



VGIscience – Geovisuelle Analyse von Location-based Social Media Daten

Prof. Dr.-Ing. habil. Dirk Burghardt, Institut für Kartographie



Gliederung

1. Interpretation nutzergenerierter, raumbezogener Daten
 - Begriff: Volunteered Geographic Information
 - Analyse raumbezogener Daten von Social-Media-Plattformen
 - DFG-Schwerpunktprogramm VGIscience
2. Ableitung von Mehrfachrepräsentationen
 - Ableitungsschema für Mehrfachrepräsentationen
 - Anwendungsbeispiele für
 - interne Ableitung (Tag Maps) und
 - externe Zuordnung (Micro Diagramme)

Begriffsbildung

- VGI -Volunteered Geographic Information (Goodchild, 2007)
 - eingeführt von Michael Goodchild (2007)
 - Spezialfall nutzergenerierter Daten (engl. user generated content / UGC), mit direktem oder indirektem Raumbezug
 - Mehrwert besteht u.a. in der freien Verfügbarkeit
- Citizen Science (dt. Bürgerforschung)
 - Projekte engagierter, interessierter Personen
 - Konzept der „Humans as Sensors“ – Nutzung kostengünstiger Sensoren für verschiedenste Aufgaben
 - Datenerzeugung erfolgt aktiv und zweckbezogen



Raumbezogene Daten von Social-Media-Plattformen (location-based social media data)

- soziale Netzwerke bieten Plattformen zum wechselseitigen Austausch von Meinungen, Erfahrungen und Informationen
 - Ereignisse finden an bestimmten Plätzen statt - Sportveranstaltungen, Hochwasser, Staus, Demo's, Kommentare zu Restaurants, ...
 - Erzeugung der raumbezogene Daten eher als Nebeneffekt (passiv)
- Charakteristika der Daten
 - große Datenmengen bei starker Heterogenität
 - Aktualität, Echtzeit, Streaming
 - semantische interpretiert – nutzerspezifisch
 - Datenschutz, Privatsphäre

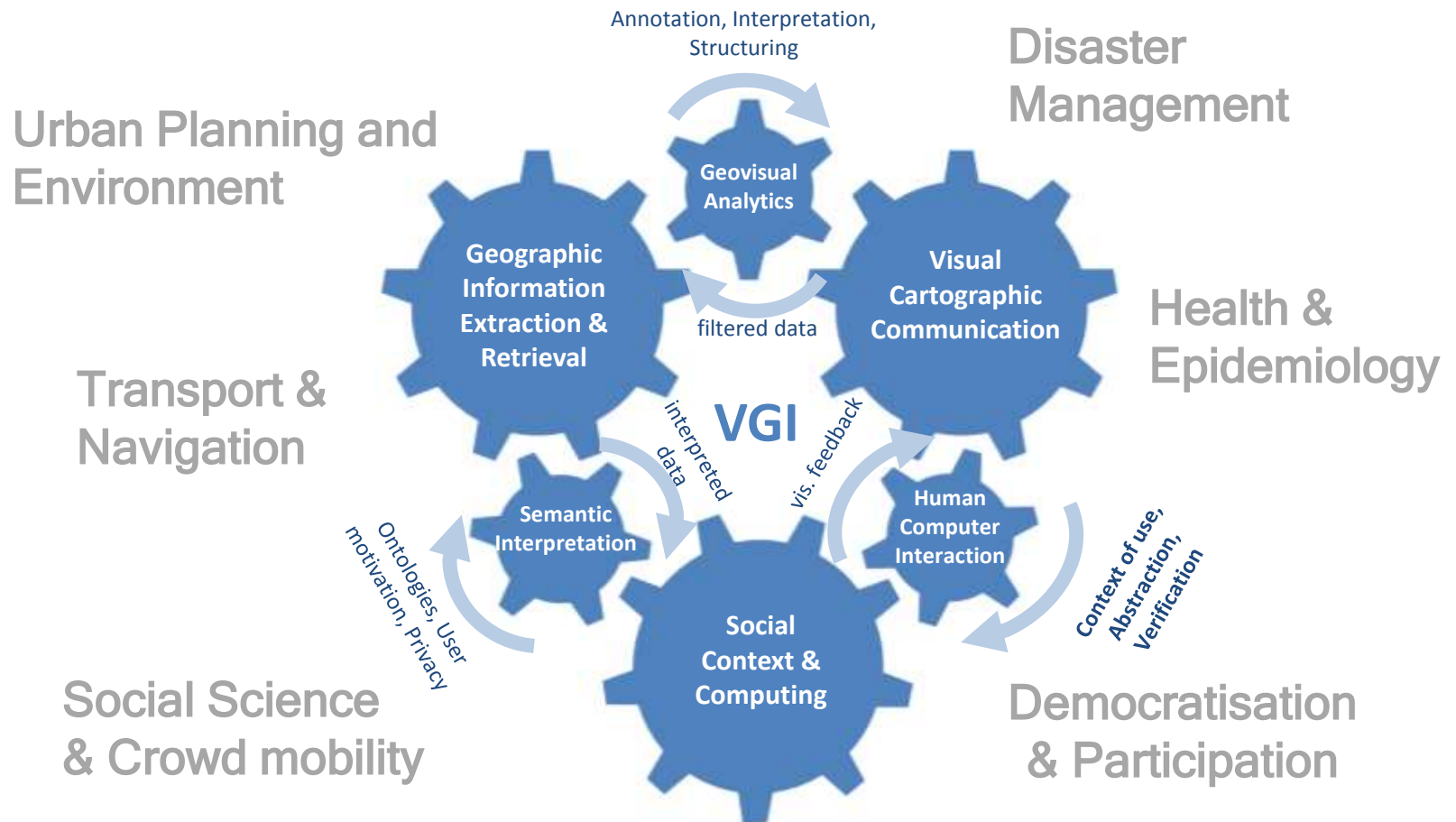


Herausforderung durch VGI in Verbindung mit “Big Data” - 4 x V (Laney, 2001; Robinson et al. 2017)

- **volume**
 - bezieht sich auf die Datengröße und variiert je nach Disziplin erheblich (von Millionen Punkten in Bewegungsdaten bis zu Petabyte in Bildquellen)
- **velocity**
 - meint einerseits die Geschwindigkeit, mit der VGI erzeugt werden kann (kontinuierliche Datenströme) und andererseits den Zeitaufwand für deren Analyse (Echtzeit)
- **variety**
 - bezieht sich auf die Heterogenität der Daten hinsichtlich Formaten, Grad der Strukturierung, räumlicher-zeitlicher Auflösung etc.
- **veracity**
 - Herausforderungen hinsichtlich Qualität, Vertrauenswürdigkeit, Unsicherheit, Subjektivität und Privatsphäre

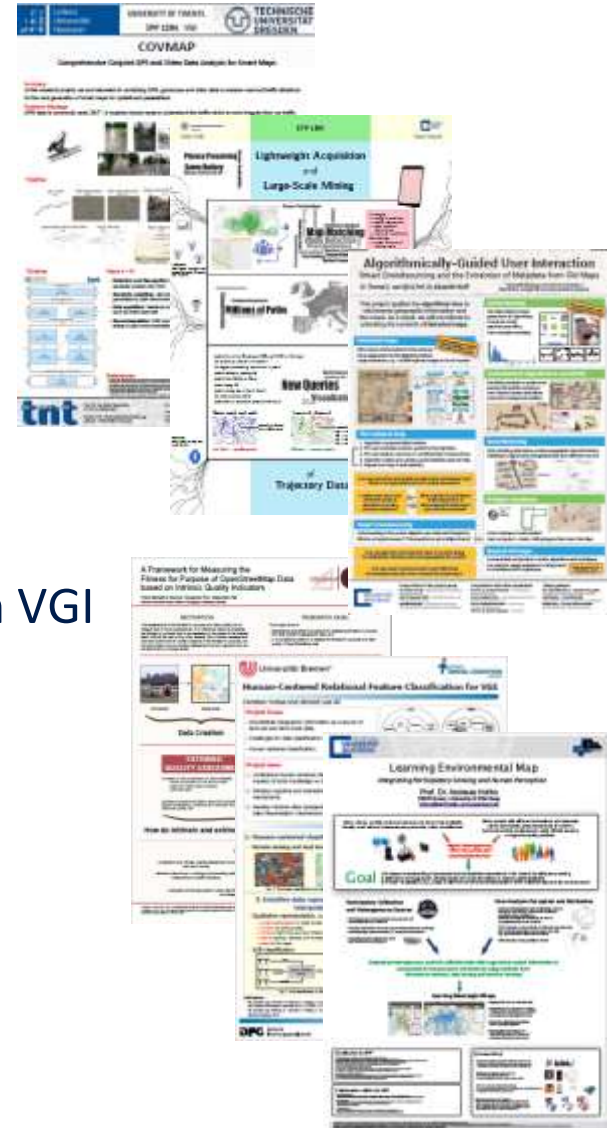
DFG-Schwerpunktprogramm VGIscience

<https://vgiscience.org/>



Forschungsprojekte in VGIscience

- aktive Beteiligung und Erfassung von Geodaten (COVMAP, LearnEnviMaps, topikos, TrajectoryVGI, OldMapsVGI)
 - gemeinsame GPS- und Video-Sammlung für Verkehrsmanagement
 - partizipative Erfassung von Umweltdaten
 - Extraktion von Metadaten aus alten Karten
- Qualität /-sicherstellung und Verbesserung von VGI (QualityOSM, HC-VGI, LearnEnviMaps)
 - Qualität von OpenStreetMap für verschiedene Anwendungszwecke
 - Optimierung von Sensormessungen durch erweiterte Kalibriermechanismen



Forschungsprojekte in VGIscience

- Geovisualisierung und Datenanalyse
(VA4VGI, EVA-VGI, ENAP, SpatialCorrelationVGI, UncertaintyTrustVA, VaguePlaceVGI)
 - Echtzeit-Event-Bewertung auf Basis von LBSM
 - Analyse der Reaktionen von Menschen
 - räumliche Korrelationen in LBSM
 - Interpretation von vagen Ortsbeschreibungen
- Soziale Aspekte und menschliche Wahrnehmung
(MotivationHumanVGI, UncertaintyTrustVA, LandmarksVGI, EVA-VGI)
 - Motivation und Teilnahme im Bereich humanitärer Hilfsprojekte
 - Wahrnehmung von Landmarken in VGI-basierten Karten
 - Vertrauenswürdigkeit und Privatsphäre

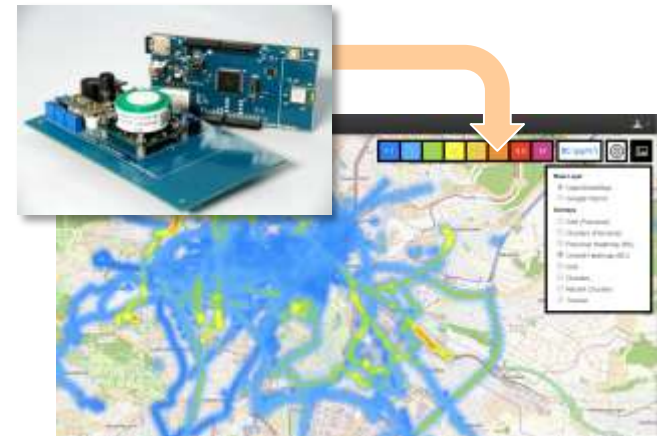


Anwendungsgebiete in VGIscience

- Katastrophenmanagement, Gefahrenreaktion und Hochwasserschadensmodellierung (COVMAP, QualityOSM, UncertaintyTrustVA, MotivationHumanVGI, ENAP)
- Stadt- und Landschaftsplanung, Umweltmonitoring, Verkehrsmanagement (LearnEnviMaps, OldMapsVGI, QualityOSM, EVA-VGI)
- Analyse menschlicher Aktivitäten und von Personenmobilität (TrajectoryVGI, UncertaintyTrustVA, EVA-VGI, VA4VGI)
- (Indoor-)Navigation und Orientierung (topikos, QualityOSM, LandmarksVGI)
- VGI und Landbedeckungskartierung (HC-VGI)



Relevance Feedback (ENAP)



Sensorbox zur Ableitung intelligenter Luftverschmutzungskarten (LearnEnviMaps)

VGIScience Summer School

(11.-15. Sep. 2017 in Dresden)



- 25 PhD und 12 Postdoc (Deutschland, Schweiz, Österreich, Japan, Indien)
- Vision talk (Prof. Alan MacEachren), Lectures und praktische Übungen
- projektbezogene Gruppenarbeit (institutionsübergreifend)

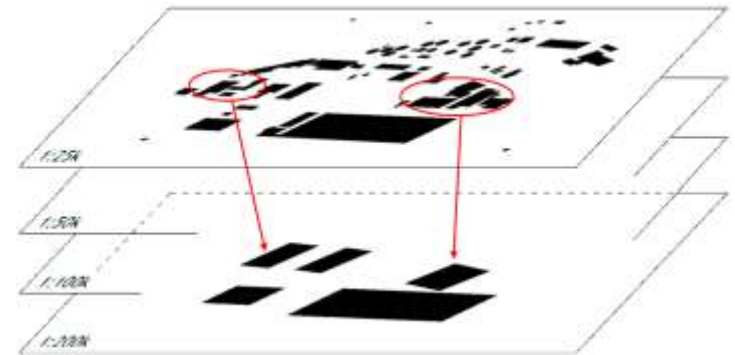


Gliederung

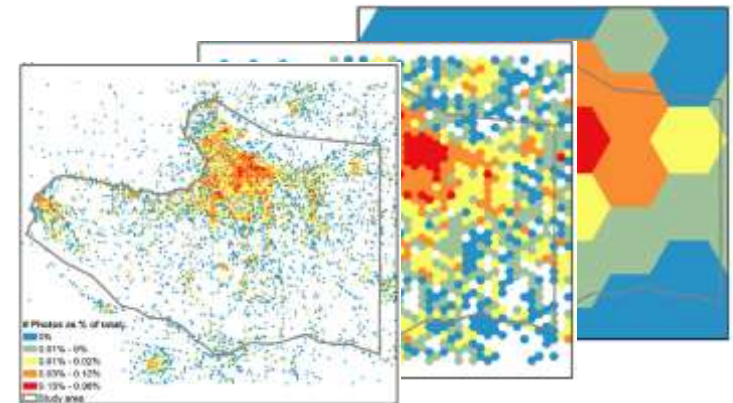
1. Interpretation nutzergenerierter, raumbezogener Daten
 - Begriff: Volunteered Geographic Information
 - Analyse raumbezogener Daten von Social-Media-Plattformen
 - DFG-Schwerpunktprogramm VGIscience
2. Mehrfachrepräsentationen von LBSM
 - Ableitungsschema für Mehrfachrepräsentationen
 - Anwendungsbeispiele für
 - interne Ableitung (Tag Maps) und
 - externe Zuordnung (Micro Diagramme)

Terminologie – “Multiple representation” vs. “Multi-scale views”

- Begriff: “multiple representations”
 - wurde bisher im Kontext topographischer Kartenproduktion für maßstabsabhängige Speicherung von Objektgeometrien des gleichen Realweltobjektes verwendet (Sarjakoski, 2007)
 - explizite Verknüpfung von Repräsentation zur Unterstützung von Updates, Konsistenzüberprüfungen und kontinuierlichem Zoom
- angewendet auf Analyse von LBSM-Daten wird auch “multi-scale view” verwendet
 - Fokus liegt auf Analyse variierender Muster anstelle von Verknüpfung einzelner Objekte



Verlinkung innerhalb MRDB (Cecconi, 2003)



Multi-scale Ansatz angewendet auf räumlich annotierte Photos (Feik and Robertson, 2015)

Ableitung von Mehrfachrepräsentation

Interne Ableitung
aus den Daten

Hierarchisches Clustering
(Dendrogramm)

Tag Maps

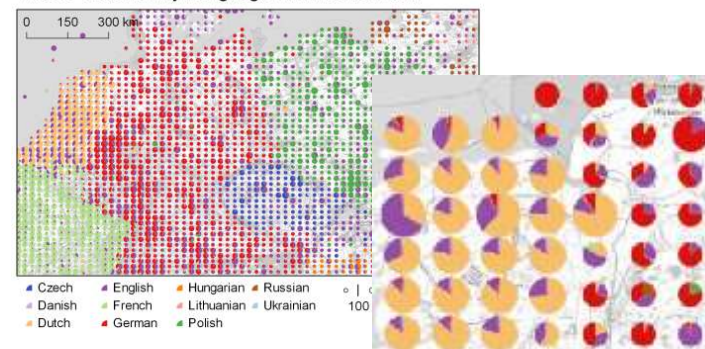


Externe Zuordnung
von Bezugseinheiten

Hierarchische Tessellation
mittels Gitter, Quadtree, Hexagon, Geohash

Micro-Diagramme

Tweets Coloured by Language within One Week

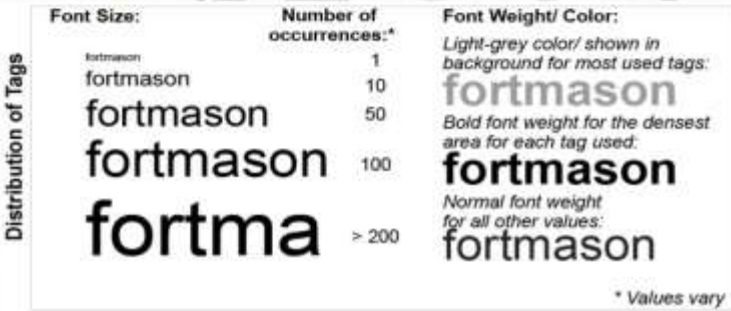
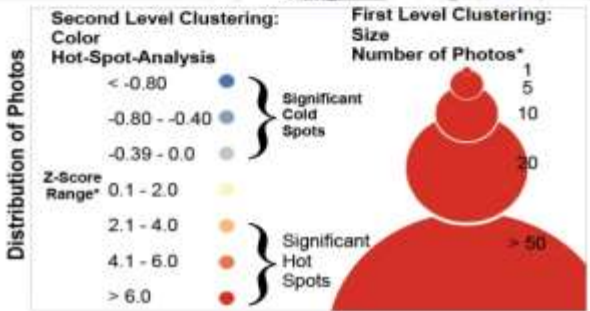
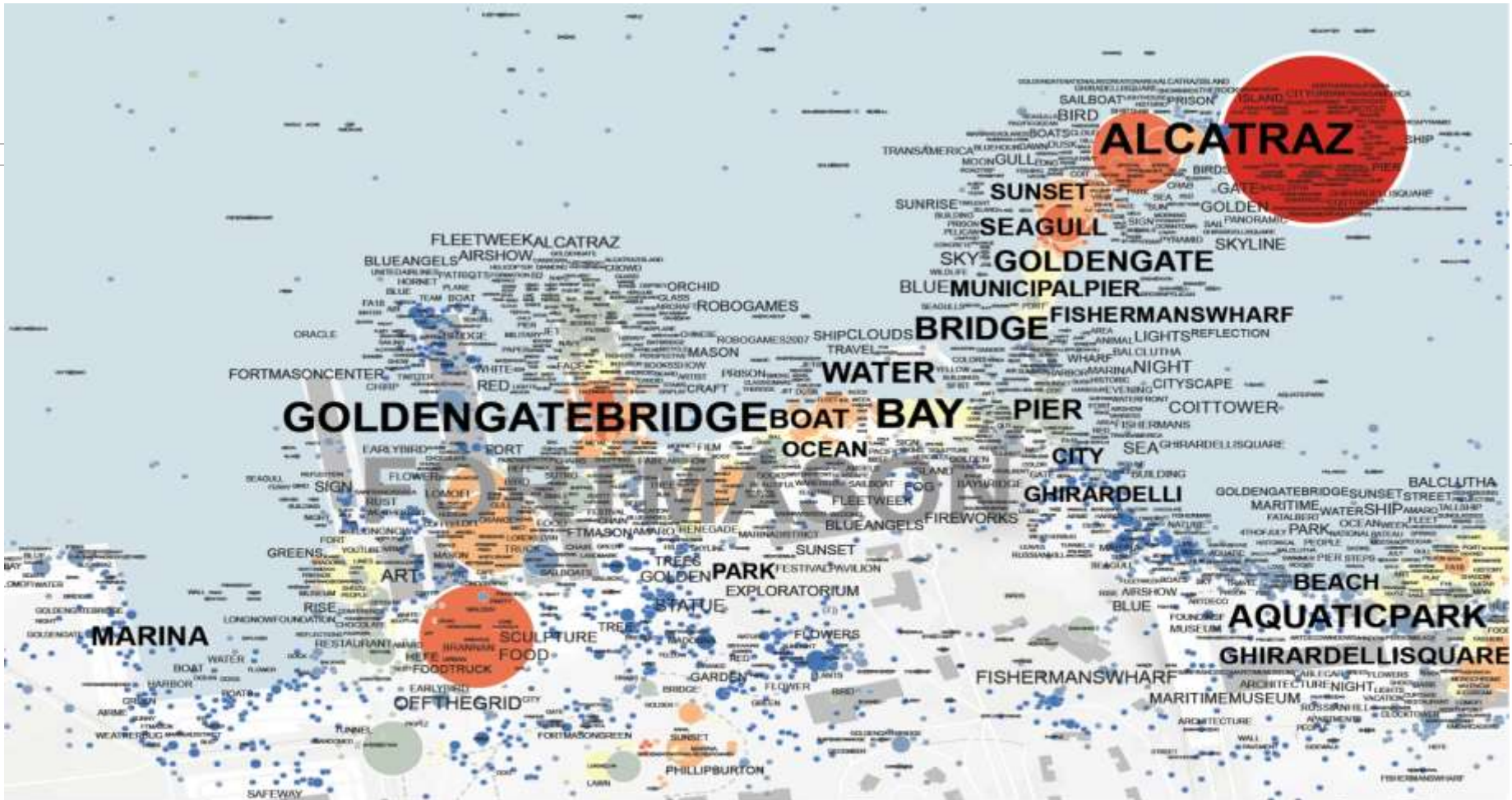


Tag Maps (interne Ableitung)

- Input: punktbezogene Daten, aus georeferenzierten Photodatenbanken (Flickr, Panoramio), multivariat (viele Attribute bzw. tags)
- „tag maps“ für Landschaftsplanung (Dissertation Dunkel, 2016)
 - Hierarchisches Clustering für Aggregation der Photostandorte
- Visualisierung der am häufigsten verwendeten Tags



Dissertation: A. Dunkel (2016)
Assessing the perceived environment through crowdsourced spatial photo content for application to the fields of landscape and urban planning.

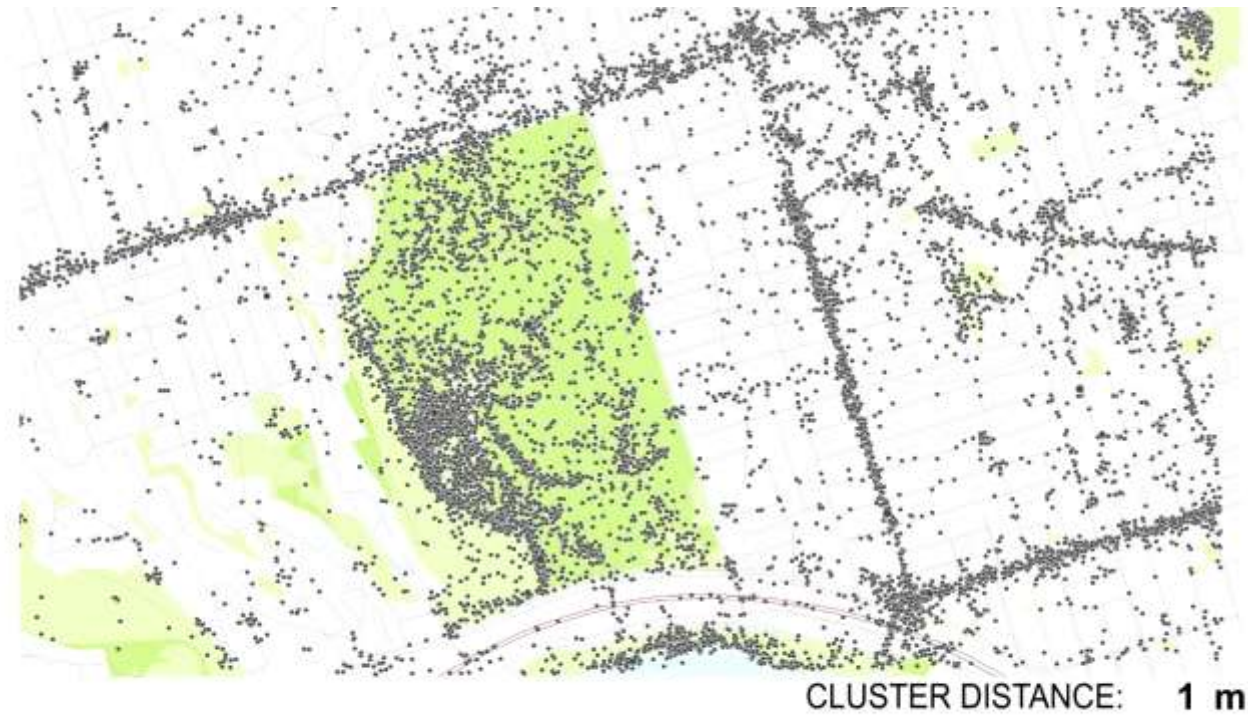


Source: Dr.-Ing. Alexander Dunkel, Institute of Cartography, TU Dresden

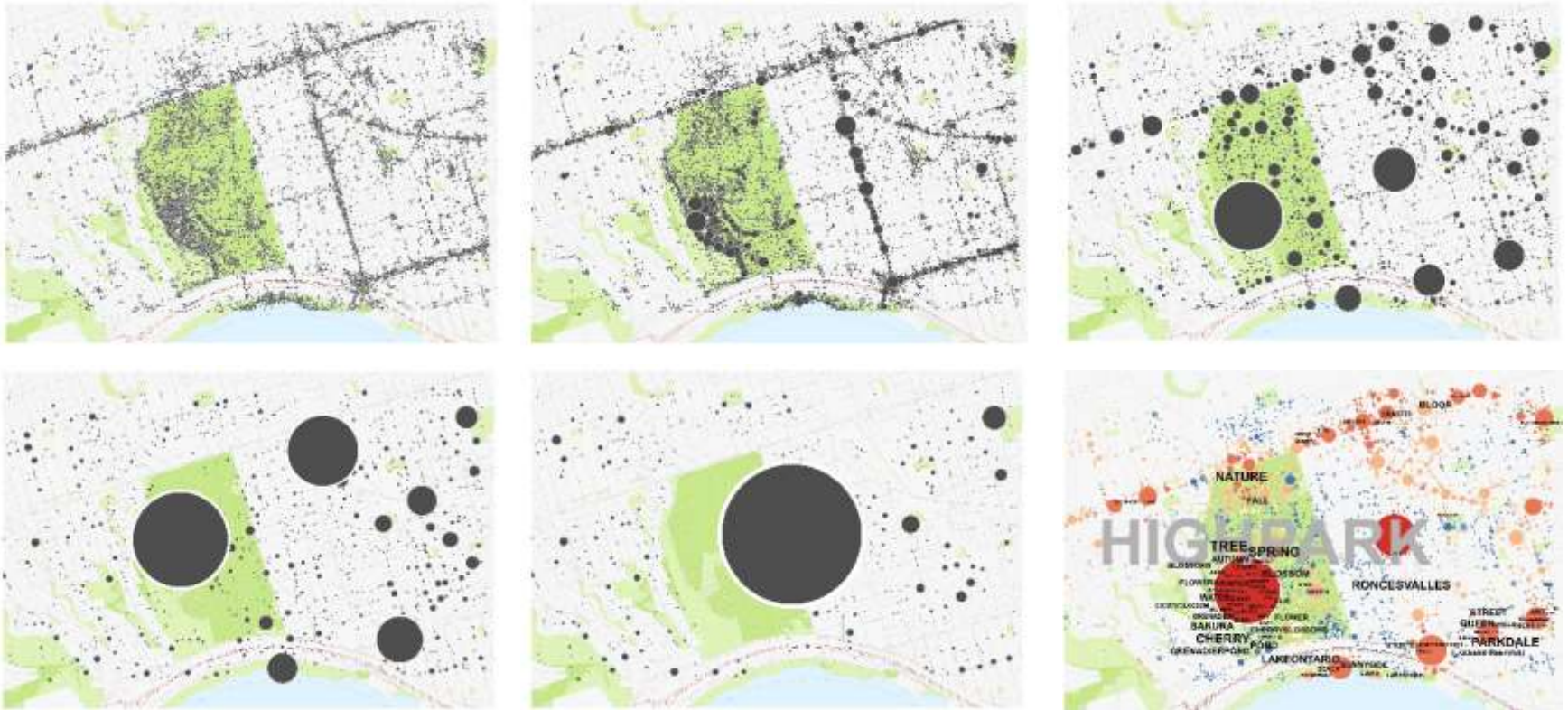
Hierarchisches Clustering

Variation der Clusterdistanz →

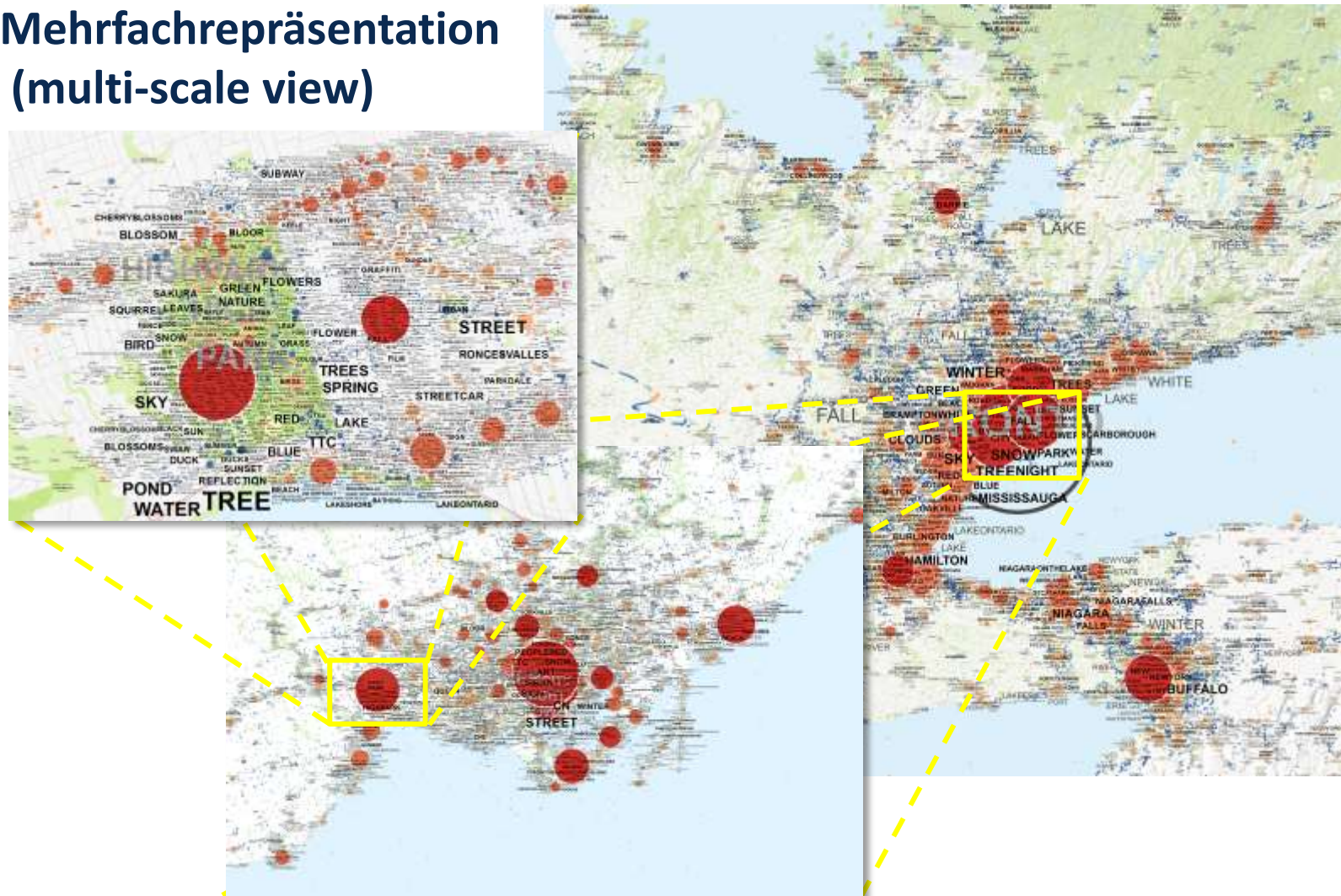
Ableitung von Mehrfachrepräsentation für verschiedene Maßstäbe



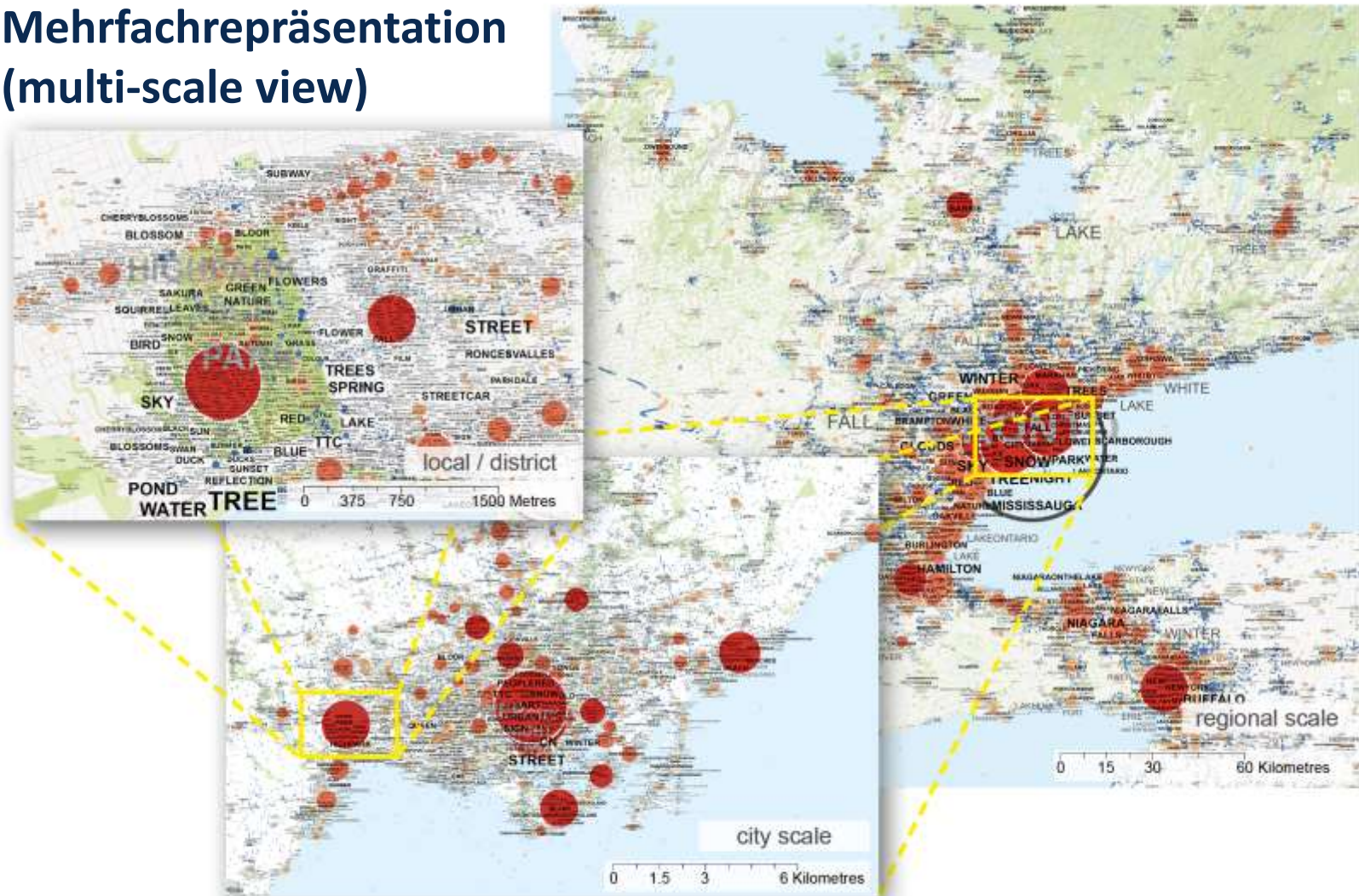
Erzeugung unterschiedlicher Muster durch hierarchisches Clustering



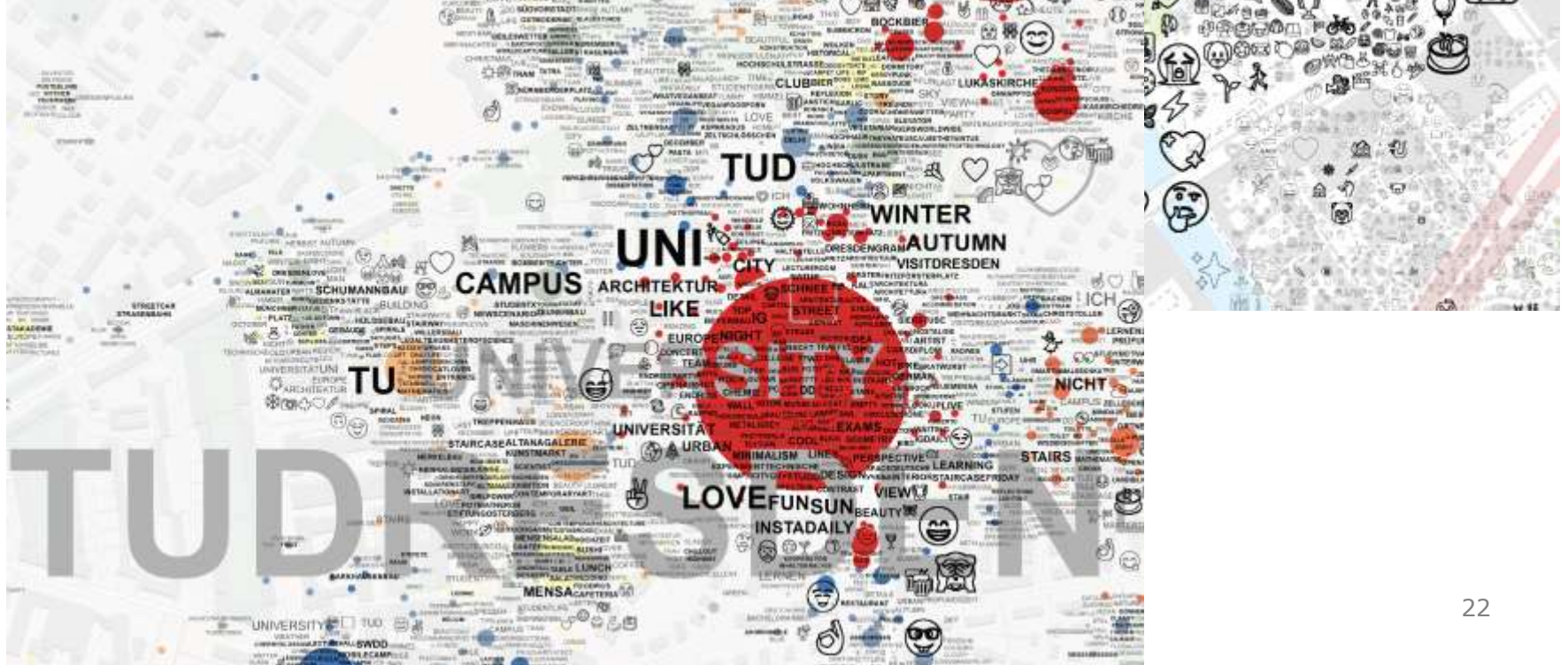
Mehrfachrepräsentation (multi-scale view)



Mehrfachrepräsentation (multi-scale view)



Tag Clouds & Emoji Maps

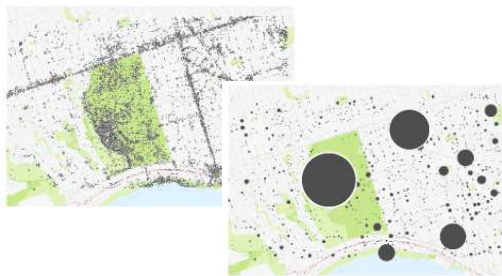


Ableitung von Mehrfachrepräsentation

Interne Ableitung
aus den Daten

Hierarchisches Clustering
(Dendrogramm)

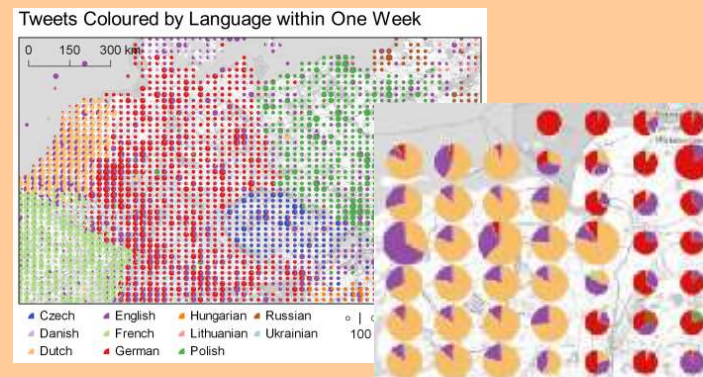
Tag Maps



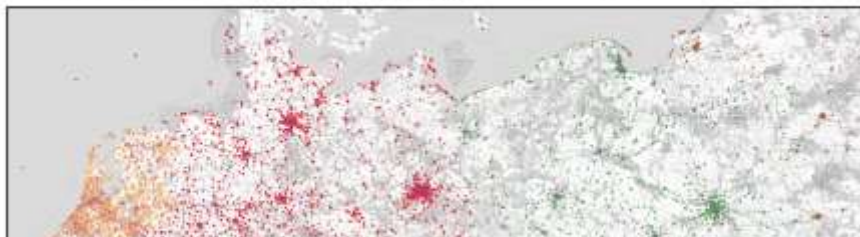
Externe Zuordnung
von Bezugseinheiten

Hierarchische Tessellation
mittels Gitter, Quadtree, Hexagon, Geohash

Micro-Diagramme

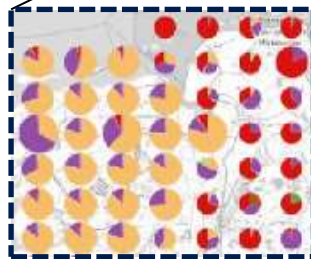
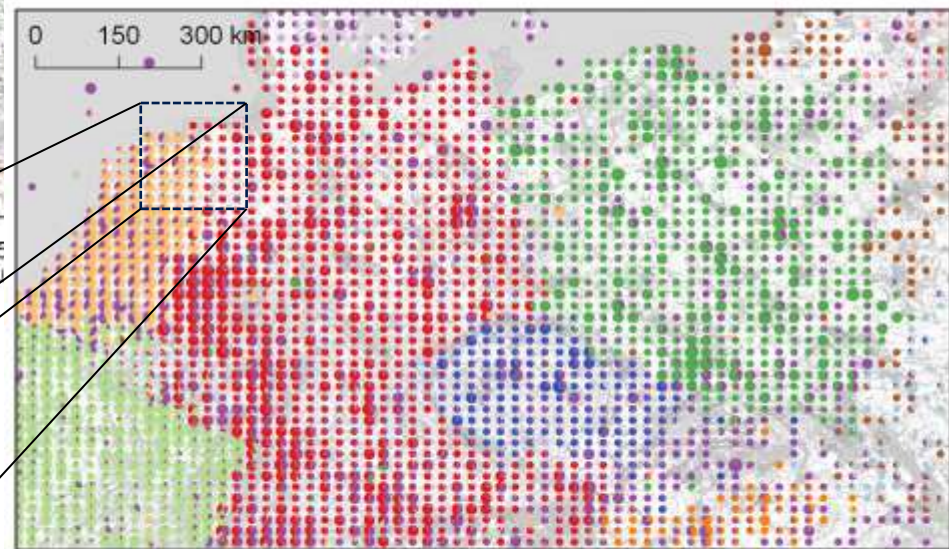


Tweets Coloured by Language within One Week

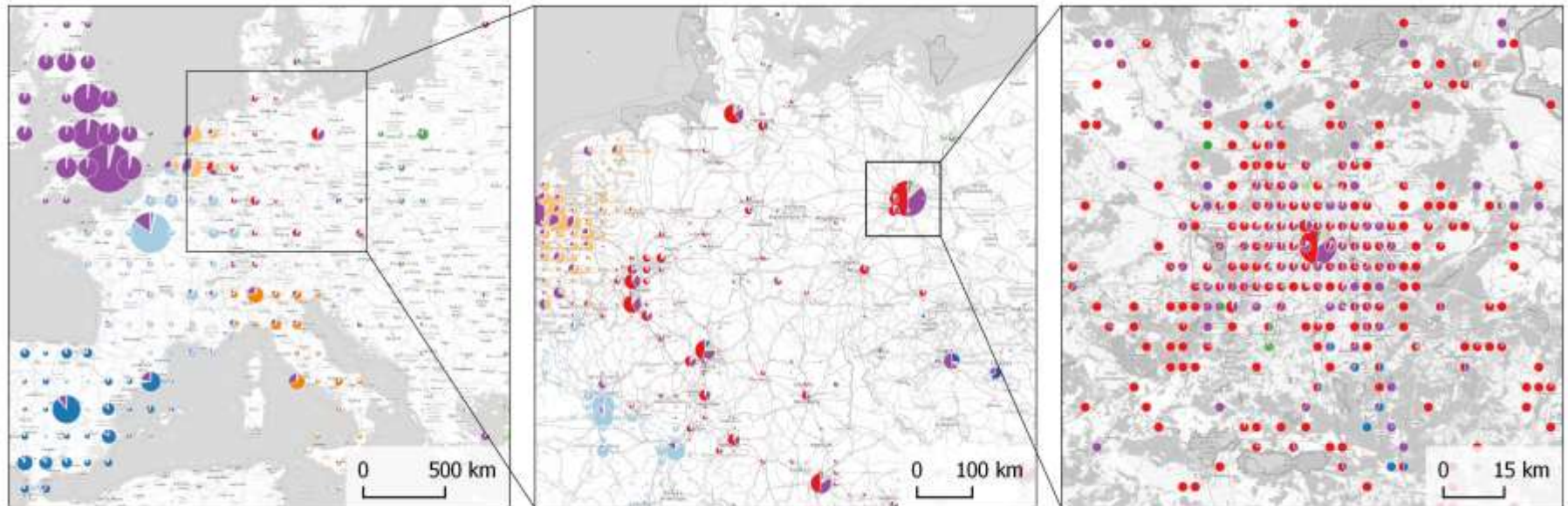


Micro Diagramme – multivariate Visualisierung zur Sprachverteilungen von Tweets

Tweets Coloured by Language within One Week



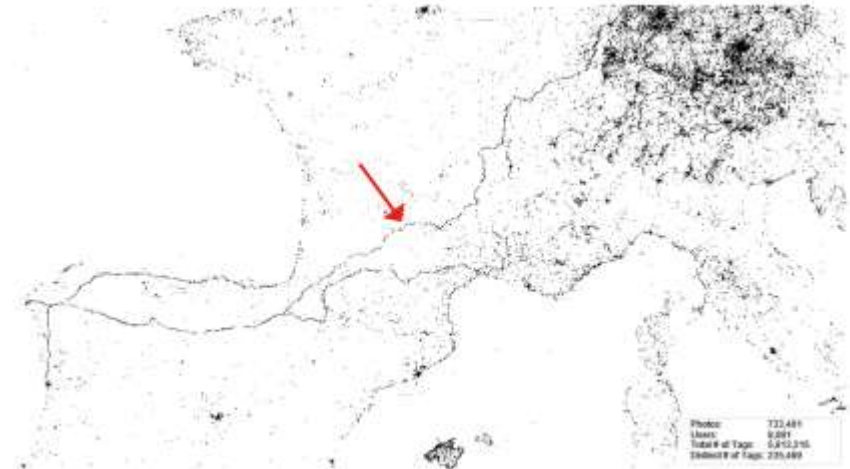
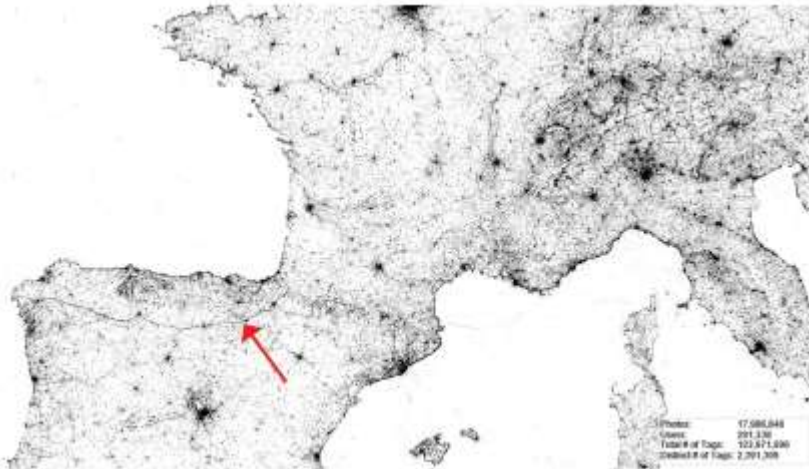
Ableitung von Mehrfachrepräsentation



Multiple Repräsentation zur Analyse variierender Wertzuschreibung verschiedener Nutzergruppen

im weiteren Sinne lassen sich “Mehrfachrepräsentationen” einsetzen für die Sichtweisen von Nutzergruppen (subjektive Wahrnehmungen)

- erfordert Spezifikation von unterschiedlichen Nutzergruppen in den Daten, z.B. bzgl. Alter oder Heimatstadt (Nutzerprofil)
- Abbildung → Jakobsweg ist erkennbar aus den Photostandorten dt. Nutzer



Flickr Photostandorte für Photos aller Nutzer (links) und Nutzer aus Deutschland (rechts)

Fazit

- raumbezogene, nutzergenerierte Daten in Form von VGI besitzen großes Potential für verschiedene Anwendungsfelder
- für deren Erschließung hat DFG Schwerpunktprogramm eingerichtet VGIScience - Volunteered Geographic Information (SPP 1894)
- ein Forschungsschwerpunkt liegt im Bereich interaktiven, geovisuellen Analyse von Location-Based Social Media (LBSM)
- Mehrfachrepräsentation ermöglicht Analyse unterschiedliche Muster auf verschiedene Maßstäben
 - Interne Ableitung aus den Daten – hierarchisches Clustering
 - Externe Zuordnung von Bezugseinheiten – Gitter, Quadtree, Geohash
- Beispiele: *Tag Maps* und *Micro Diagramme* als angepasste Visualisierungsmethoden für LBSM



Prof. Dr.-Ing.habil Dirk Burghardt
Institut für Kartographie
TU Dresden

dirk.burghardt@tu-dresden.de