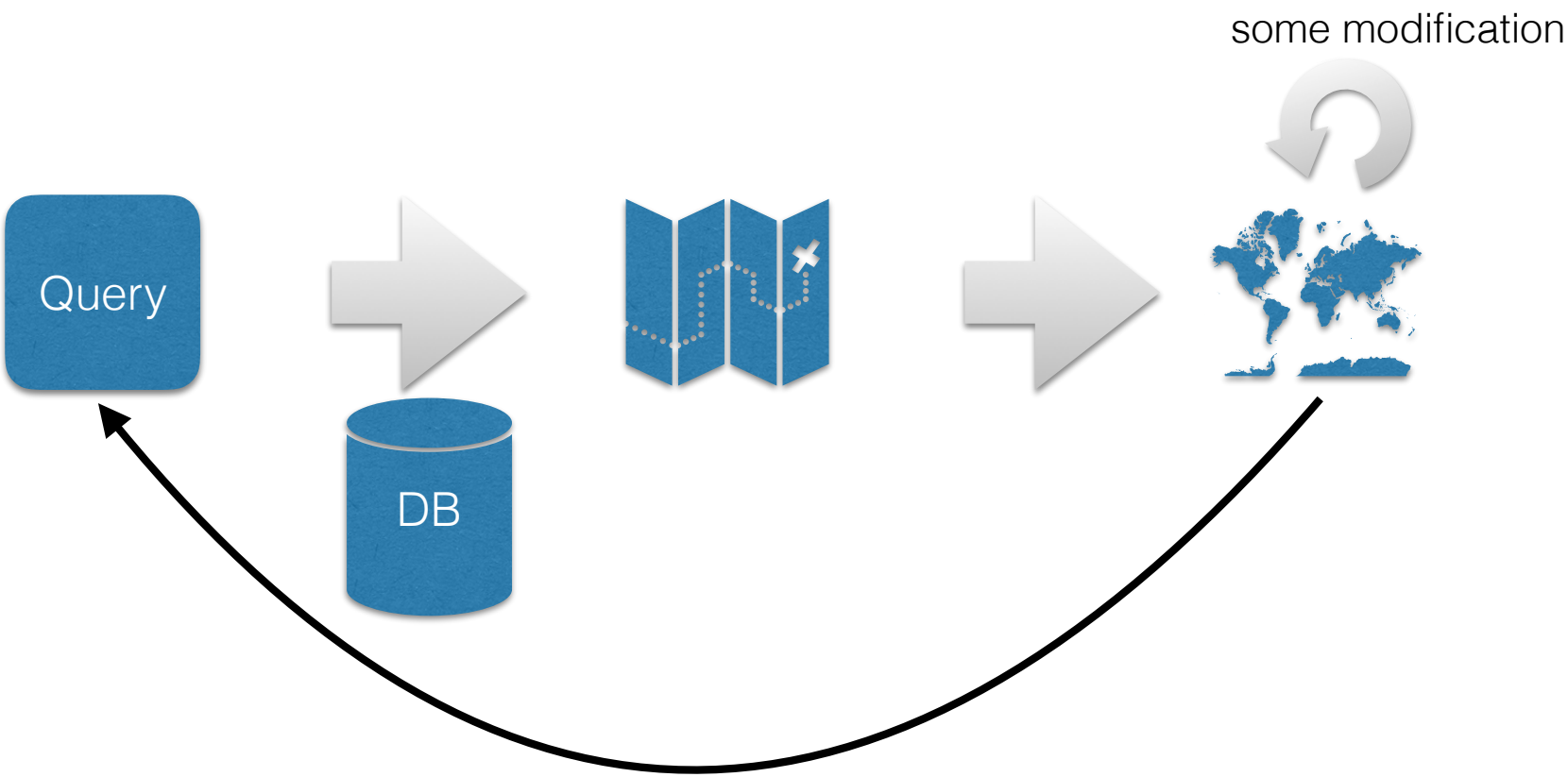
- Popular approaches/systems:
  - **VDM** (Visual Dialectometry, Goebl/Haimerl, http://ald.sbg.ac.at/dm/)
  - **Gabmap** (Nerbonne et al., https://www.gabmap.nl)
  - **Geoling** (various authors, https://www.geoling.net)
- **But**:
  - task-, application- or system-specific approaches without geo-referenced content
  - interaction/exploration is restricted and/or rigid

„Dialectometric chain of processing“ (Goebl/Haimerl)



Old conception

Methods-to-system/application approach
Query-Maps-Presentation-Cycle



New conception:
From presentation to interaction/(exploratory) visual analysis

Query-/ data-/ purpose-oriented operations

selection, combination, comparison, choice of statistical method etc.

Display-oriented operations

Zooming, blend in/out, highlighting, symbol/ color choice etc.

Geo-map-centred
platform independent, non-proprietary, browser-based
rather colours than symbols
general javascript-based procedures/libraries (e.g., Open Layers, D3)
for presenting geo-referenced content
interest-/analysis-driven changes of presentation
(from data-based to data-driven research, Pickl 2013)

or, more simpler:

old conception:

data → maps → presentation → analysis

new conception:

data → visual analysis

Maxim(s)

- Synthesis of Shneiderman (1996), Keim et al. (2010) and Ertl (p.c.), slightly modified:

  1. Overview (show important things) first
  2. offer intuitive options (e.g., zoom/filter, limit, focus, navigate)
  3. details on demand
  4. show relationships
  5. easy switching between aspects to be presented

Desiderata

- on the fly generation of
  - dialect maps
  - types (names for aggregations of similar data)
- presentation of information which is
  - zoom-adaptive
  - clearly arranged
  - adjustable (for selection and filtering)
    - e.g., names, landscape features, borders etc.
    - selecting geo-maps (e.g. OpenStreetMap)
- modes/kinds of presentation (depending on interest/task/query)
  - Symbols vs. (Voronoi) cells
  - separate layers
  - dialectometry operations

Distinctions

- **visibility** of information:
  - on click / hover (on the map)
    - -> in a tooltip
  - by selection (in some menu)
    - layers, features
  - context sensitive
    - e.g., showing names depending on scale/resolution

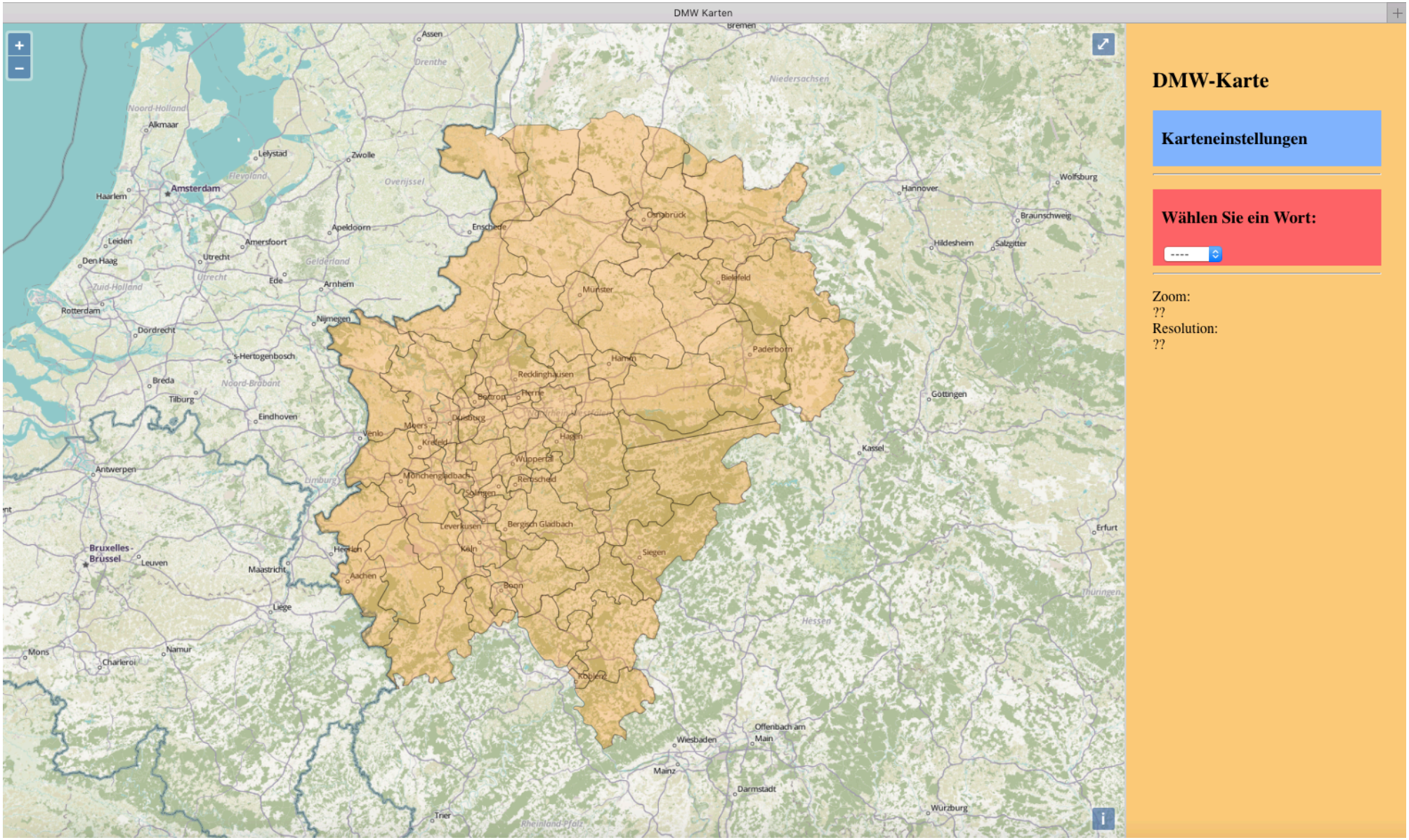
Distinctions

- „**importance**“ of information:
  - foreground: query related aspects
  - background: landscape aspects etc
  - perhaps each on more than one layer
- **presentation** of information:
  - symbols (pie, bar, other; simple/complex)
  - areas/cells
    - replacing or complementing symbols
    - colouring (Cloropleth- or distribution maps)
  - mixed versions

Distinctions

- **level of detail**
  - zoom independent complete overview vs.
  - zoom dependent (perhaps partial, less detailed …) presentation
    - e.g., imprecise presentation in indicative cluster or heat maps
- **necessity** of presentation, e.g.,
  - always showing symbol for mp3-item vs.
  - on click

Starting point/screen: the DMW region

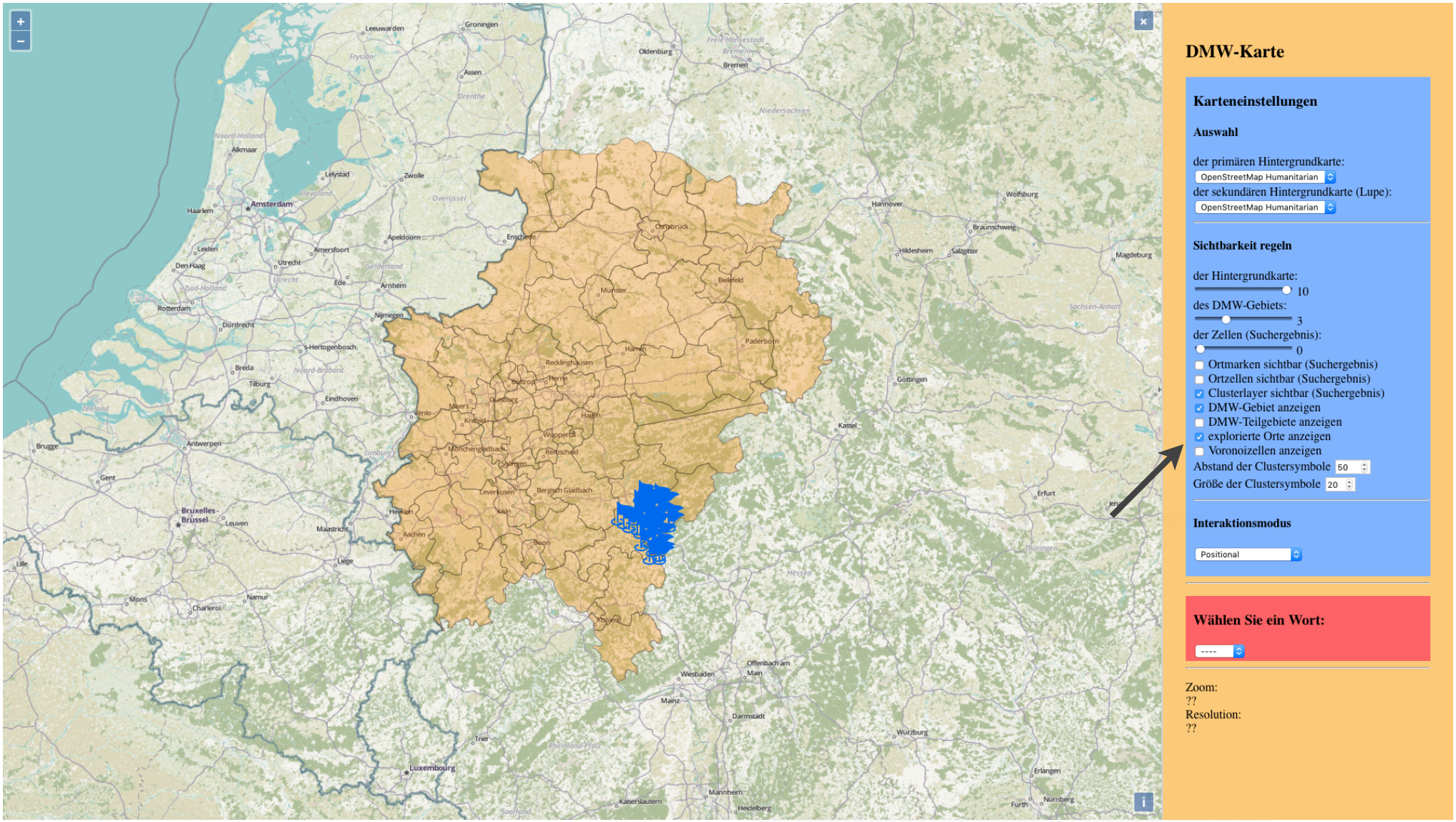


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

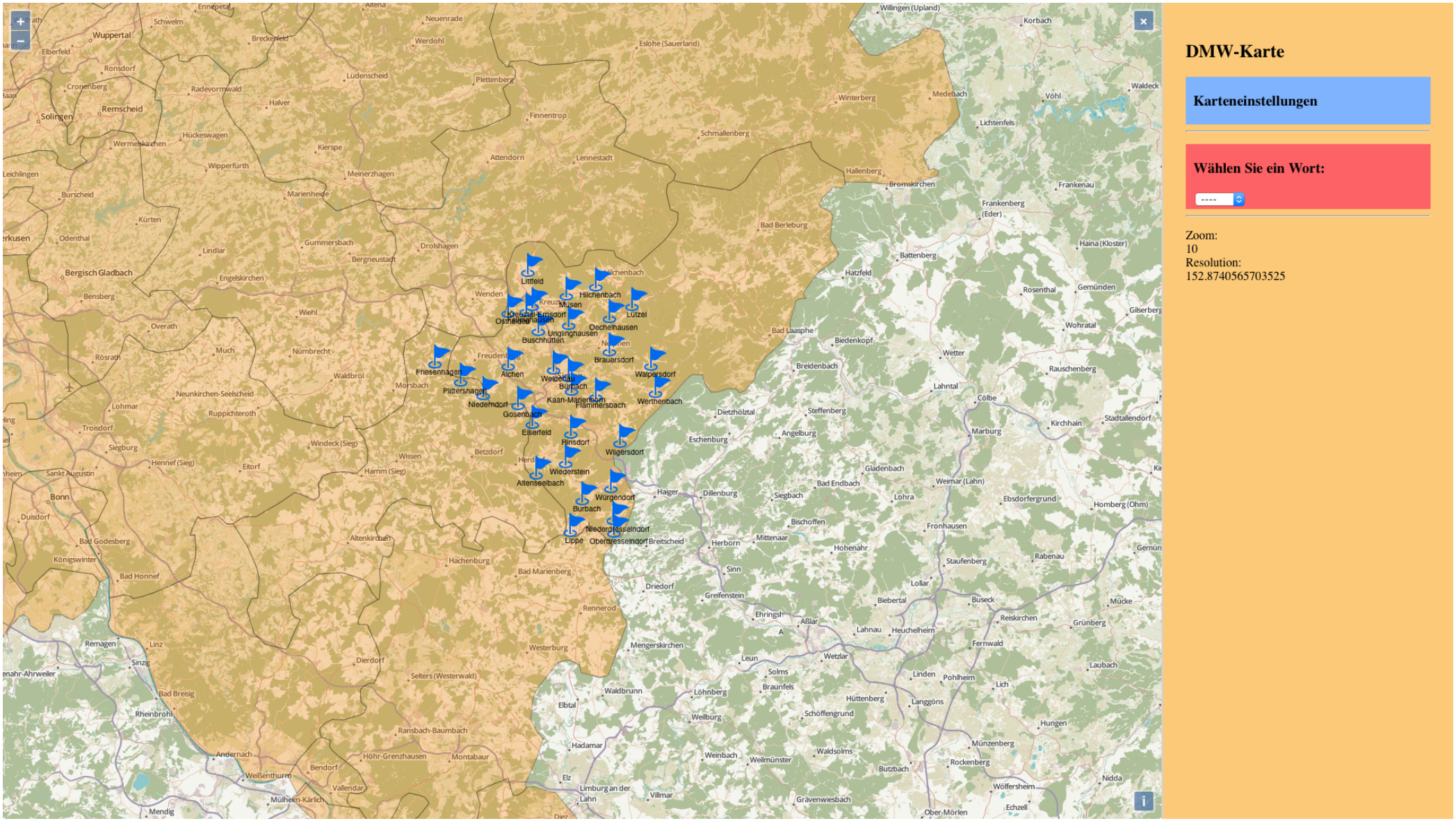
AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

The problem: visibility <-> size of region



Test data courtesy of P. Solau-Riebel

after two double-clicks

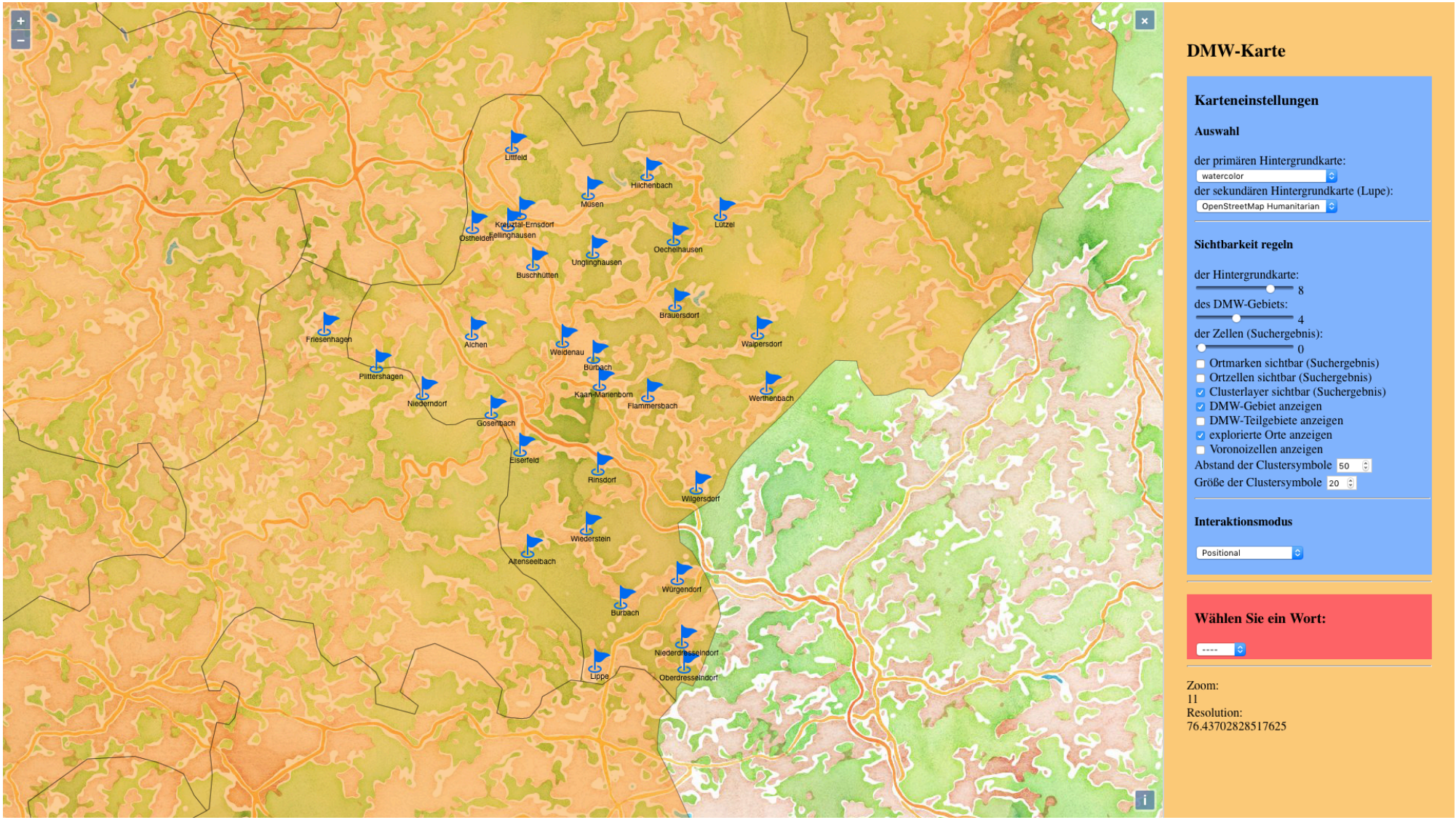


DMW
Dialektatlas Mittleres Westdeutschland

UNION
DER DEUTSCHEN AKADEMIEN
DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE
NORDRHEIN-WESTFALEN

alternative background map (watercolor)

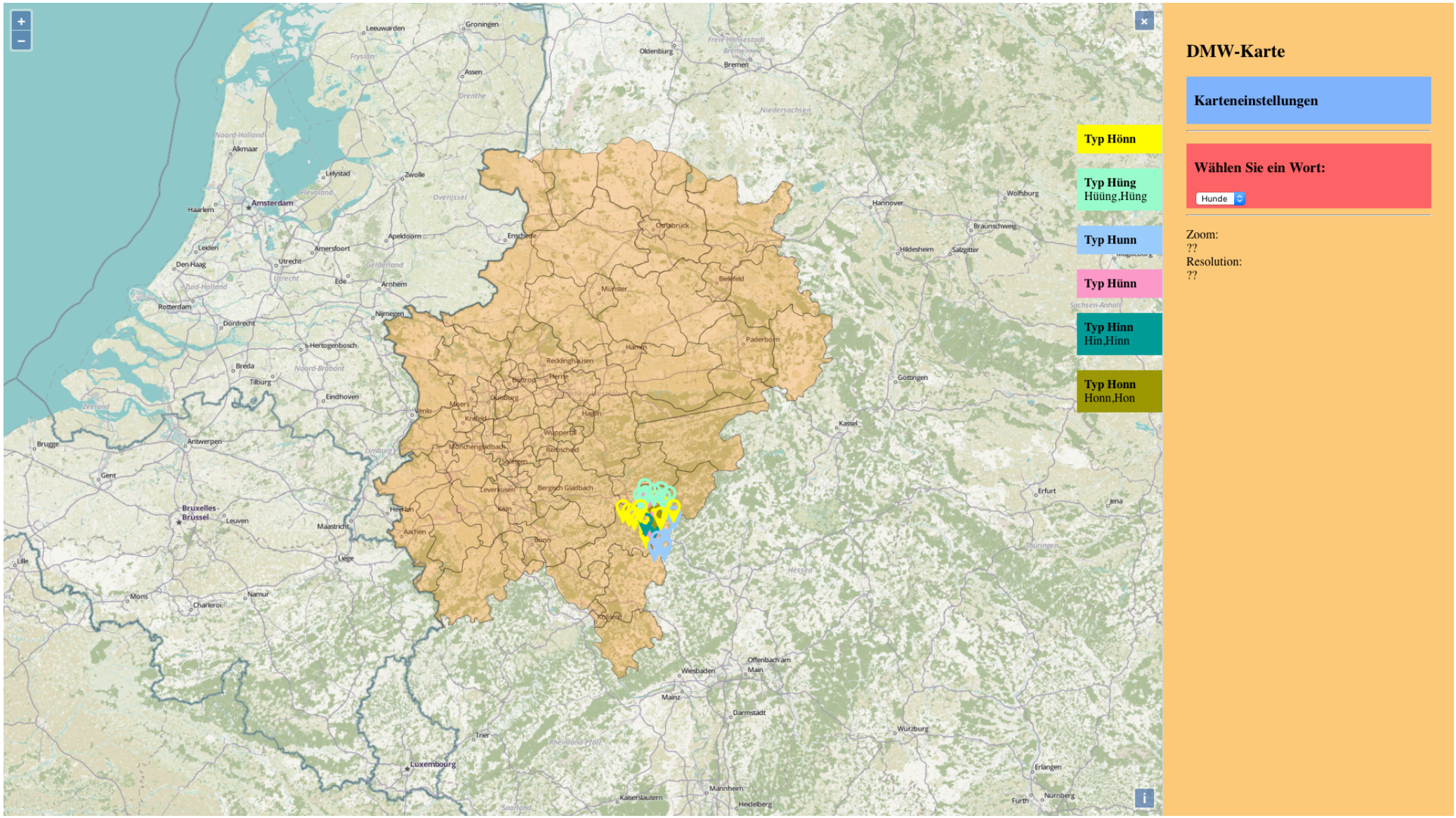


DMW Dialektatlas Mittleres Westdeutschland

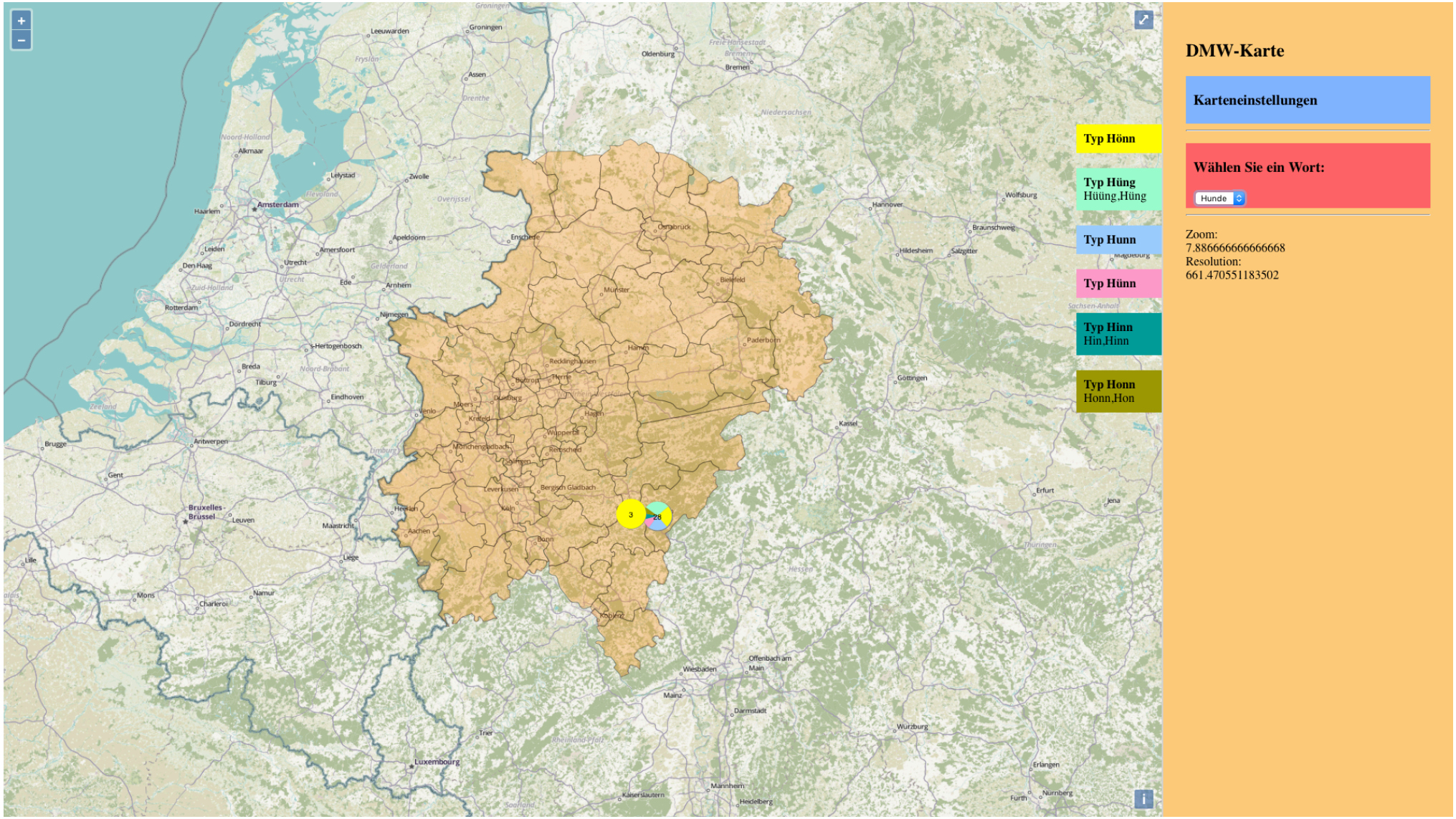
UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

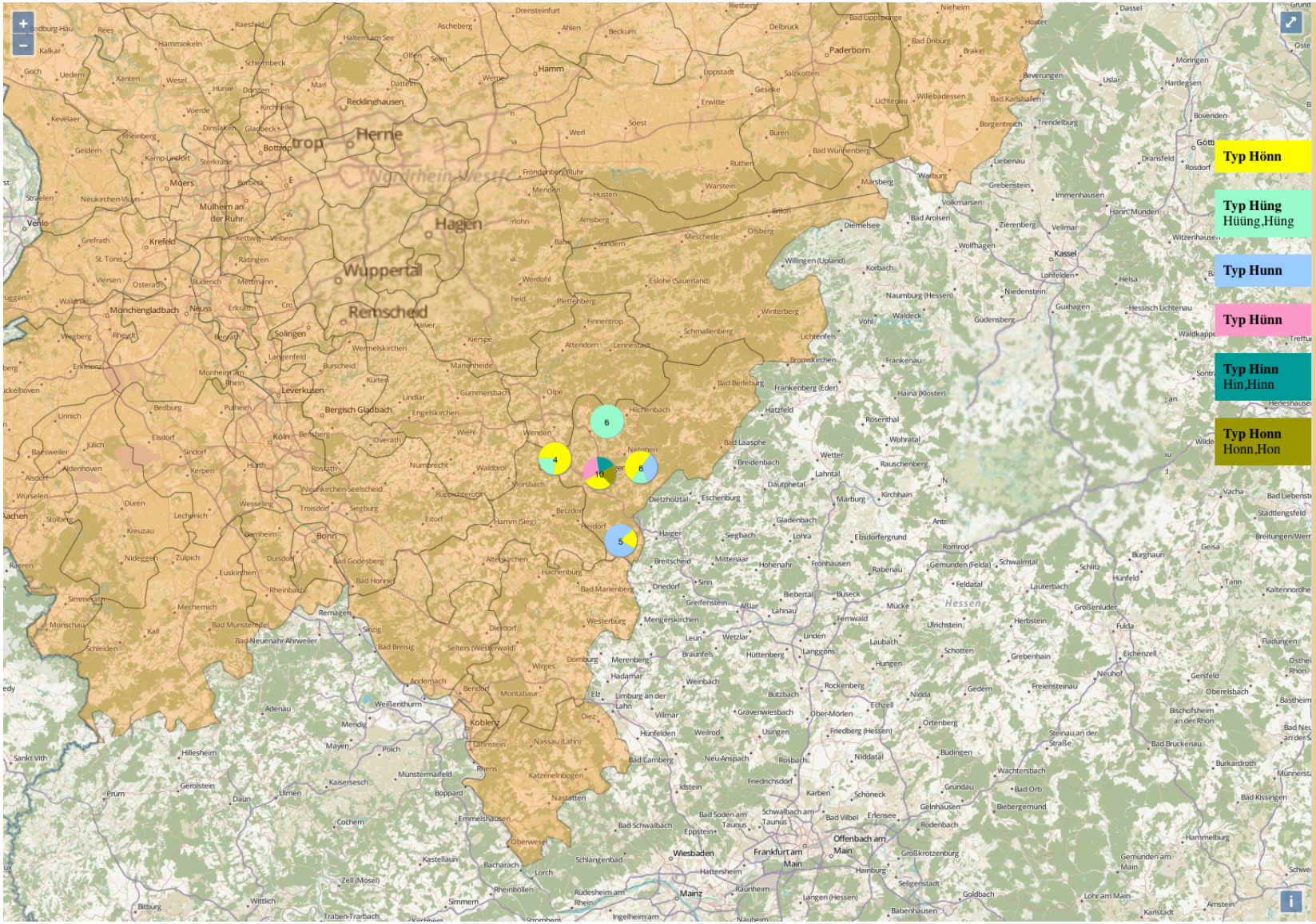
a query: back to the problem



The solution: low-granular data presentation (cluster layer)



double-click zooming in



DMW-Karte

Karteneinstellungen

Wählen Sie ein Wort:

Hunde

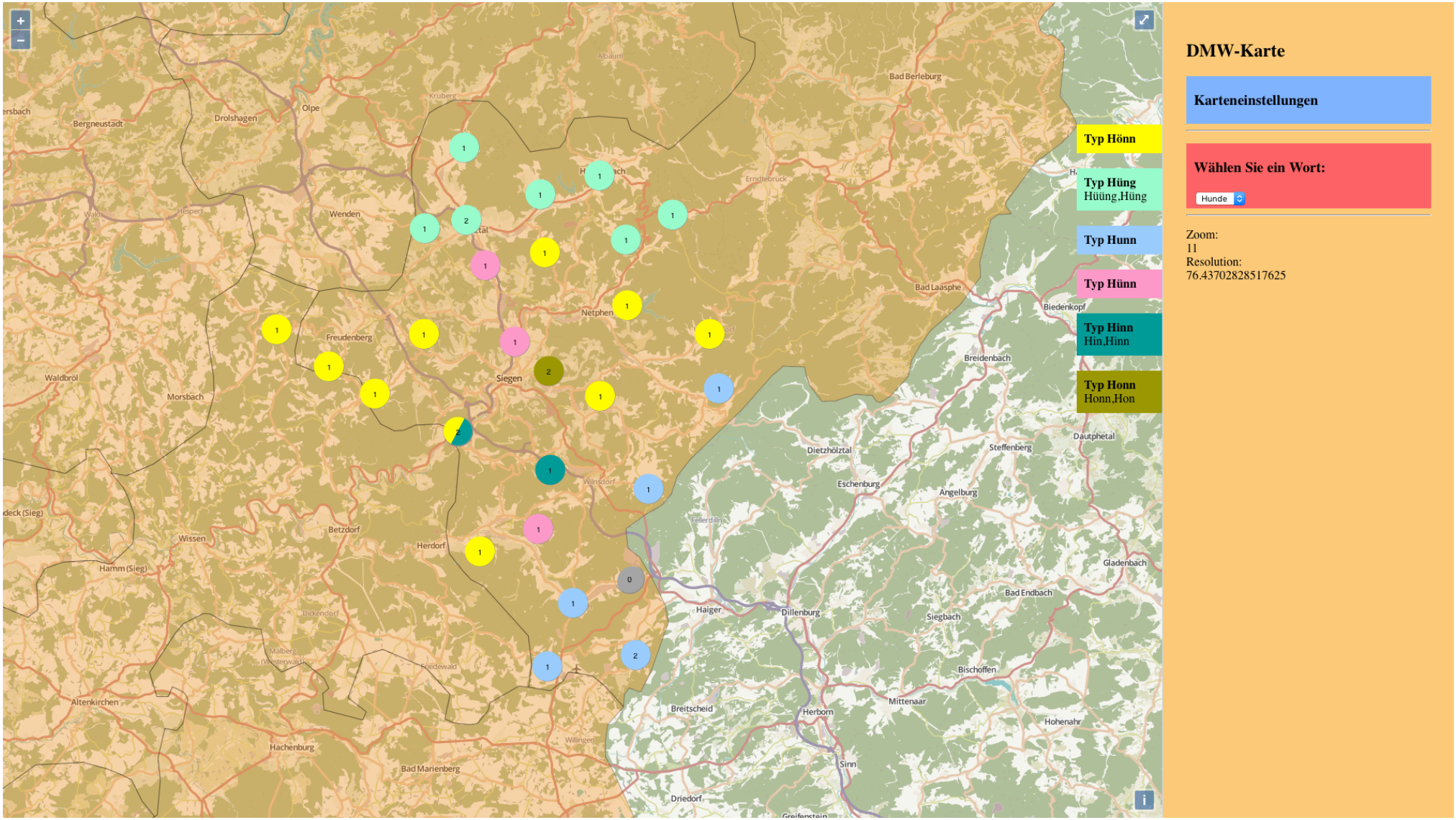
Zoom:
9
Resolution:
305.748113140705

DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

double-click zooming in

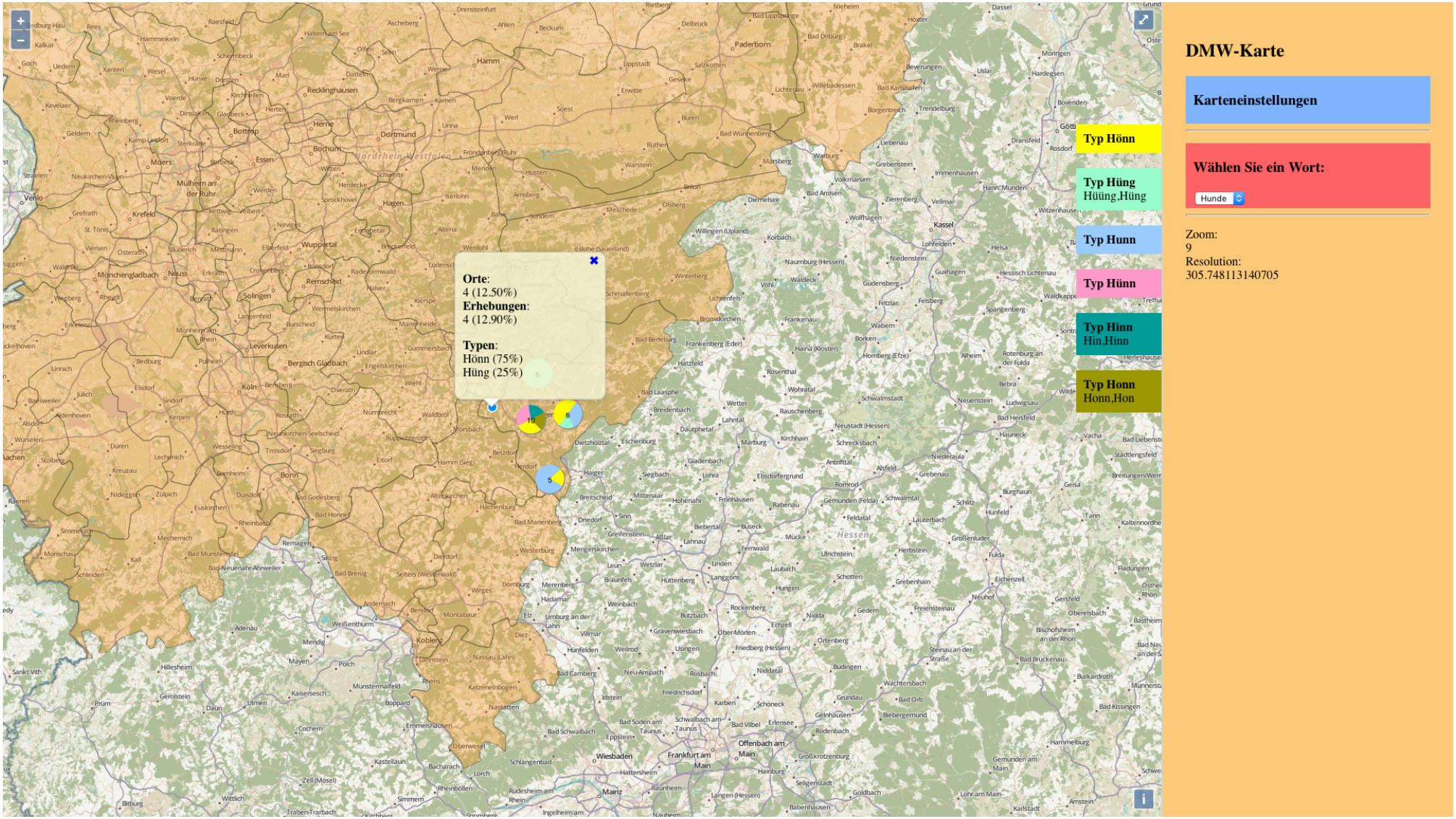


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

tooltip-information by hovering over cluster symbol

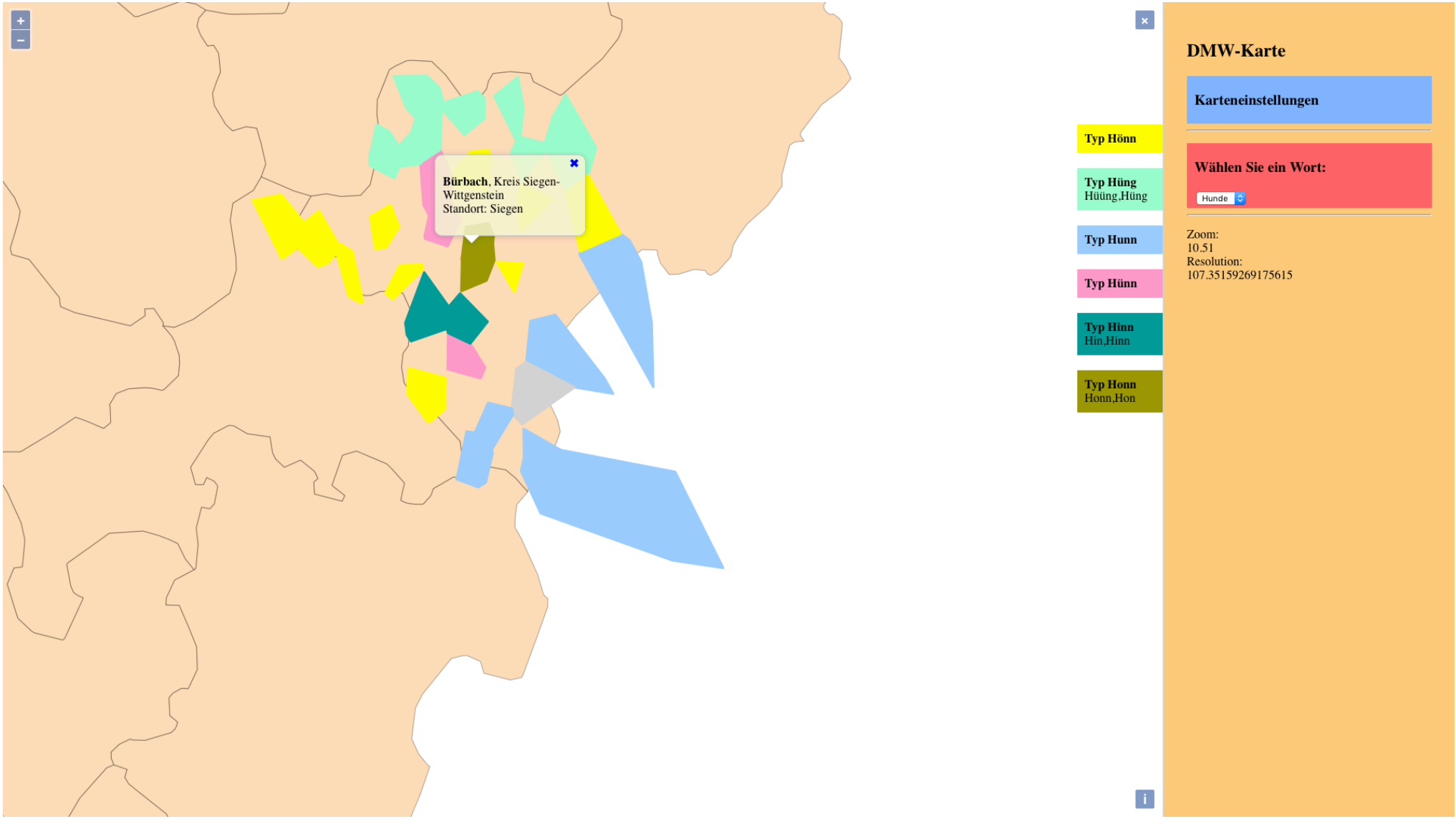


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

other presentation options selected, click in cell shows place information

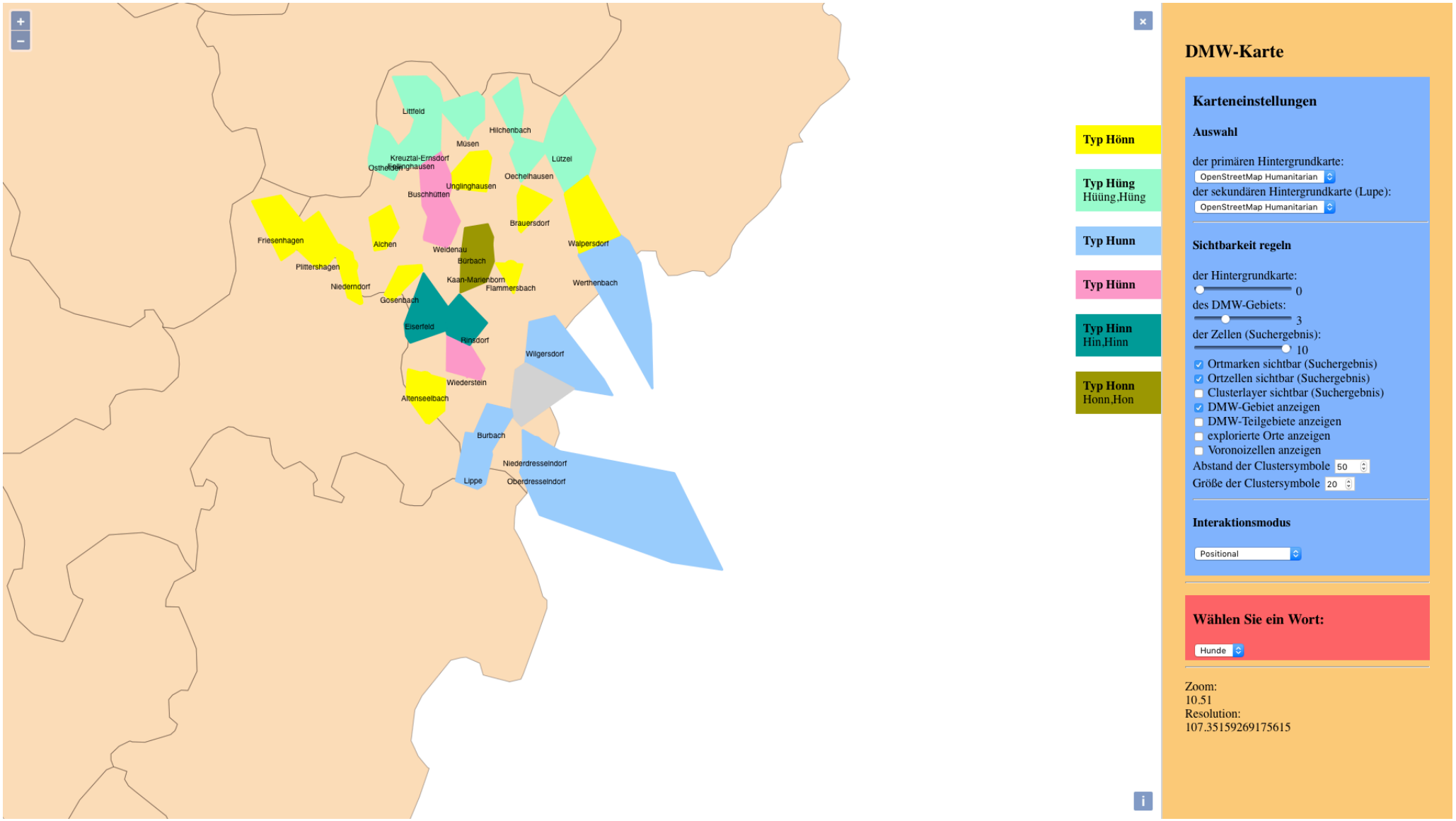


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

with place markers (+names)

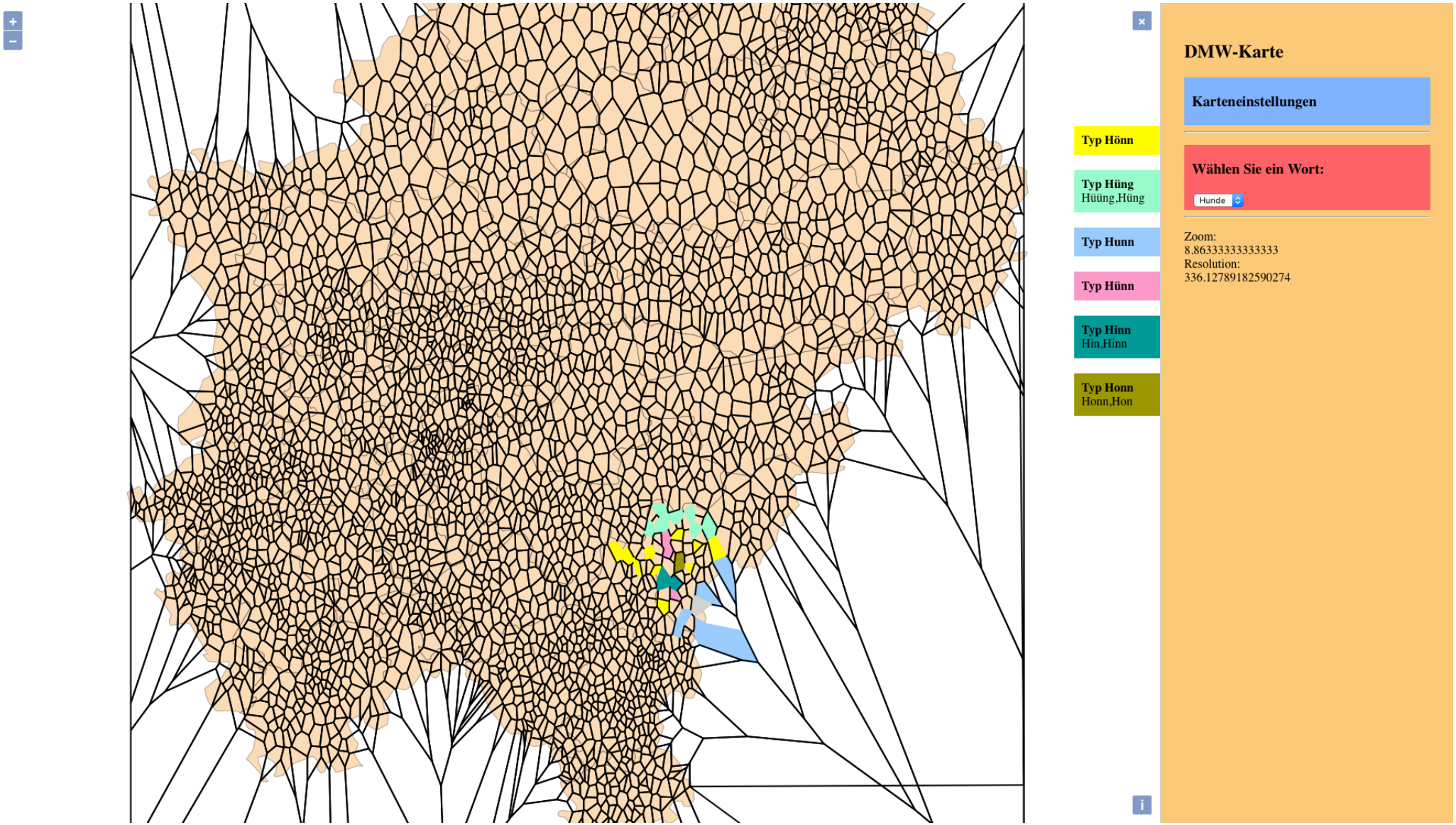


DMW Dialektatlas Mittleres Westdeutschland

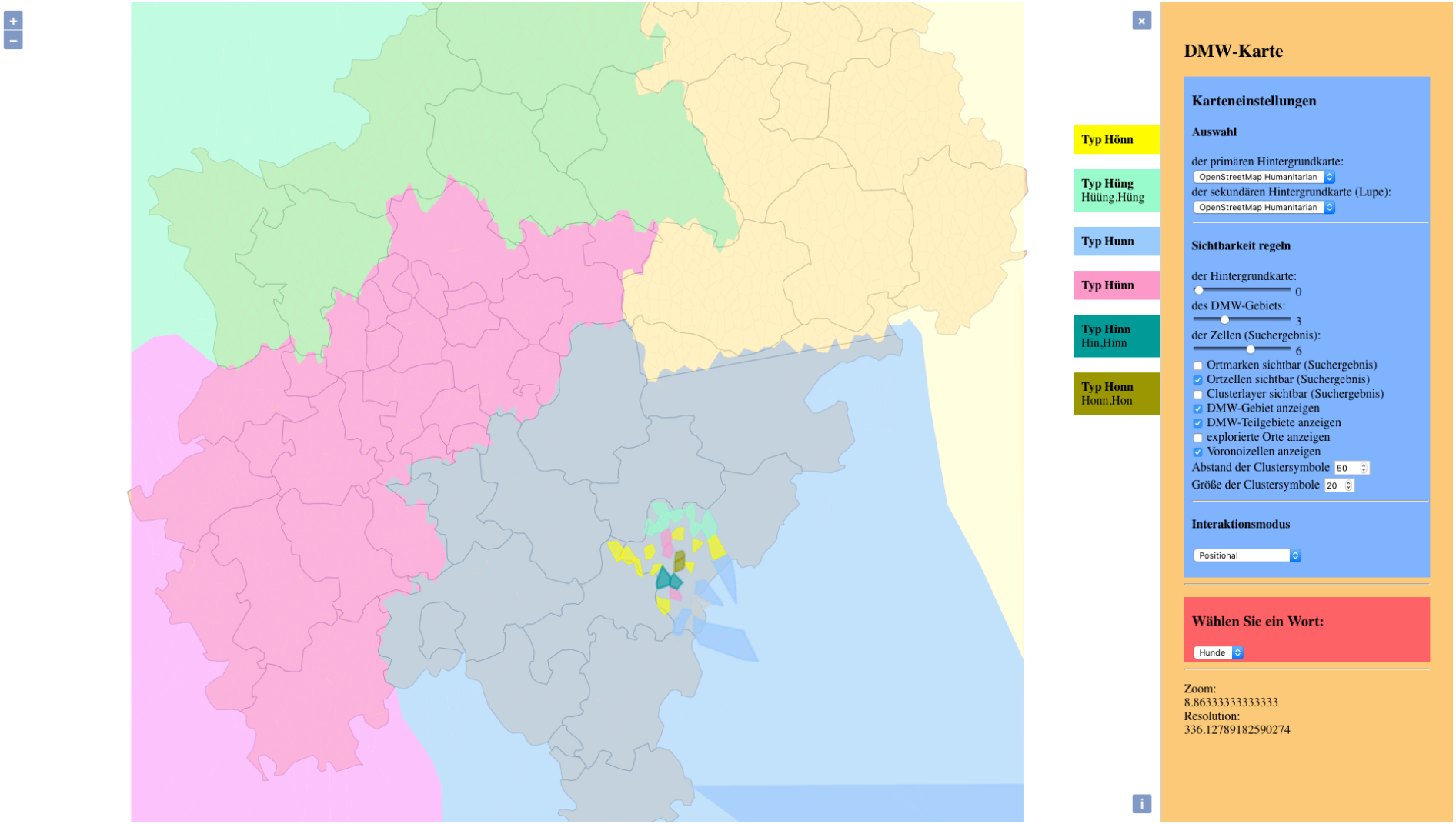
UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

all the voronoi cells of the region



the four project regions

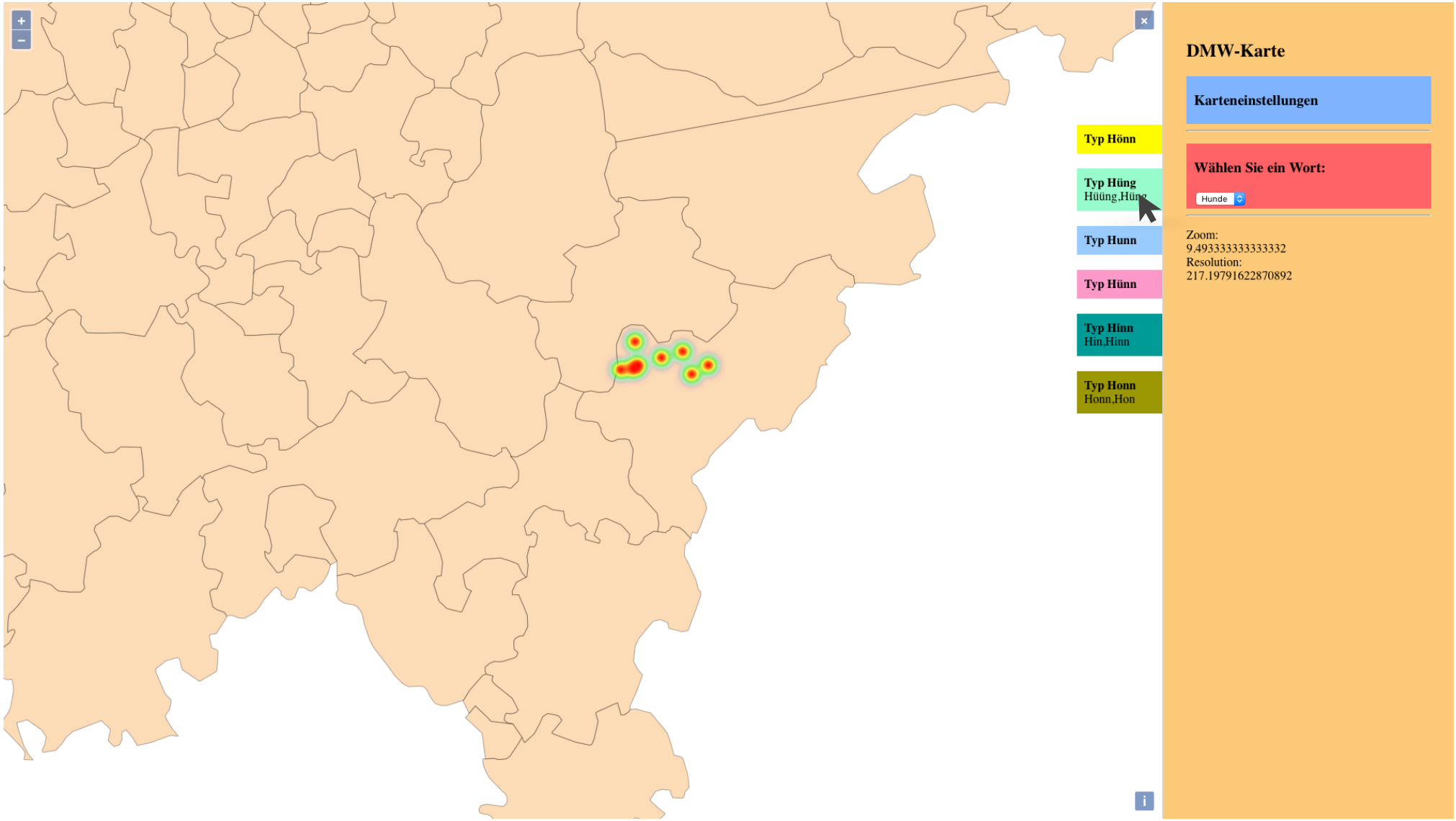


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

indicative presentation: heat map symbols when mouseover type

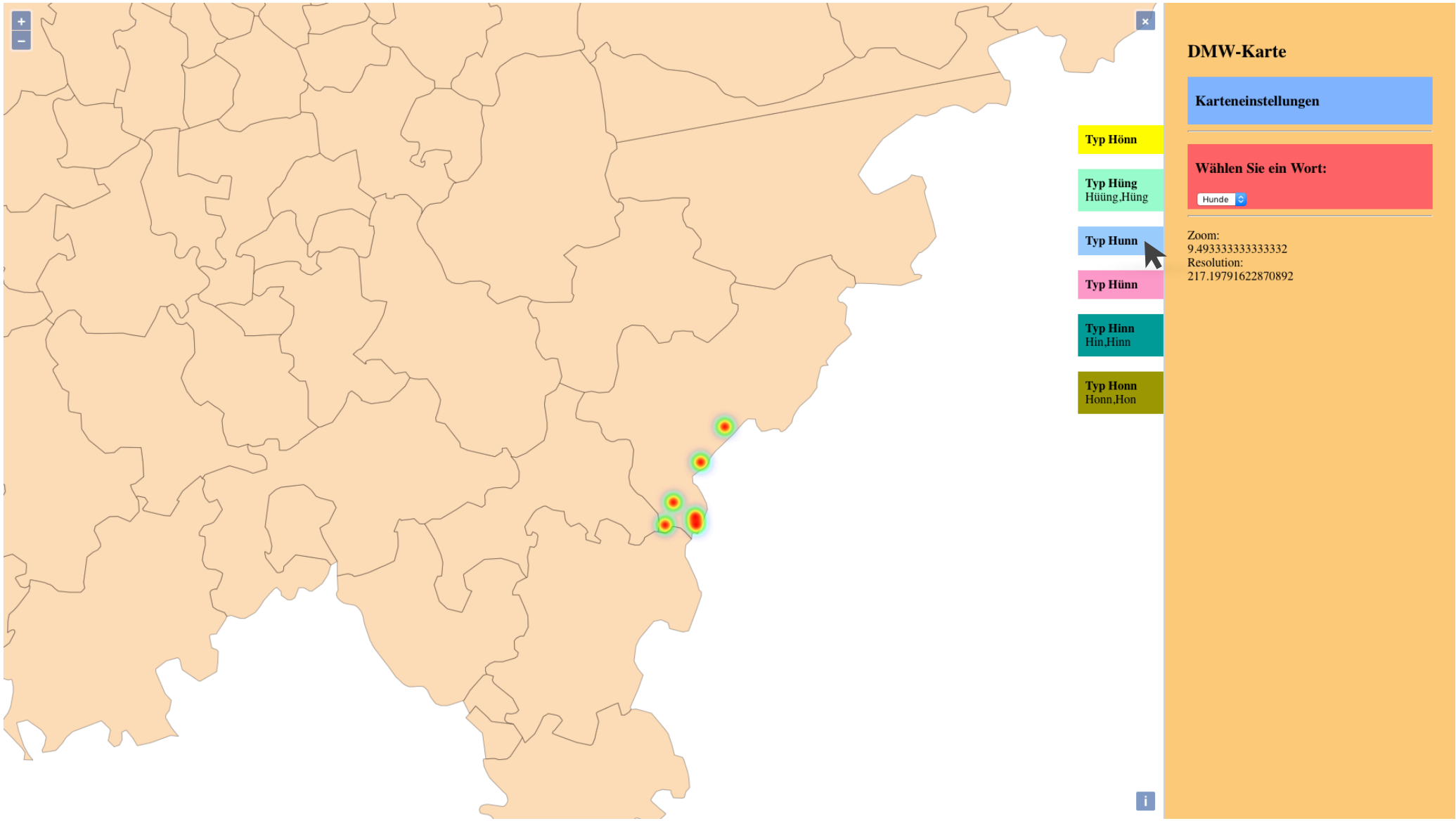


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

indicative presentation: heat map symbols when mouseover type

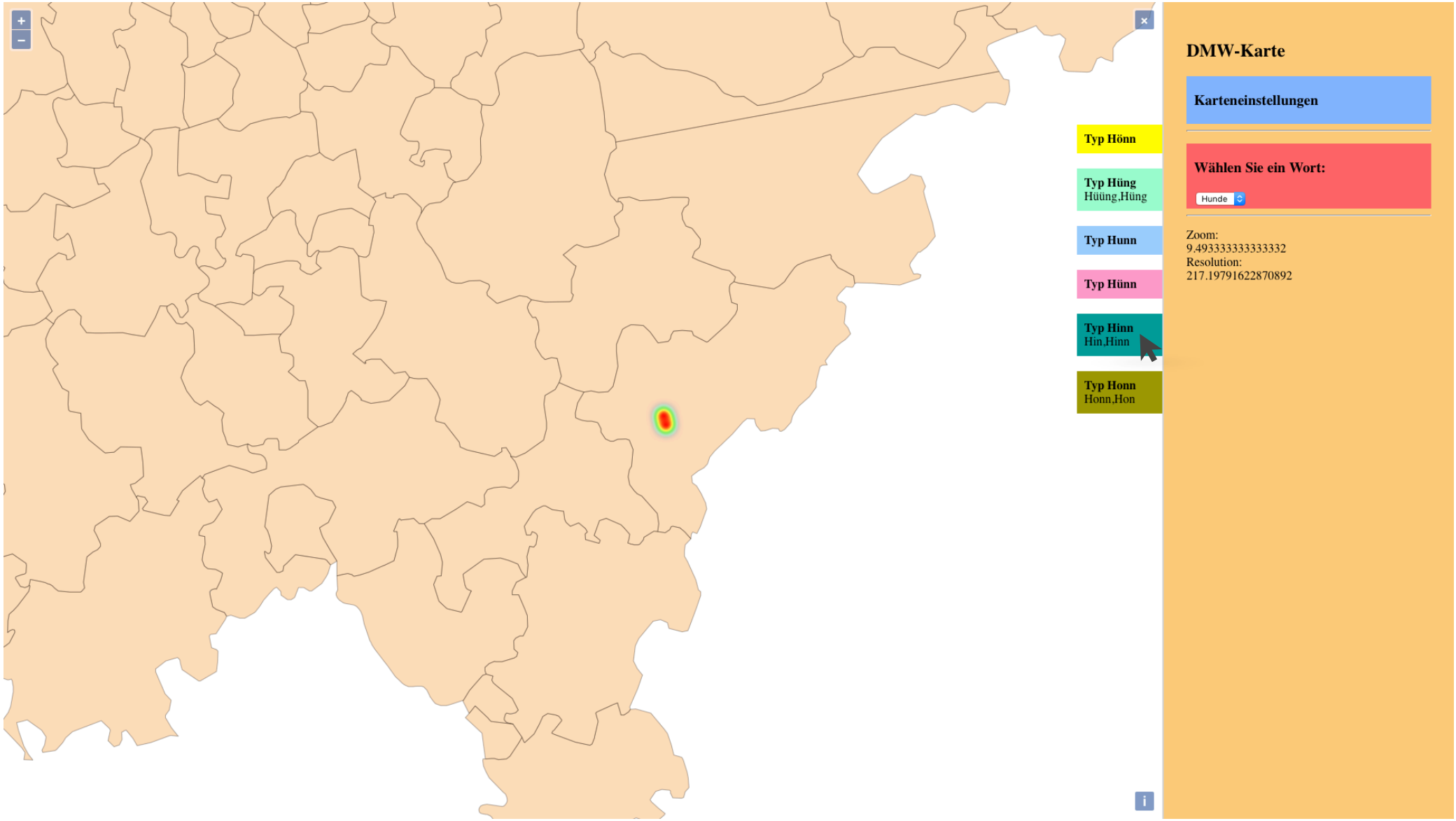


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

indicative presentation: heat map symbols when mouseover type

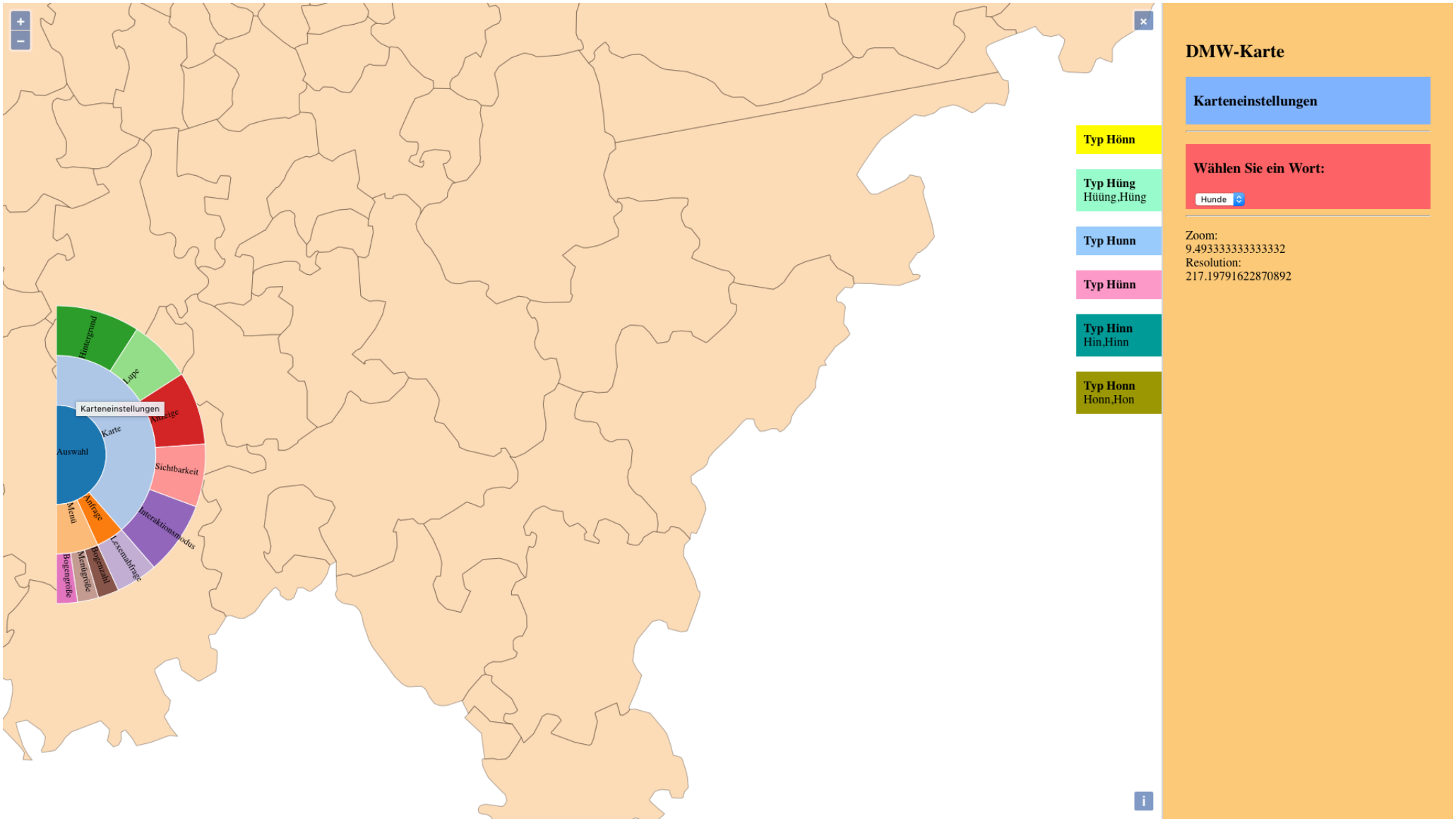


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

on-screen menu („adaptive menu palette“), easily accessible

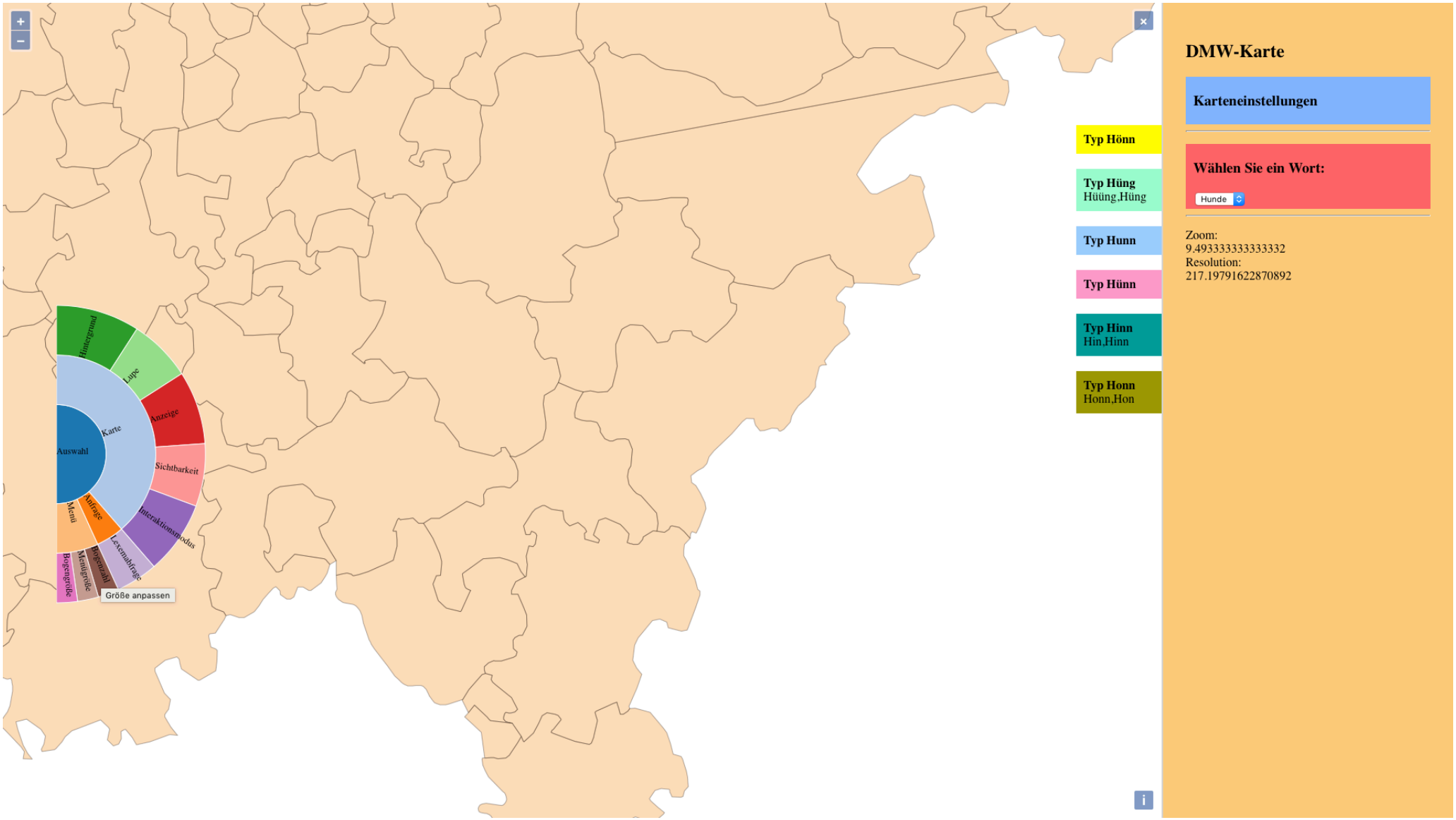


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

on-screen menu („adaptive menu palette“), easily accessible

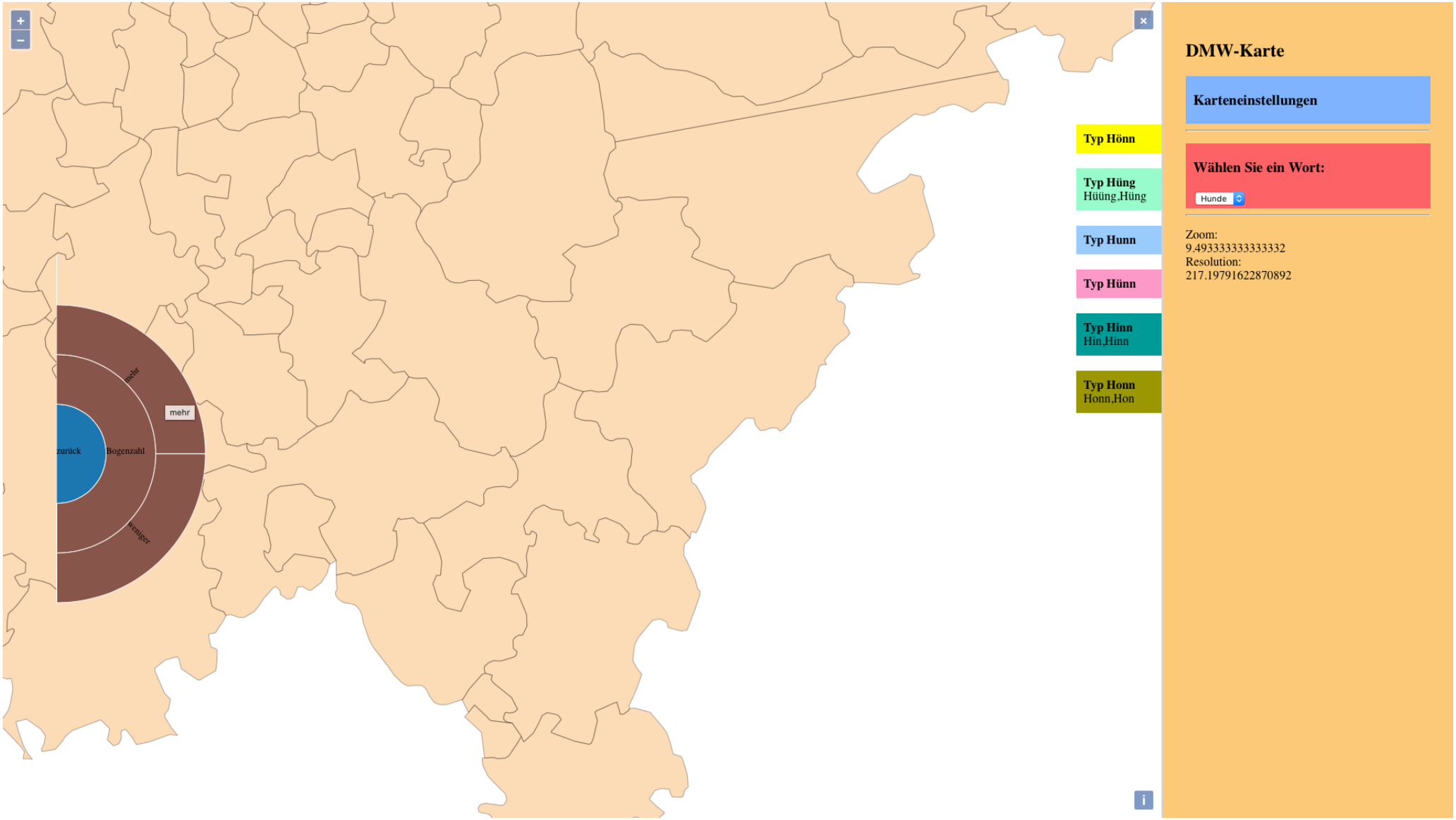


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

on-screen menu („adaptive menu palette“), easily accessible

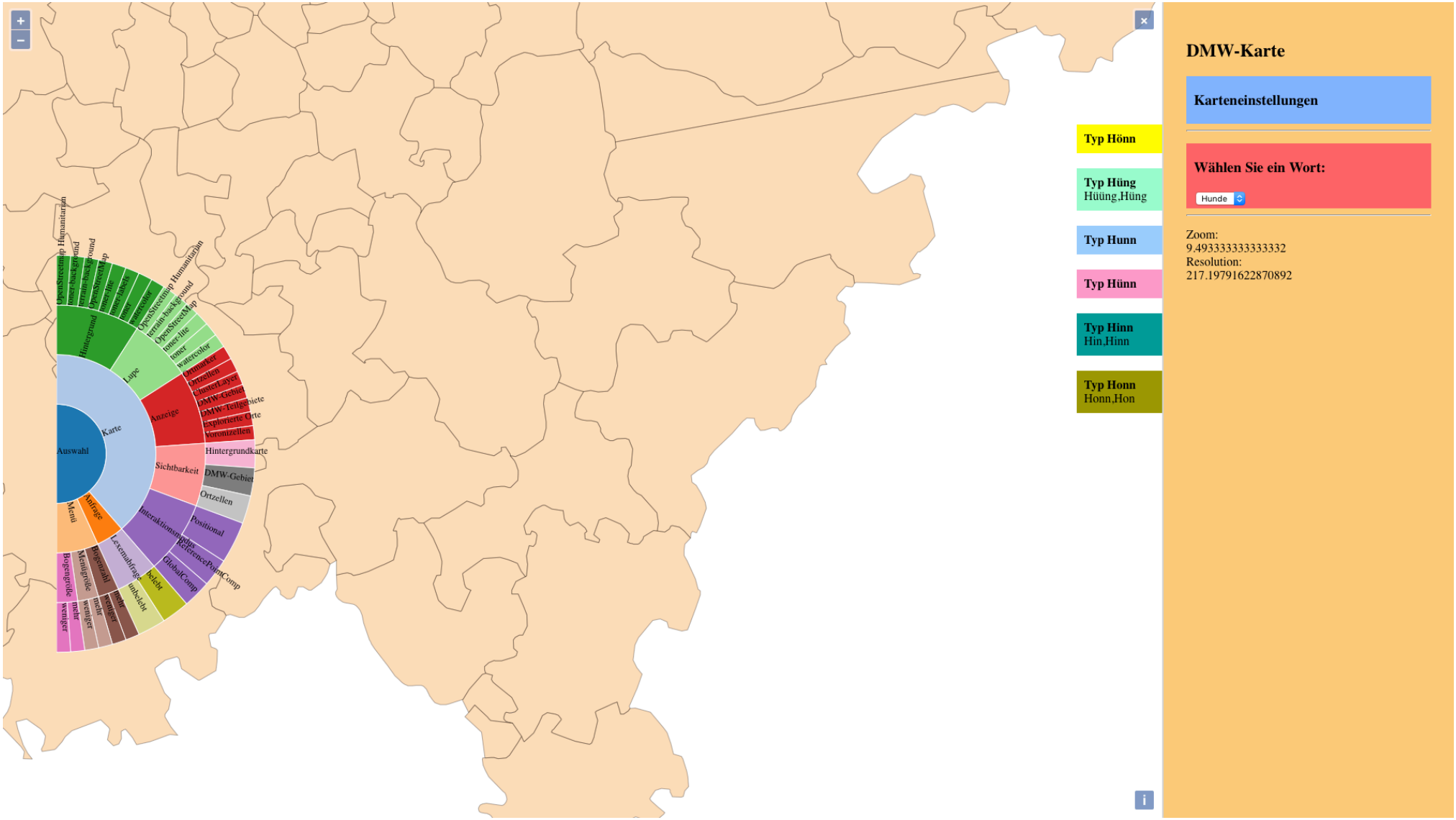


DMW Dialektatlas Mittleres Westdeutschland

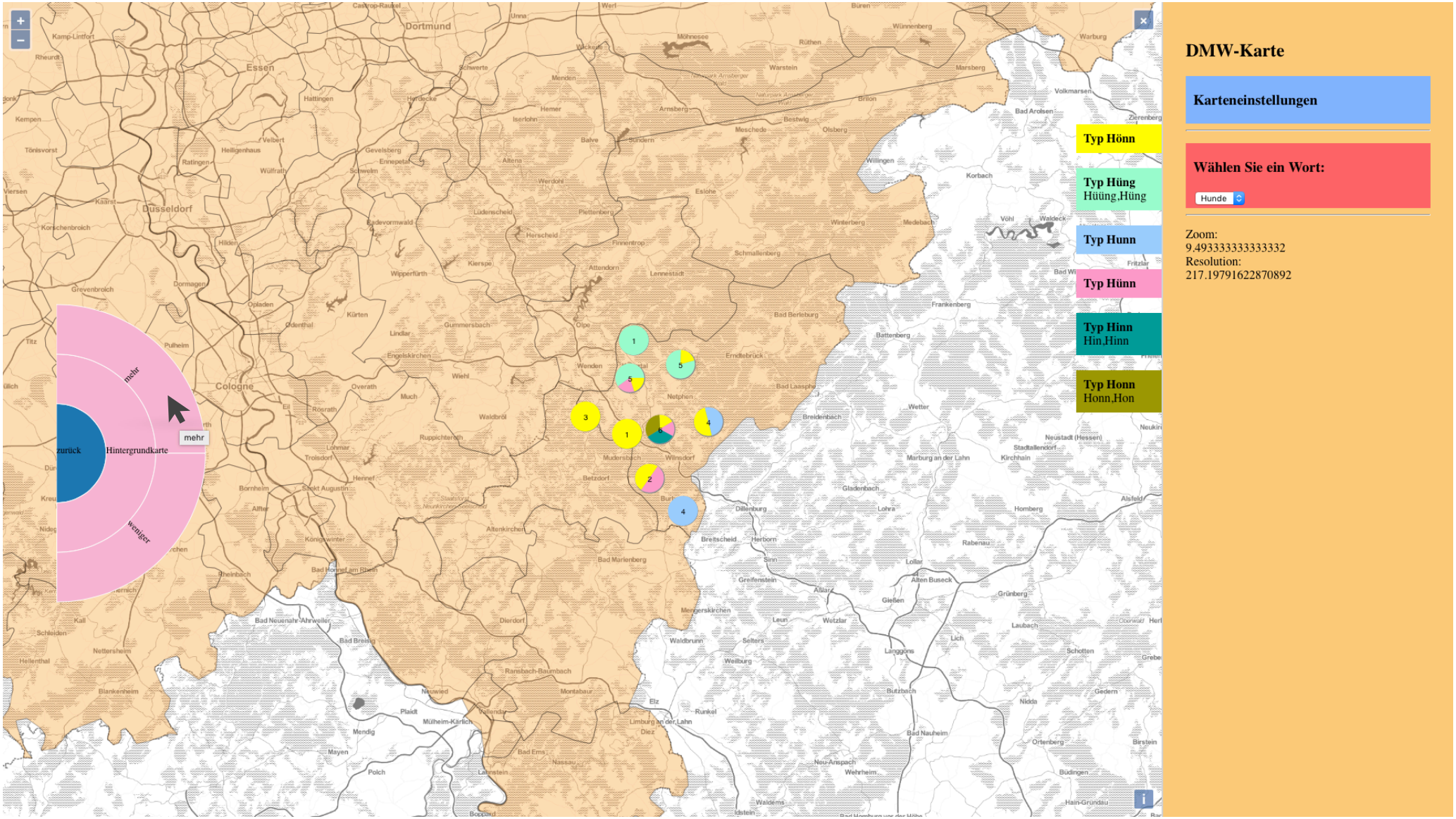
UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

on-screen menu („adaptive menu palette“), easily accessible



quick adaptions

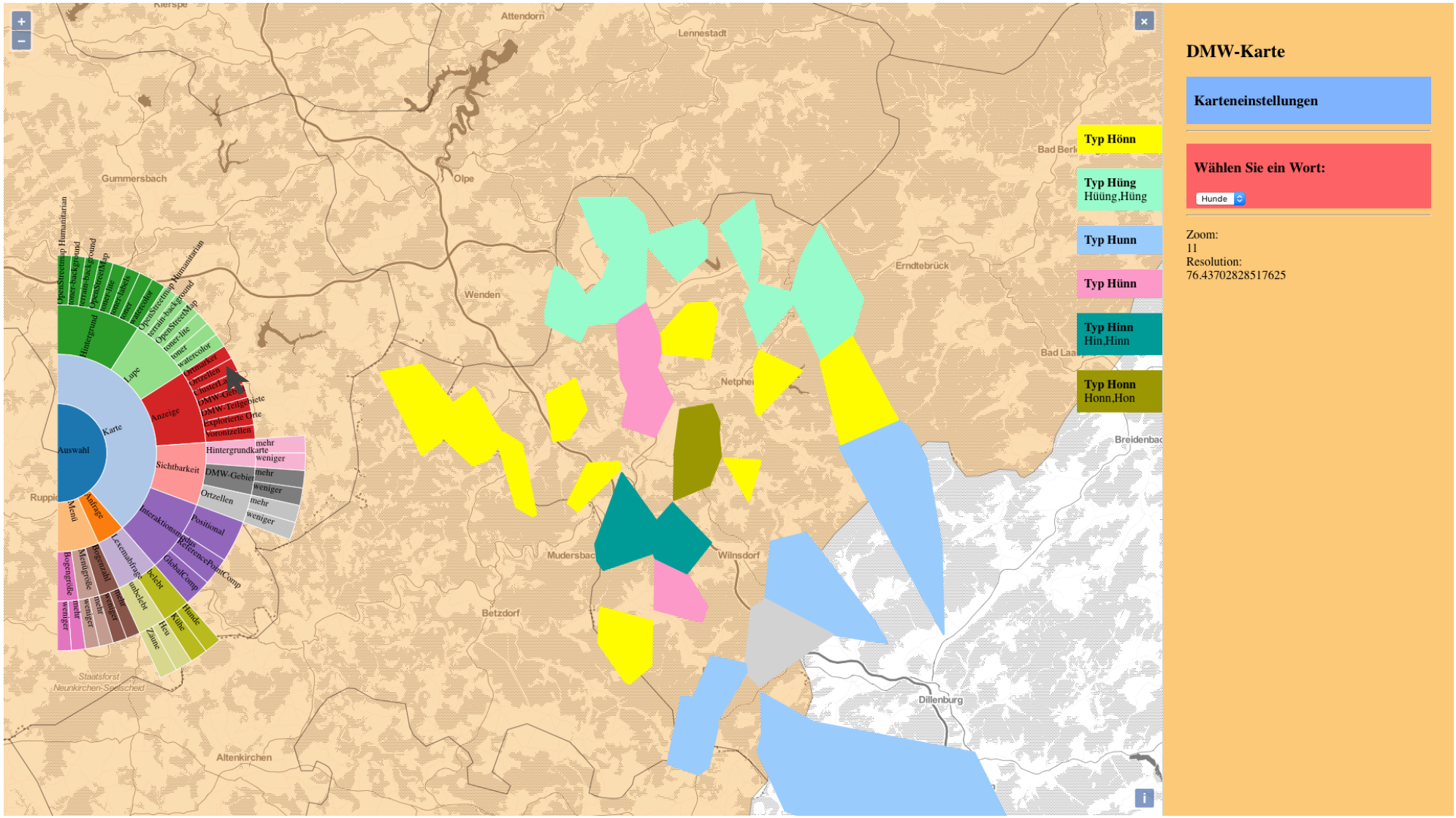


DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

quick adaptions

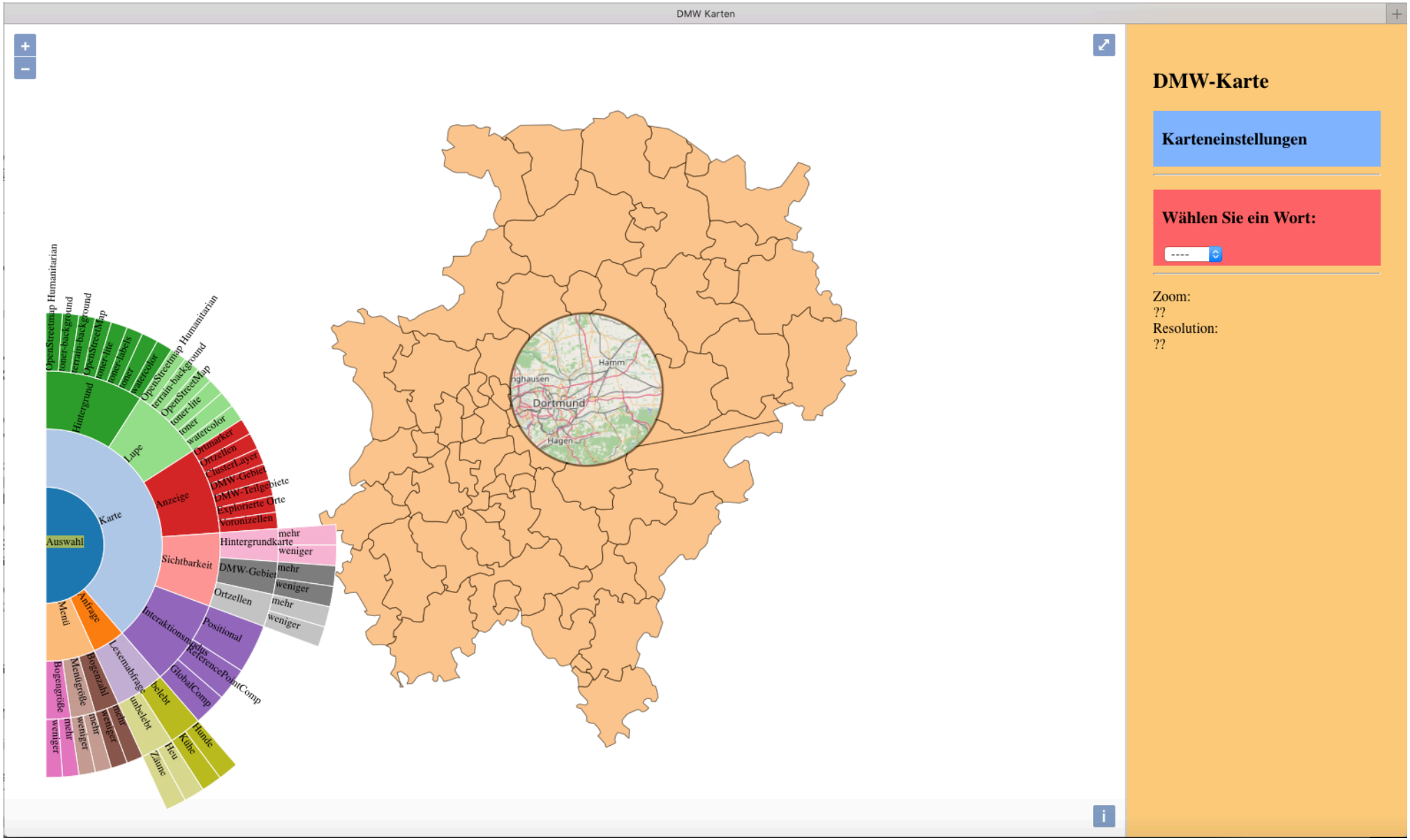


DMW Dialektatlas Mittleres Westdeutschland

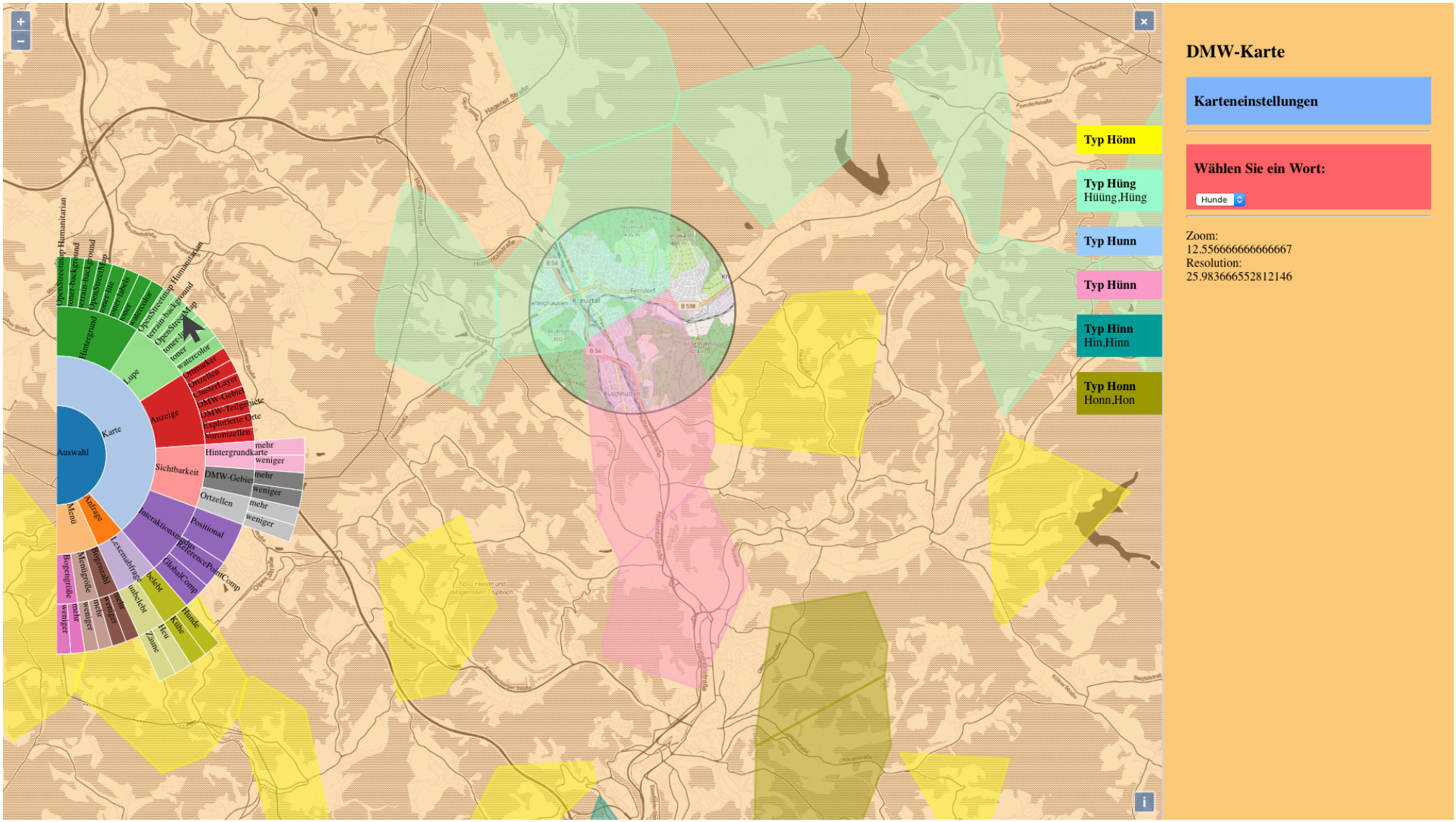
UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KUNST NORDRHEIN-WESTFALEN

Layer spy: selective presentation of different layer



Layer spy: selective presentation of different layer



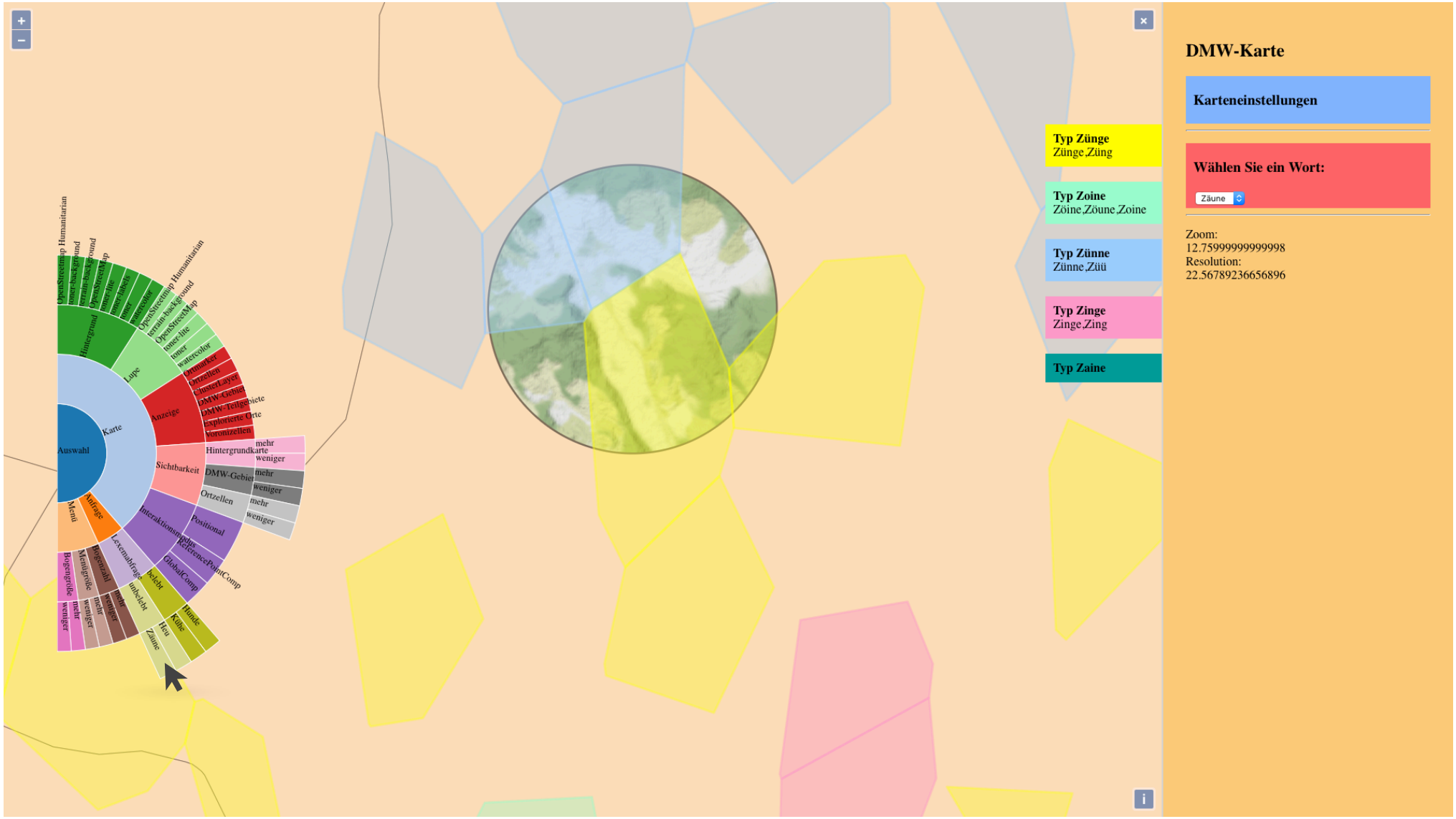
DMW
Dialektatlas Mittleres Westdeutschland

UNION
DER DEUTSCHEN AKADEMIEN
DER WISSENSCHAFTEN

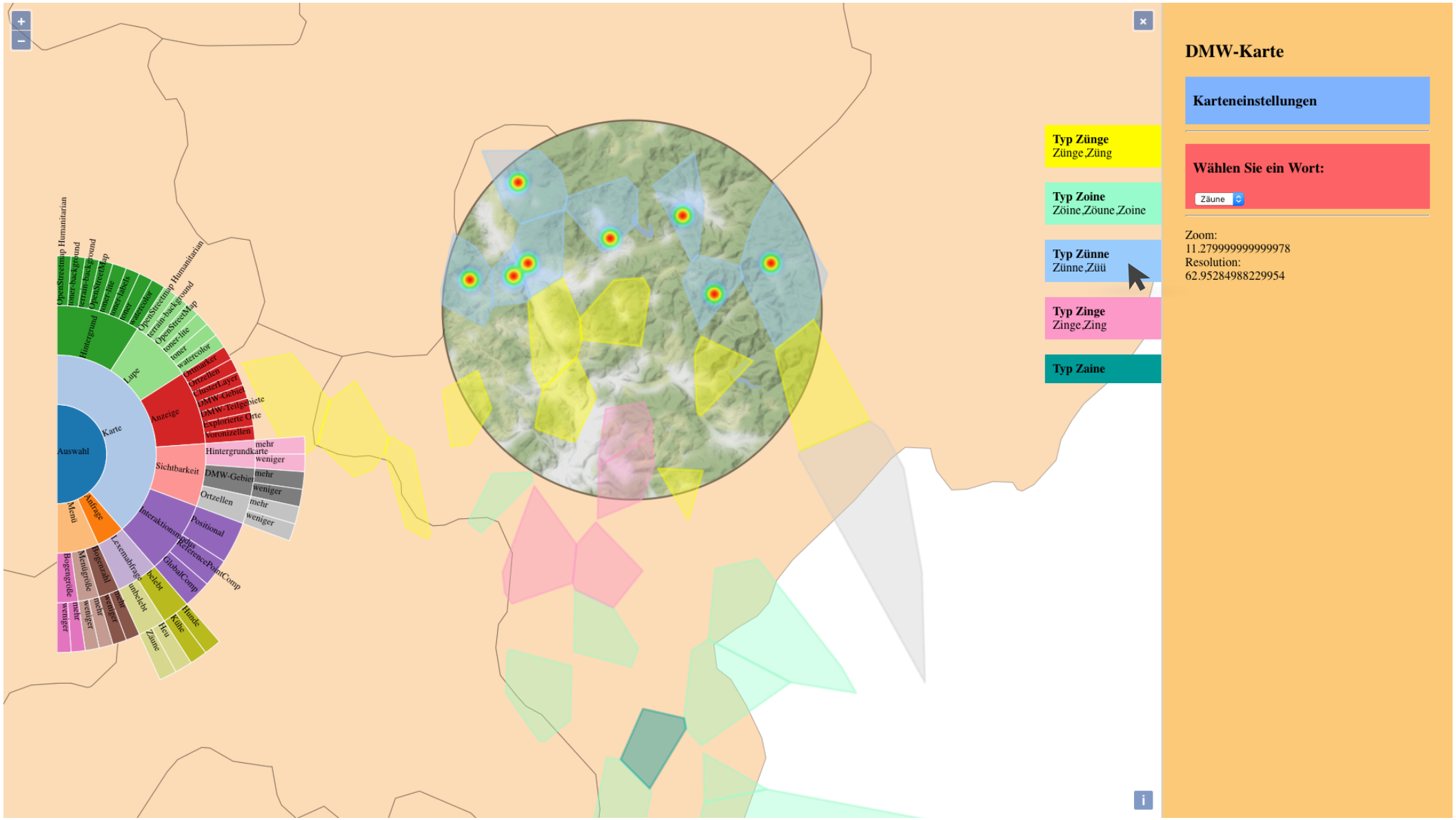
AKADEMIE WISSENSCHAFTEN UND KÜNSTE
NORDRHEIN-WESTFALEN

Layer spy: selective presentation of different layer

changing items on the fly



mouse over some type

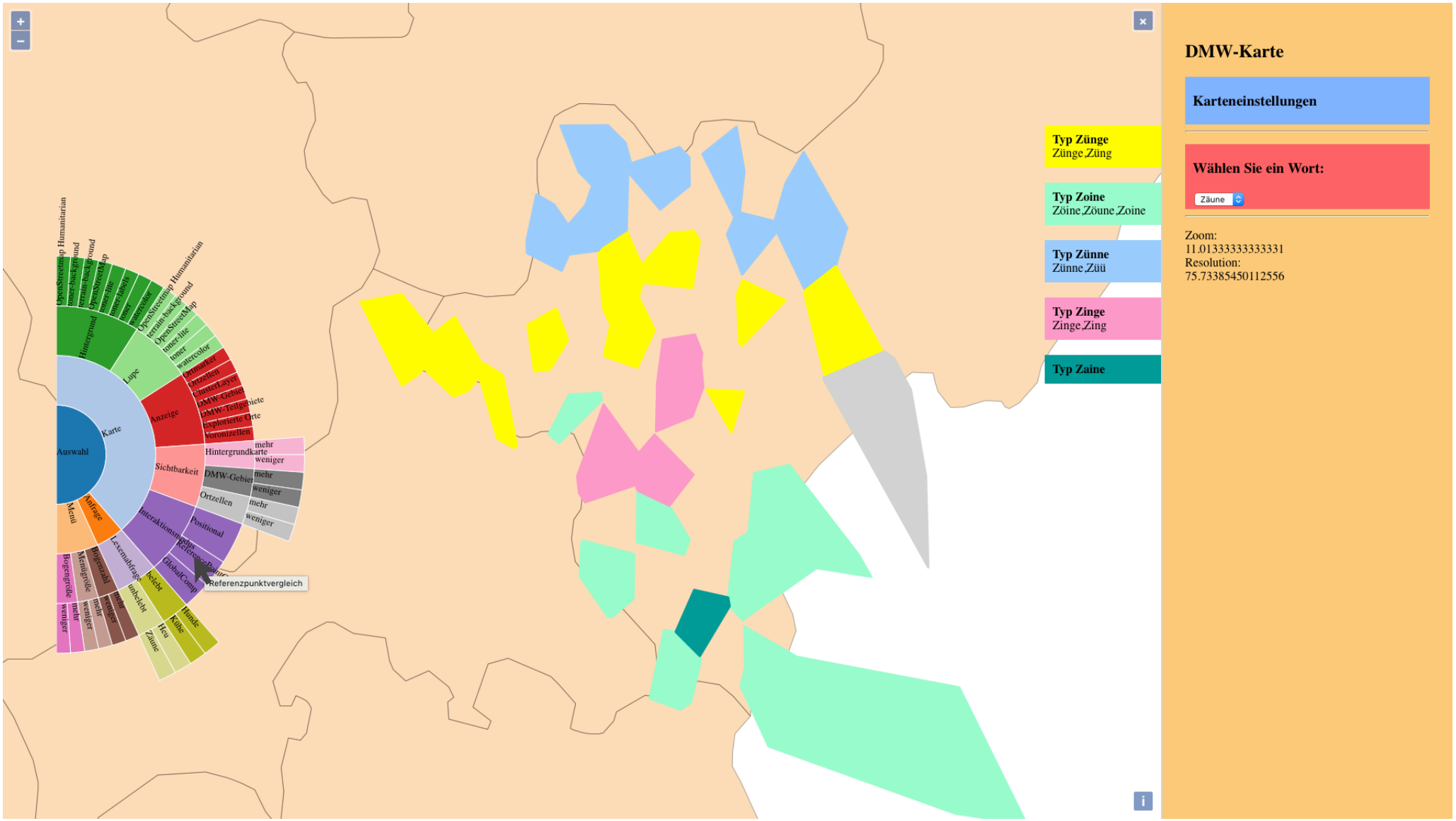


DMW Dialektatlas Mittleres Westdeutschland

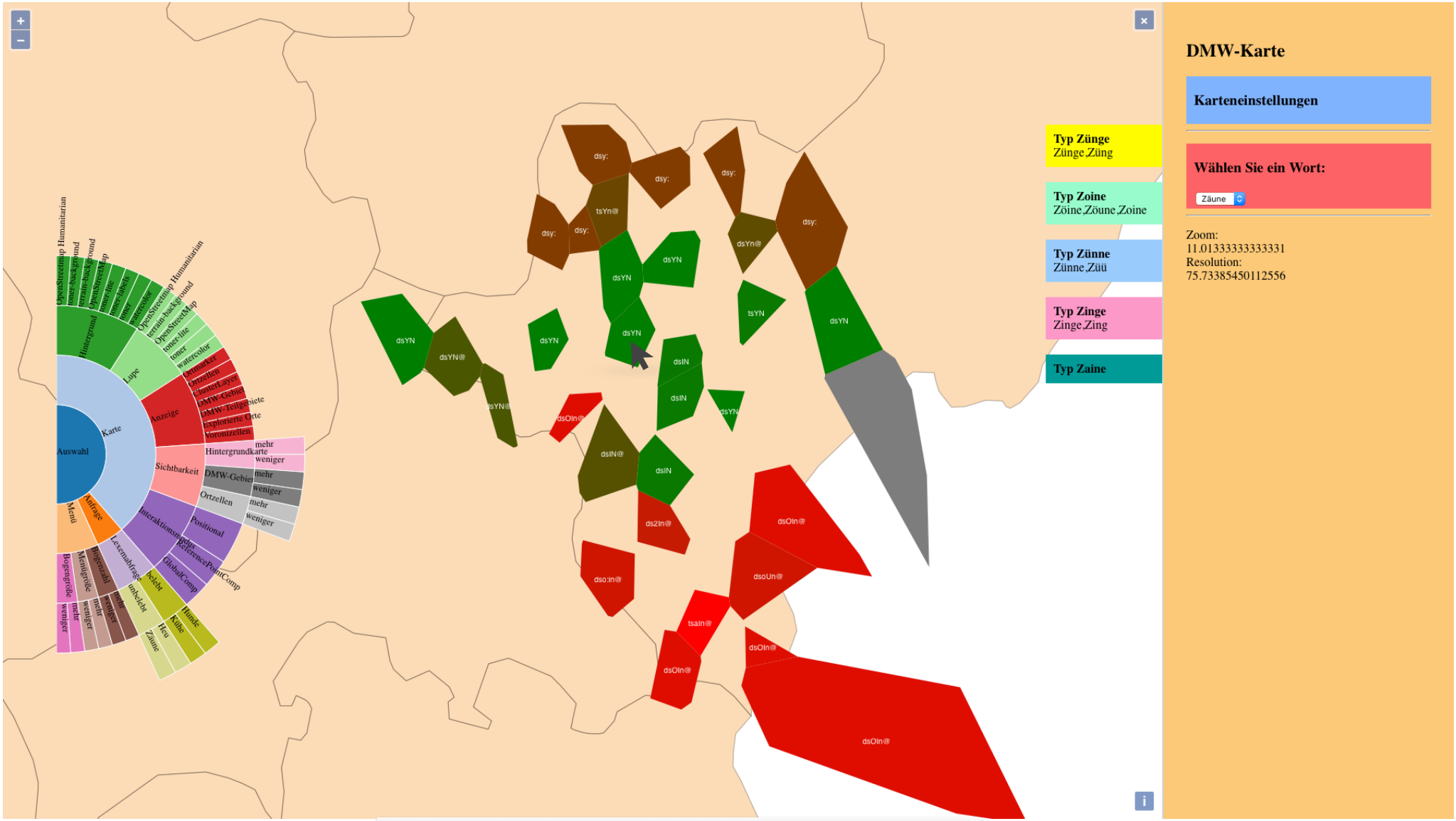
UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

different interaction/presentation mode: reference point comparison

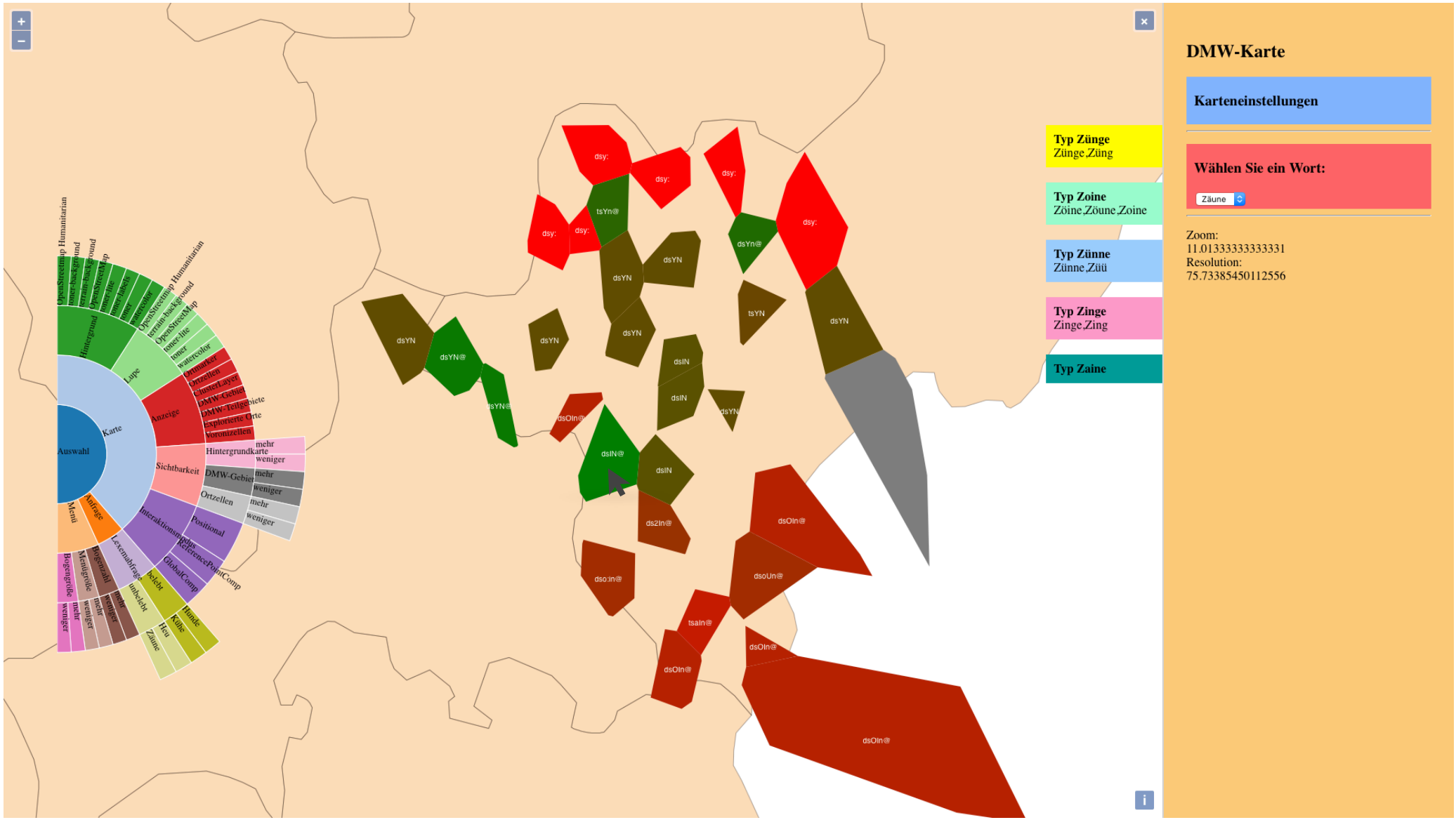


difference maps: phonetic distances relative to some clicked location

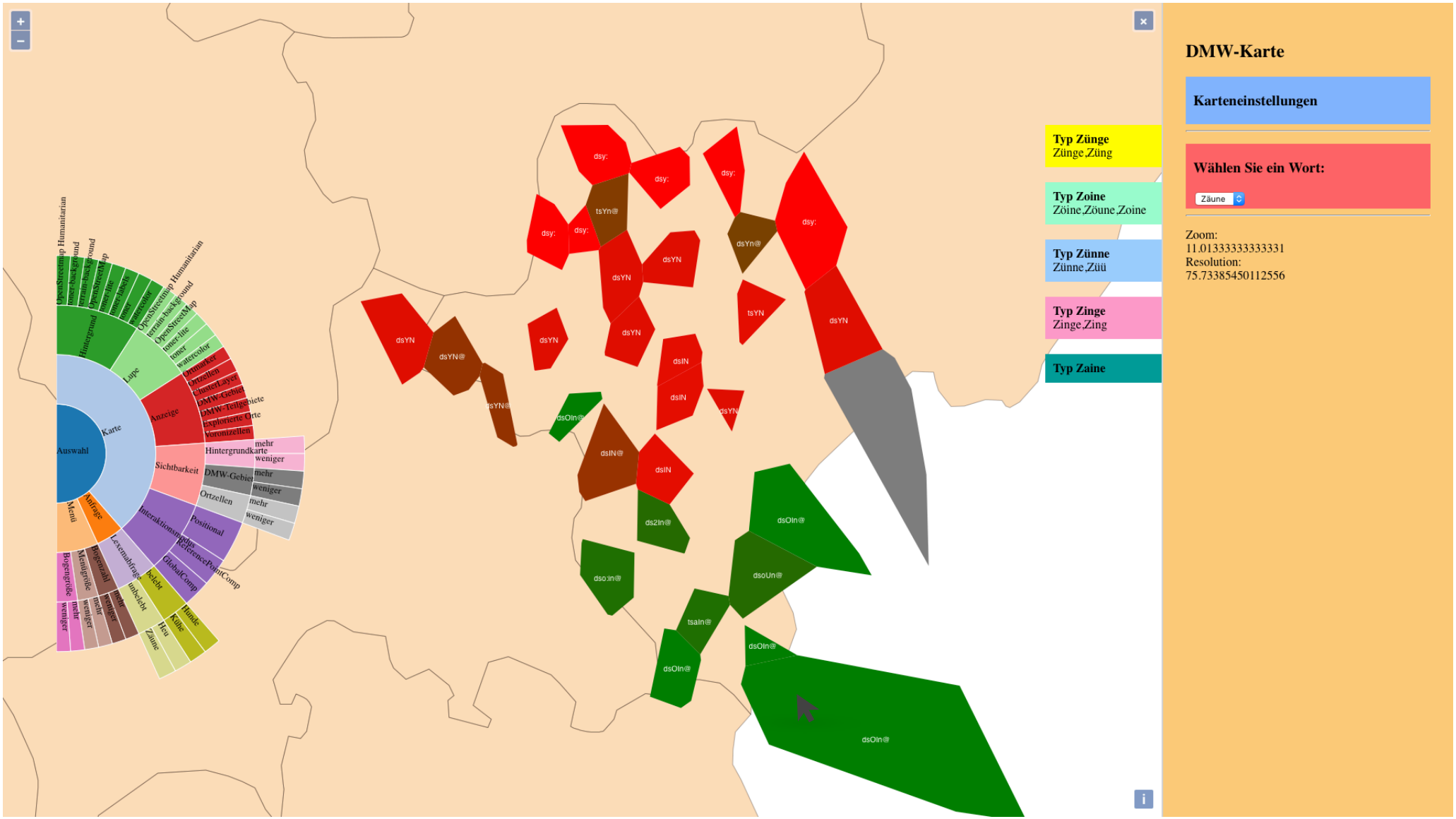


values mapped to green - red continuum

difference maps: phonetic distances relative to some clicked location



difference maps: phonetic distances relative to some clicked location

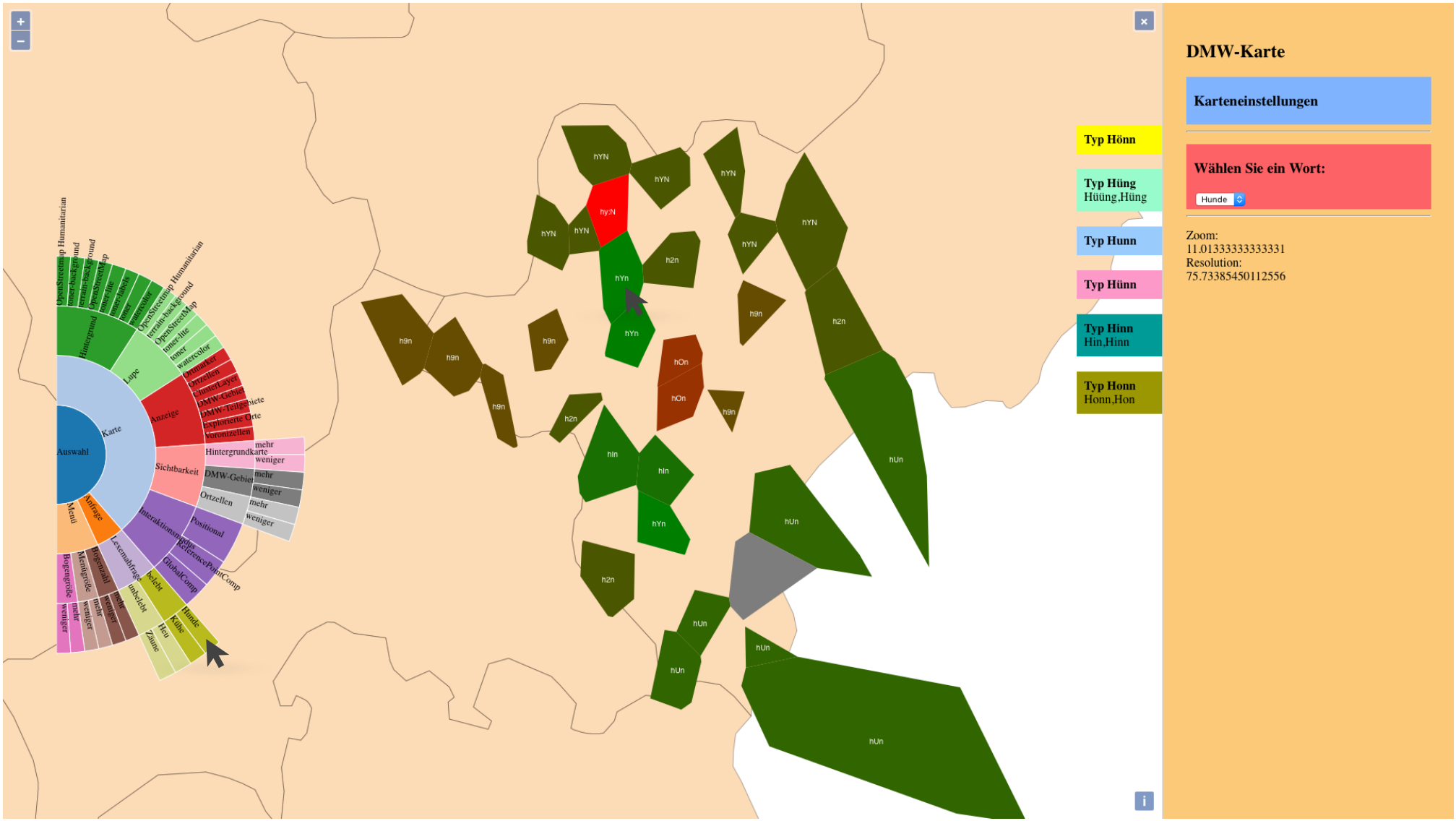


DMW Dialektatlas Mittleres Westdeutschland

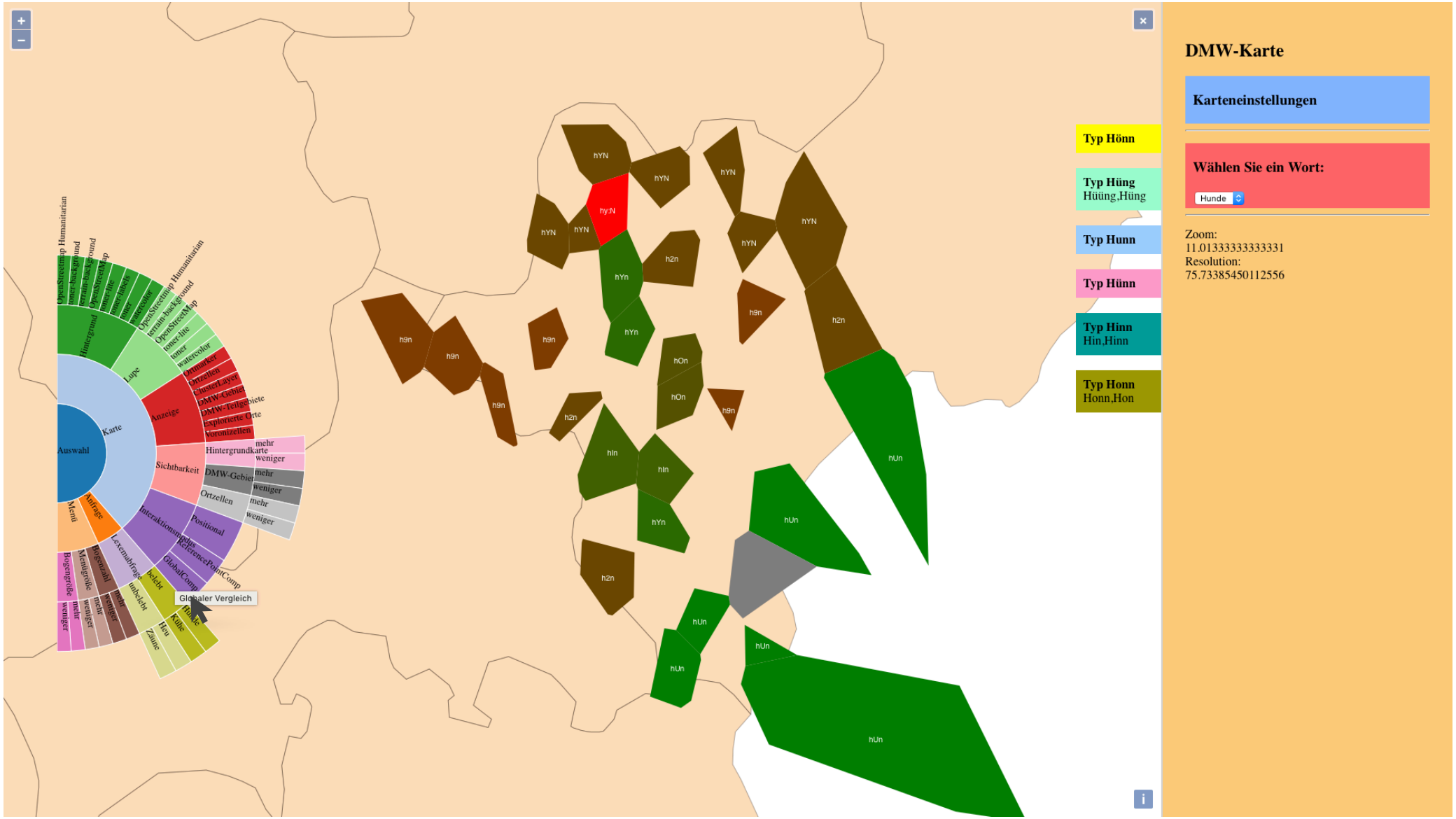
UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

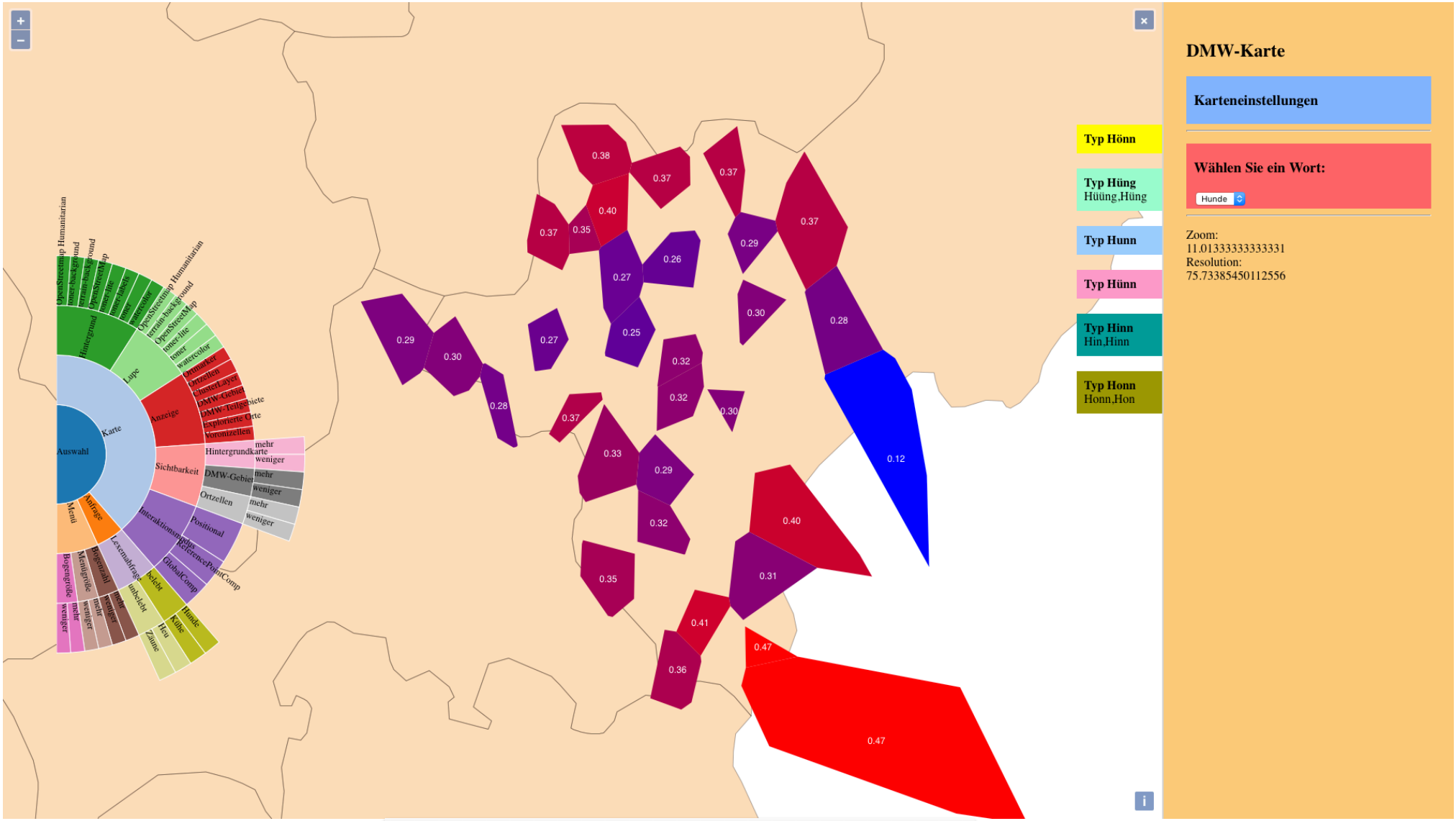
changing items: two easy clicks away



different interaction/presentation mode: global comparison

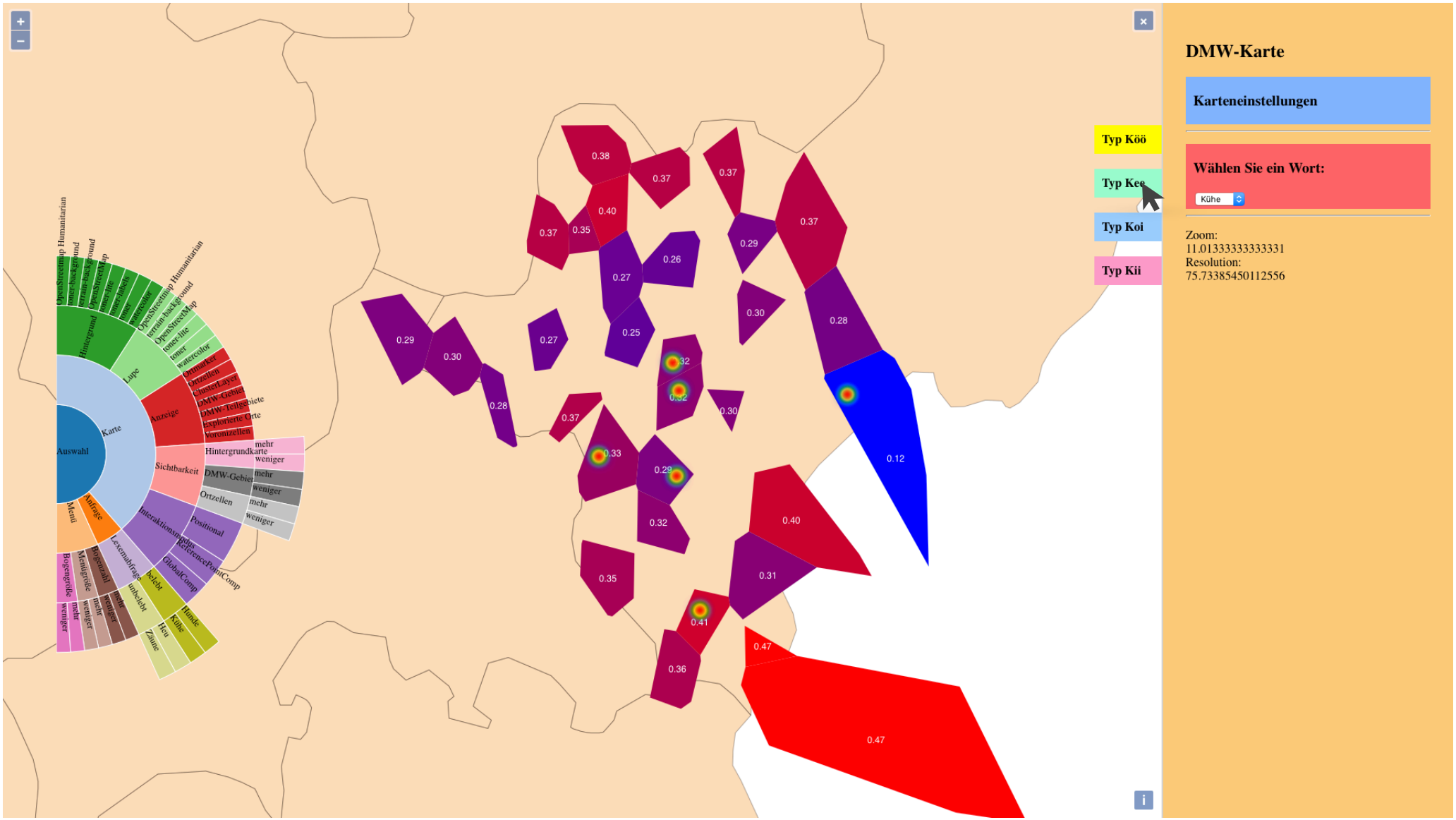


average distance to all other locations

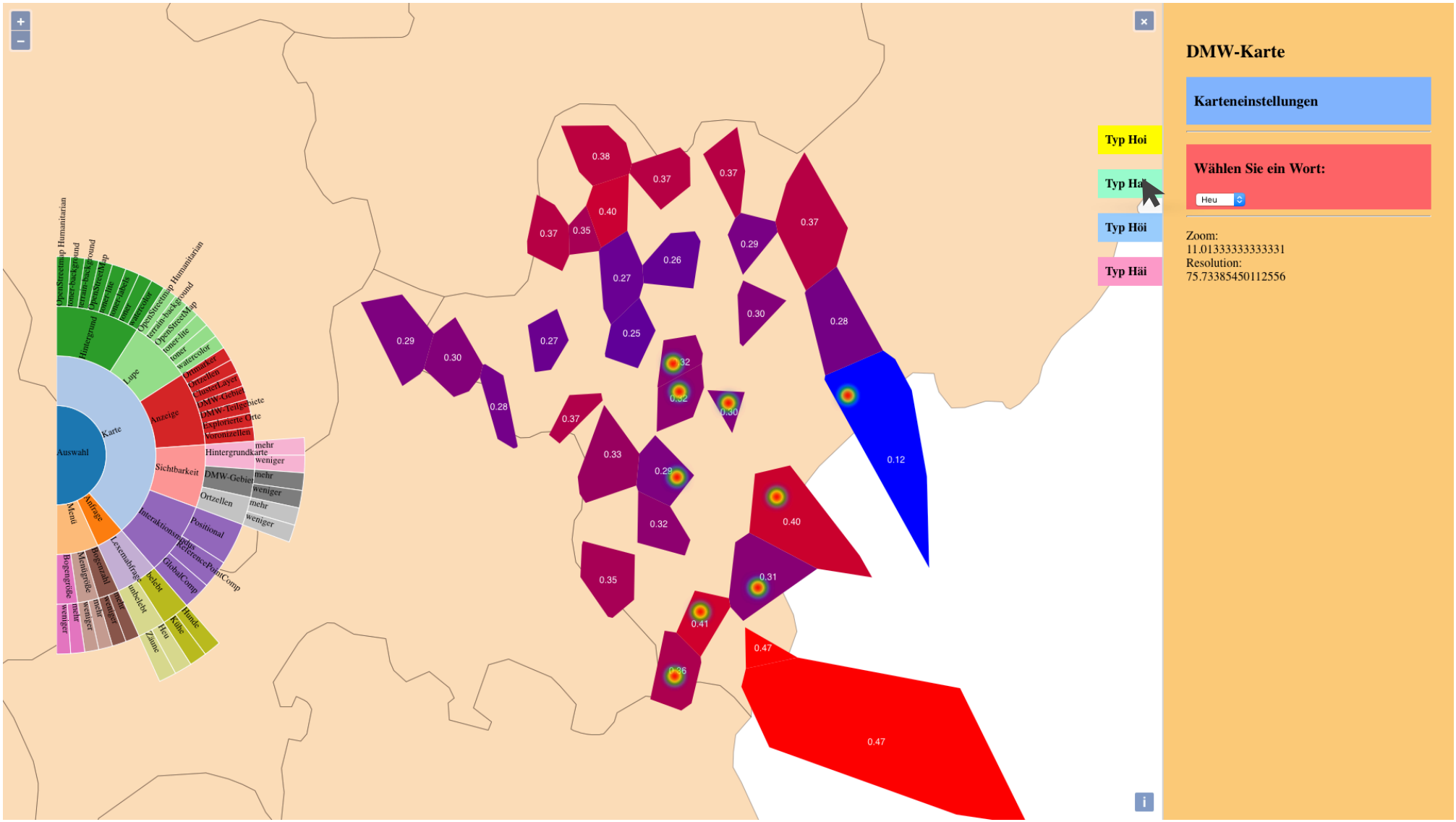


values mapped to blue - red continuum

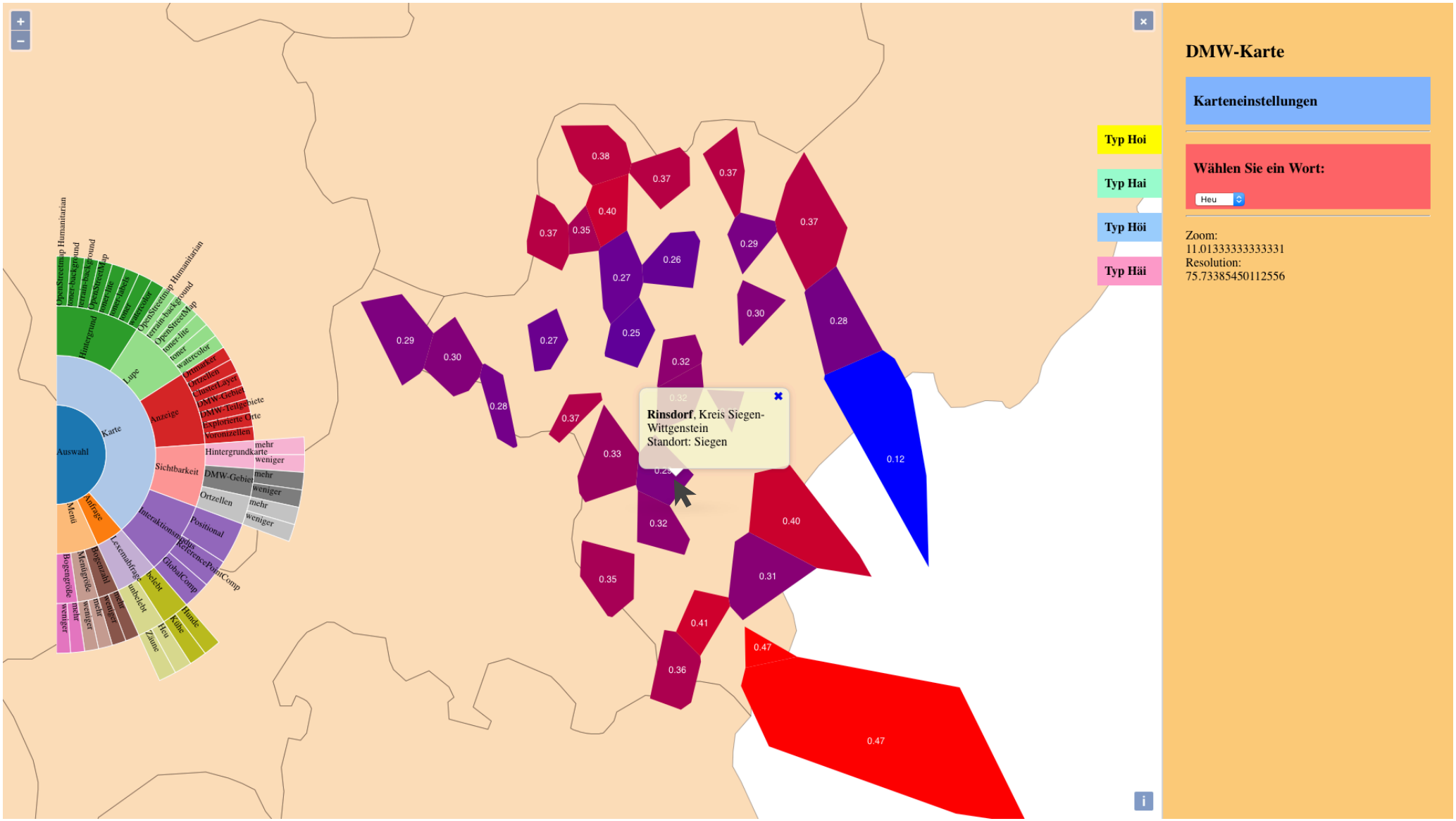
where does type kee appear?



where does type hai appear?



on-click identification of some location



DMW Dialektatlas Mittleres Westdeutschland

UNION DER DEUTSCHEN AKADEMIEN DER WISSENSCHAFTEN

AKADEMIE WISSENSCHAFTEN UND KÜNSTE NORDRHEIN-WESTFALEN

Summary: aspects of **exploratory analysis** in visual dialectology

- highly **selective** presentations
  - on various levels of granularity
  - consisting of different layers of information
  - in different modes
  - not (yet) expert level, lean statistics
- highly **interactive** choice of
  - granularity level, scale/zoom level
  - layers visible
  - presentation mode
  - symbols etc.
- **on-demand** delivery of information
  - by clicking (on map, in menu)
  - by mouseover (symbols)
- **adaptable on-screen menu**
  - quickly accessible
  - easy switching between options
- web **technology** used:
  - Html5, CSS, jQuery, Openlayers 4, D3

Thank you for your attention!