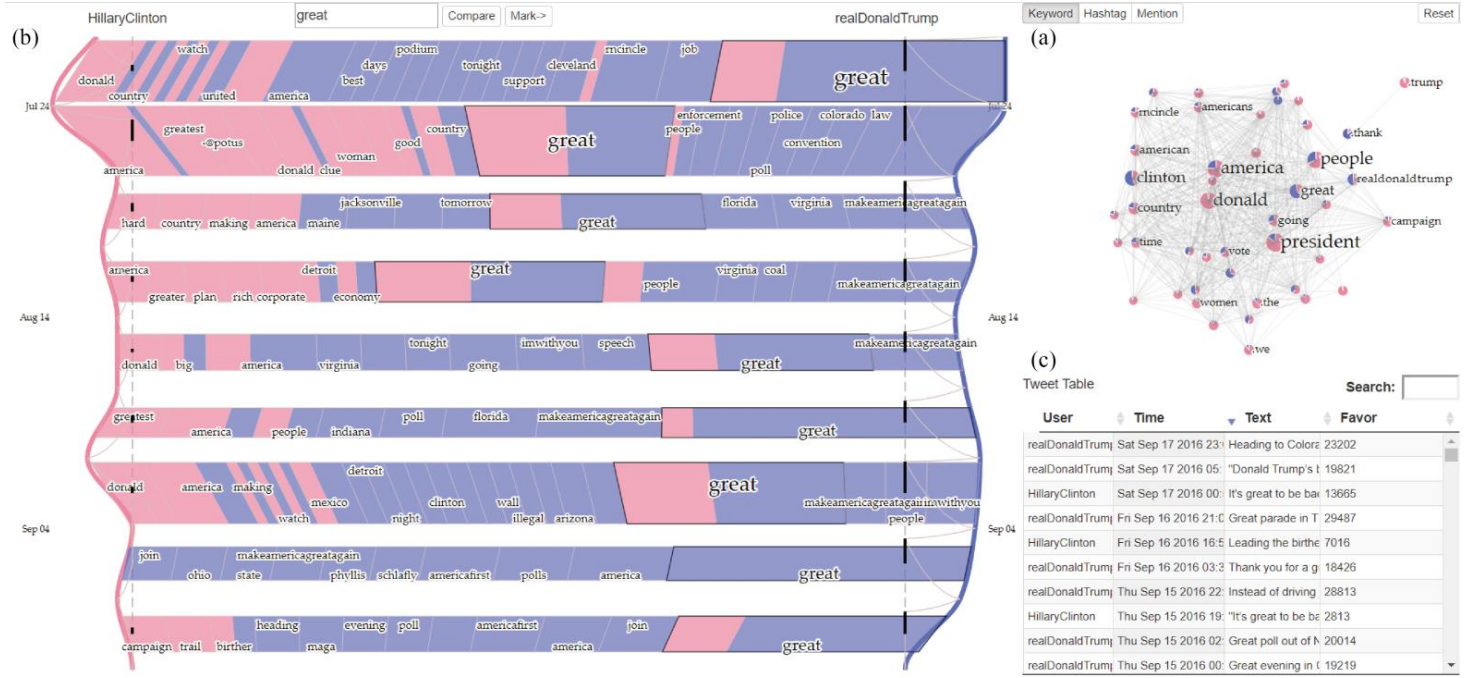
Deutsche Forschungsgemeinschaft

VGIscience
Lecture Series, April, 2021

Geovisual analysis of VGI for understanding people's behaviour in relation to multi-faceted context (EVA-VGI-II)



EVA-VGI-Team from Bonn, Zurich and Dresden

Outline

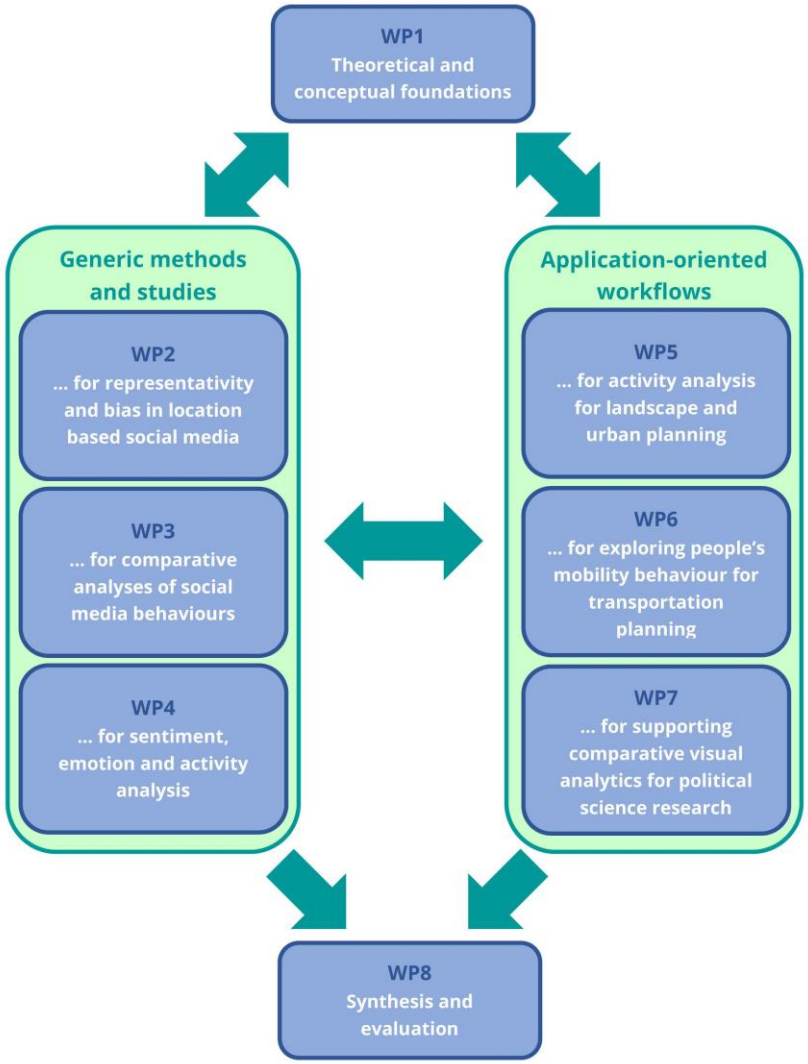
- Introduction EVA-VGI-II
  - research and work packages
- Conceptual models
  - Facet framework
  - Theoretical model for pattern discovery
- Application examples of geovisual analysis related to VGI and LBSM
  - Exploring the potential of twitter to understand traffic events
  - Identifying shared valued locations and tranquility (landscape planning)
  - Multi-item data streams during presidential election and emotional reactions to Brexit expressed by emojis (politics)
- Conclusions

Introduction EVA-VGI-II

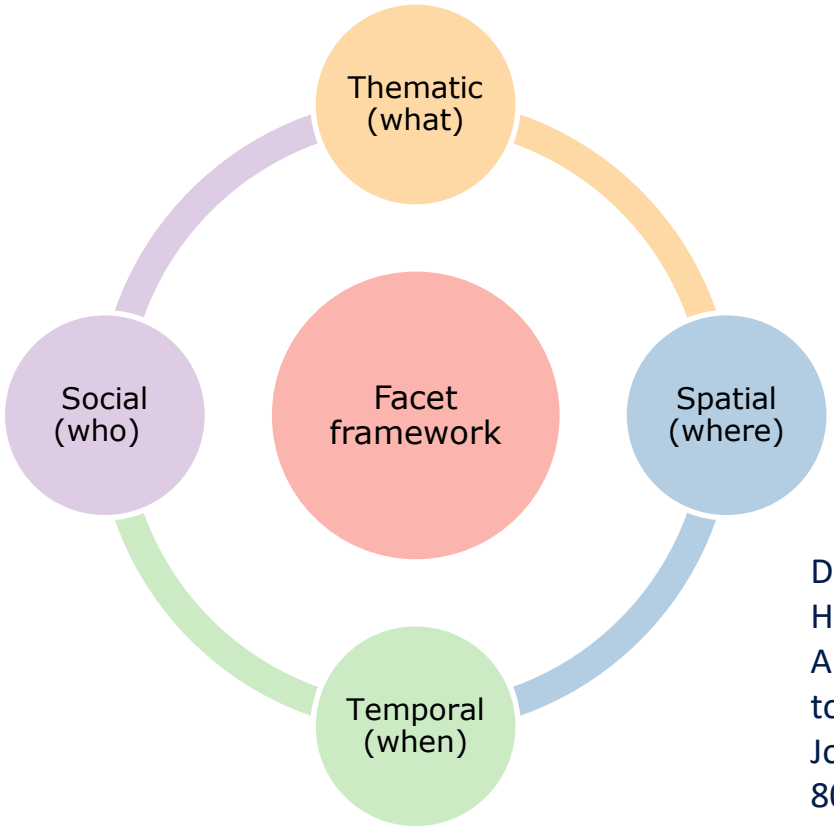
- Social media is part of everyday life for many people and used to exchange opinions, impressions, experiences or information
- Location-based social media
  - Twitter (500 mill. tweets per day 3.1.2020, 0,85% tweets are geotagged)
  - Instagram (600 mill. daily active user 28.1.2020, 21 mill. user in Germany 7.7.2020)
  - Flickr (350 mill. photo)
- aim: develop geovisual analysis methods which will help to show
  - how people interact in LBSN
  - how their interactions influence, and are influenced by, their physical and social environment

Research questions and workpackages

- two lines of research
  - with focus on generic methods (wp2-4) and application oriented workflows (wp5-7)
- generic methods
  - representativity and bias
  - comparative analysis
  - emotion and activity analysis
- applications
  - landscape and urban planning
  - mobility and transportation
  - political science research



Conceptual model (facet framework)

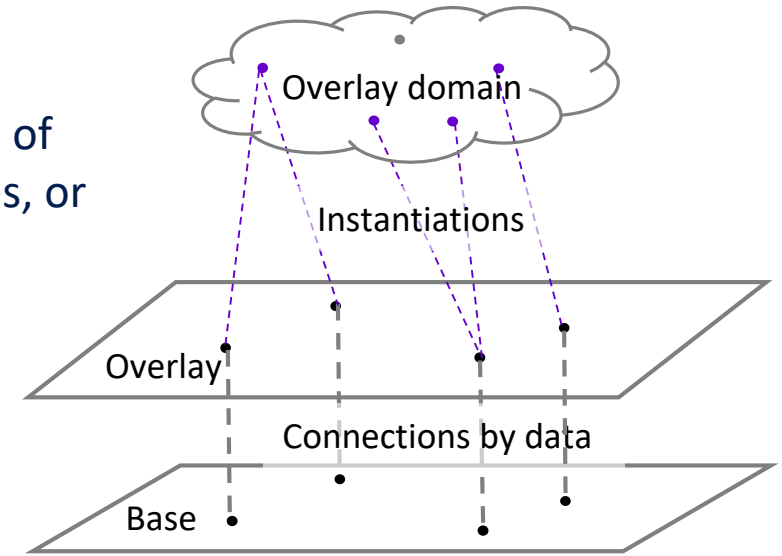


Dunkel, A.; Andrienko, G.; Andrienko, N.; Burghardt, D.; Hauthal, E. and Purves, R. (2018).
A conceptual framework for studying collective reactions to events in location-based social media. International Journal of Geographical Information Science, 33:4, 780-804. https://doi.org/10.1080/13658816.2018.1546390

A theoretical model for pattern discovery
(data distribution and its components)

- a pattern consists of relationships between elements of at least two data components
- a data component is a set of items of the same kind, e.g., a set of entities, or attribute values, or references to places or times
- differentiation between “base” and “overlay” component

Schematic representation



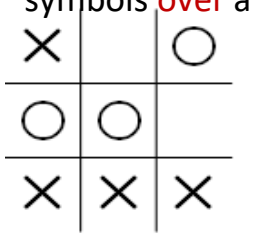
N. Andrienko, G. Andrienko, S. Miksch, H. Schumann, S. Wrobel (2020)
A theoretical model for pattern discovery in visual analytics.
Visual Informatics. https://doi.org/10.1016/j.visinf.2020.12.002

A theoretical model for pattern discovery
(data distribution and its components)

Examples

Distribution of colours over a set of apples

Distribution of X and O symbols over a grid



Distribution of moon phases over time

Patterns may refer to:

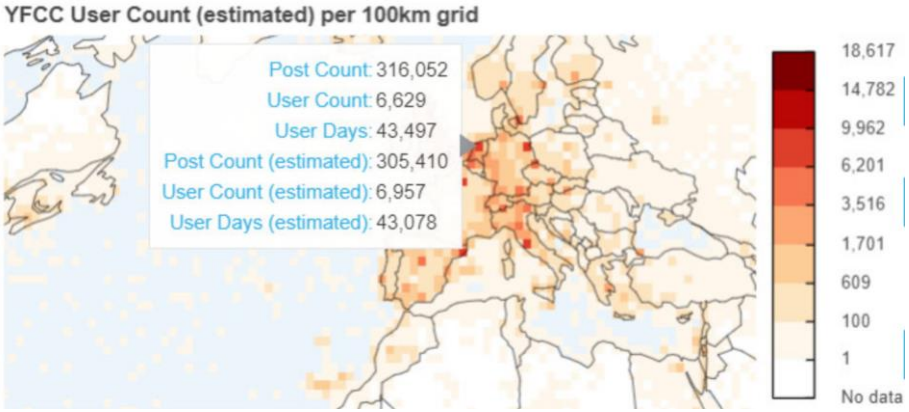
- overlay composition (frequencies of element occurrences)
- arrangement of overlay elements according to base relationships (spatial shape, temporal sequence, co-occurrences)
- overlay variation with respect to the arrangement (equivalence, differences, changes)

N. Andrienko, G. Andrienko, S. Miksch, H. Schumann, S. Wrobel (2020)
A theoretical model for pattern discovery in visual analytics.
Visual Informatics. https://doi.org/10.1016/j.visinf.2020.12.002

Open LBSN schema

https://lbsn.vgiscience.org/

- consideration of facet (base) and pattern discovery model (overlay)
  - base reference:
    spatial grid, period, hashtag
  - overlay metrics:
    number of posts (postcount),
    number of user (usercount),
    number of distinct user
    per day (userdays)

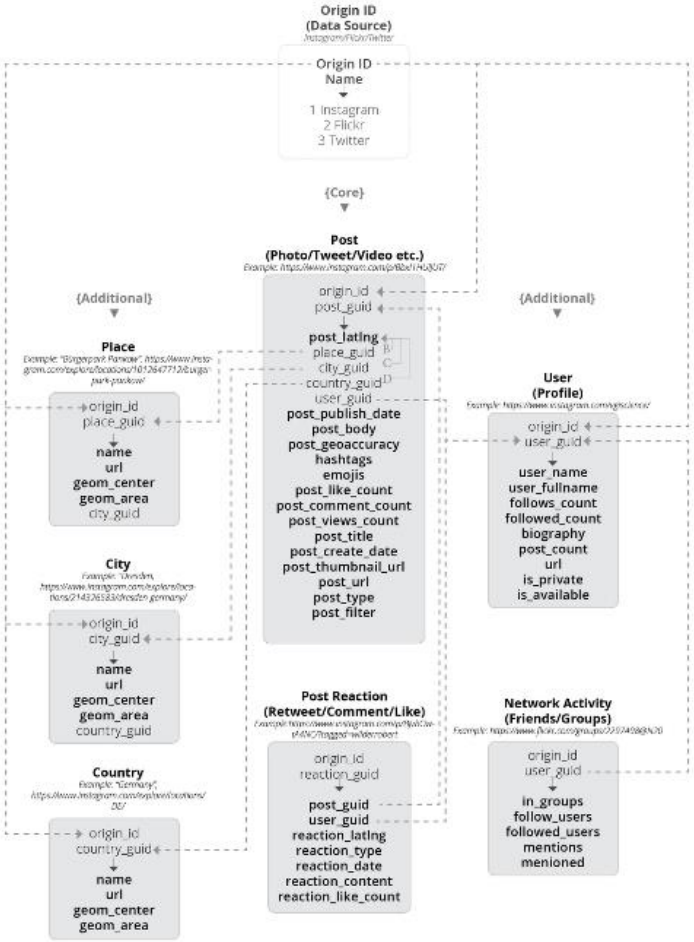


Dunkel, A., Löchner, M., & Burghardt, D. (2020). Privacy-aware visualization of volunteered geographic information (VGI) to analyze spatial activity: A benchmark implementation. ISPRS International Journal of Geo-Information, 9(10). http://doi.org/10.3390/ijgi9100607

Open LBSN schema

https://lbsn.vgiscience.org/

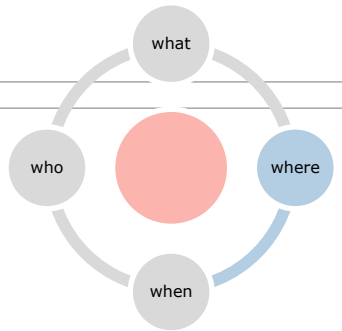
- common schema to integrate data from various social-media platform (Twitter, Flickr, Instagram)
- associated privacy aware data structure (HyperLogLog)
  - metadata from social media post are converted into statistical data
  - analysis of individual posts only with original data possible
  - high-performance quantitative analyses



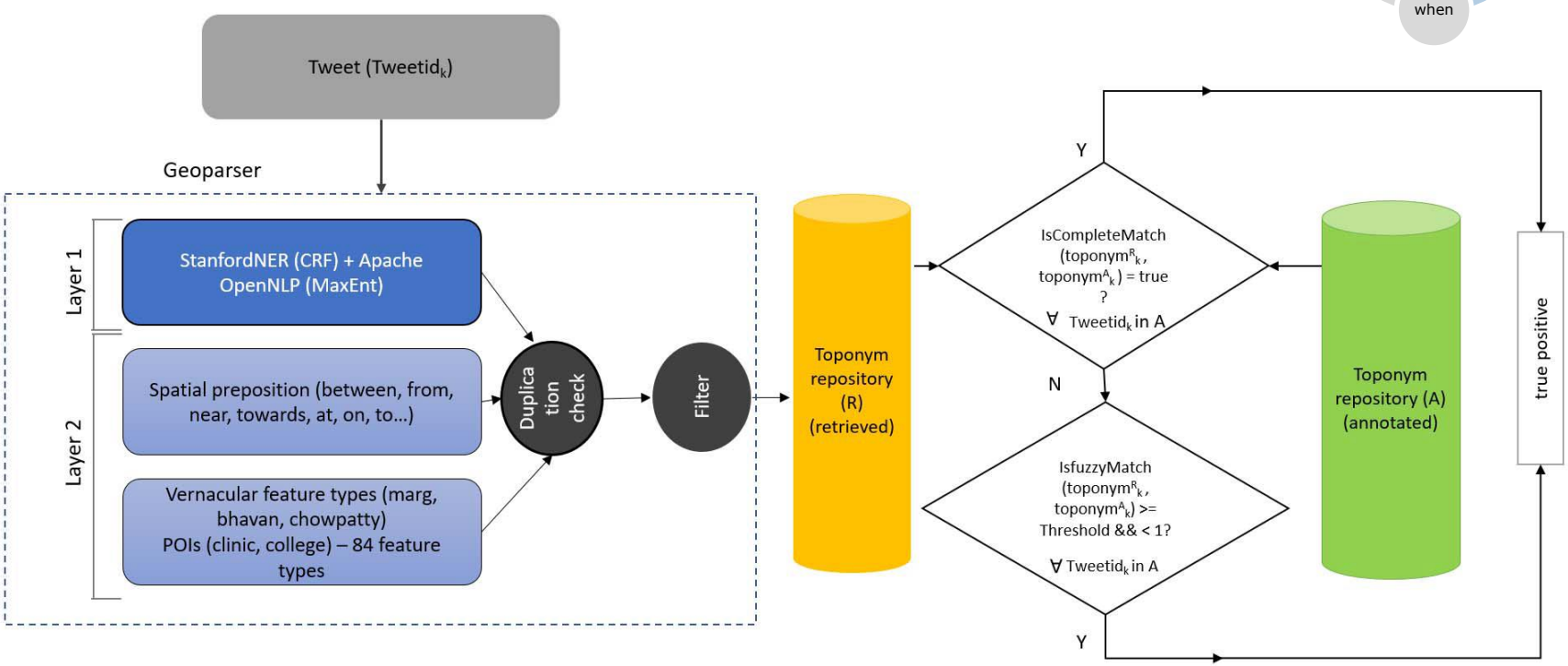
Dunkel, A., Löchner, M., & Burghardt, D. (2020). Privacy-aware visualization of volunteered geographic information (VGI) to analyze spatial activity: A benchmark implementation. ISPRS International Journal of Geo-Information, 9(10). http://doi.org/10.3390/ijgi9100607

Outline

- Introduction EVA-VGI-II
  - research and work packages
- Conceptual models
  - Facet framework
  - Theoretical model for pattern discovery
- Application examples of geovisual analysis related to VGI and LBSM
  - Exploring the potential of twitter to understand traffic events
  - Identifying shared valued locations and tranquility (landscape planning)
  - Multi-item data streams during presidential election and emotional reactions to Brexit expressed by emojis (politics)
- Conclusions

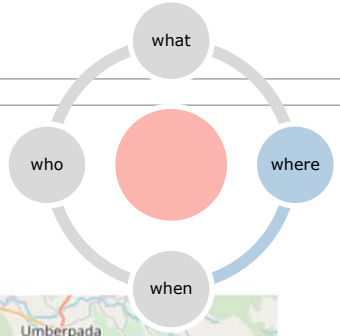


Method development for the spatial component

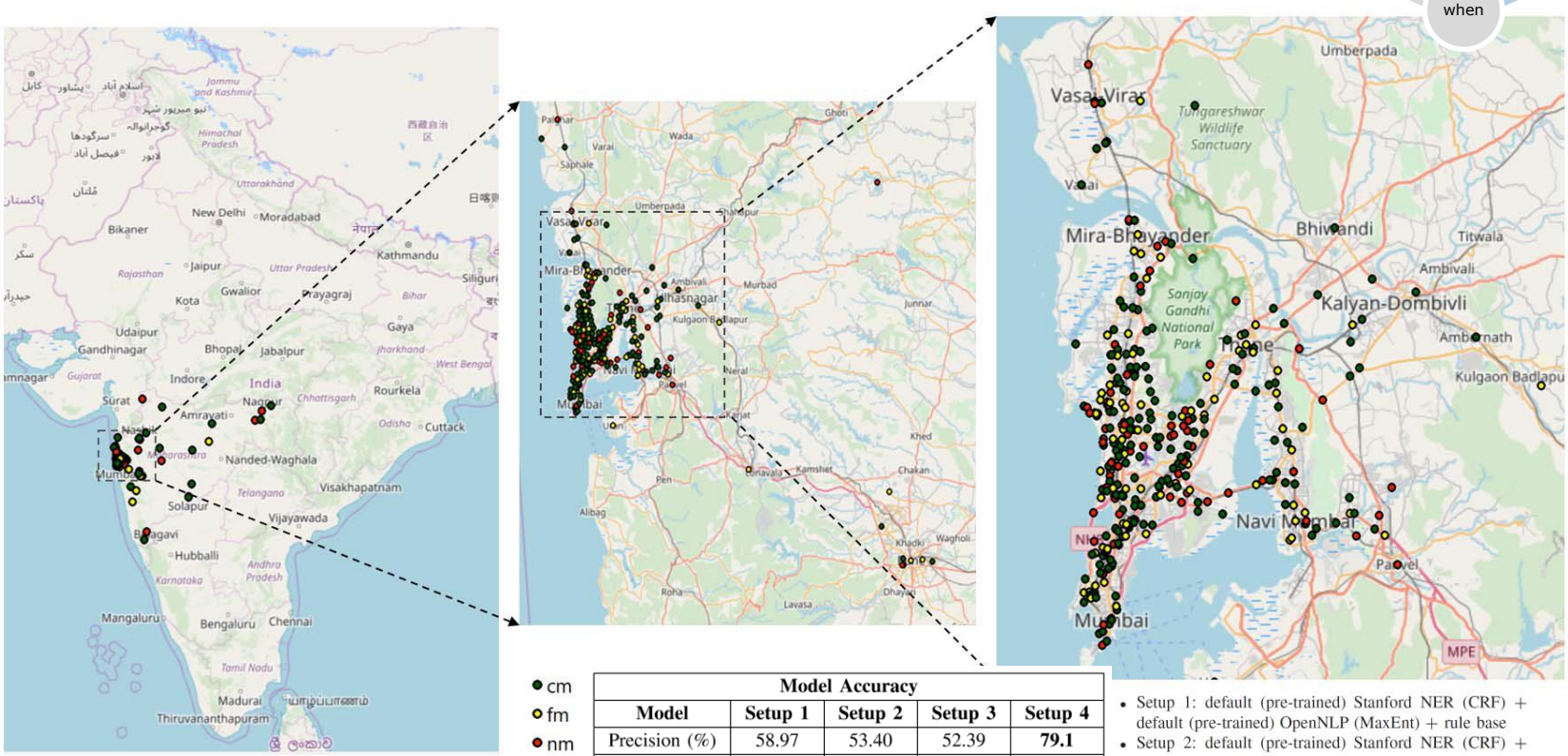


Hybrid georeferencing workflow which handles vernacular language

Das, R. D., & Purves, R. S. (2019). Exploring the Potential of Twitter to Understand Traffic Events and Their Locations in Greater Mumbai, India. IEEE Transactions on Intelligent Transportation Systems, 21(12), 5213-5222.



Method development for the spatial component



(a)

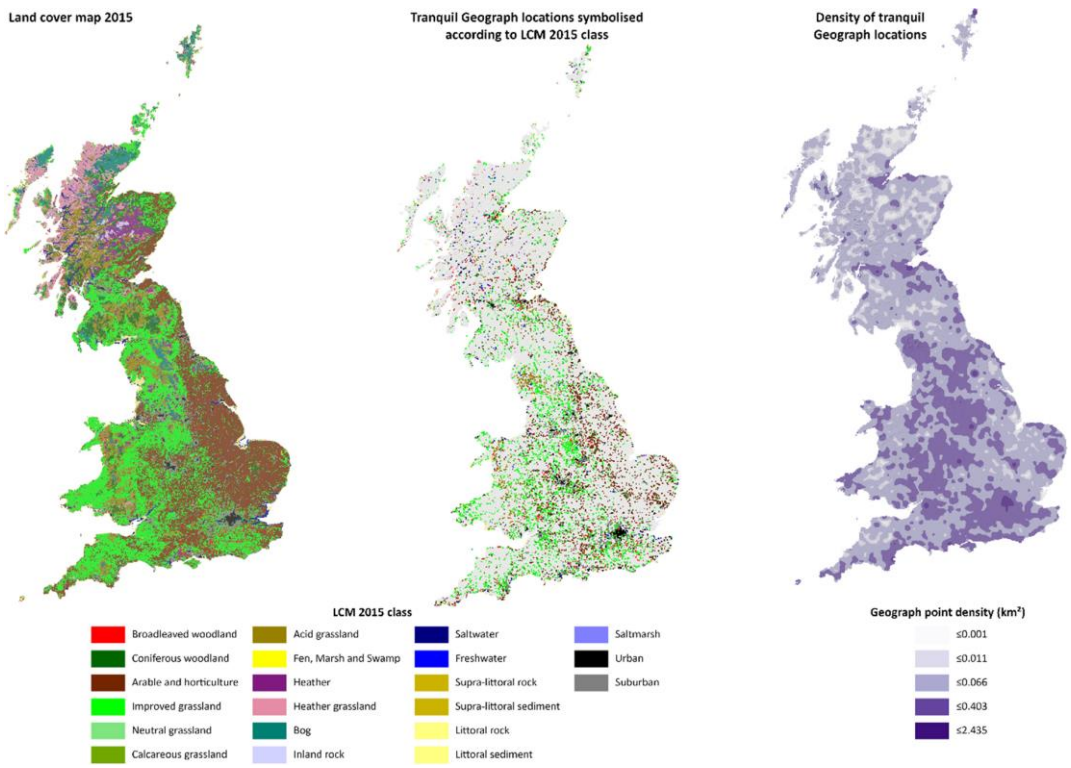
- cm
- fm
- nm

Model Accuracy				
Model	**Setup 1**	**Setup 2**	**Setup 3**	**Setup 4**
Precision (%)	58.97	53.40	52.39	**79.1**
Recall (%)	60.95	**79.14**	64.00	7.75
F1 (0-1)	0.59	**0.63**	0.57	0.14

- Setup 1: default (pre-trained) Stanford NER (CRF) + default (pre-trained) OpenNLP (MaxEnt) + rule base
- Setup 2: default (pre-trained) Stanford NER (CRF) + retrained OpenNLP (MaxEnt) + rule base
- Setup 3: retrained OpenNLP (MaxEnt)
- Setup 4: default (pre-trained) OpenNLP (MaxEnt)

Das, R. D., & Purves, R. S. (2019). Exploring the Potential of Twitter to Understand Traffic Events and Their Locations in Greater Mumbai, India. IEEE Transactions on Intelligent Transportation Systems, 21(12), 5213-5222.

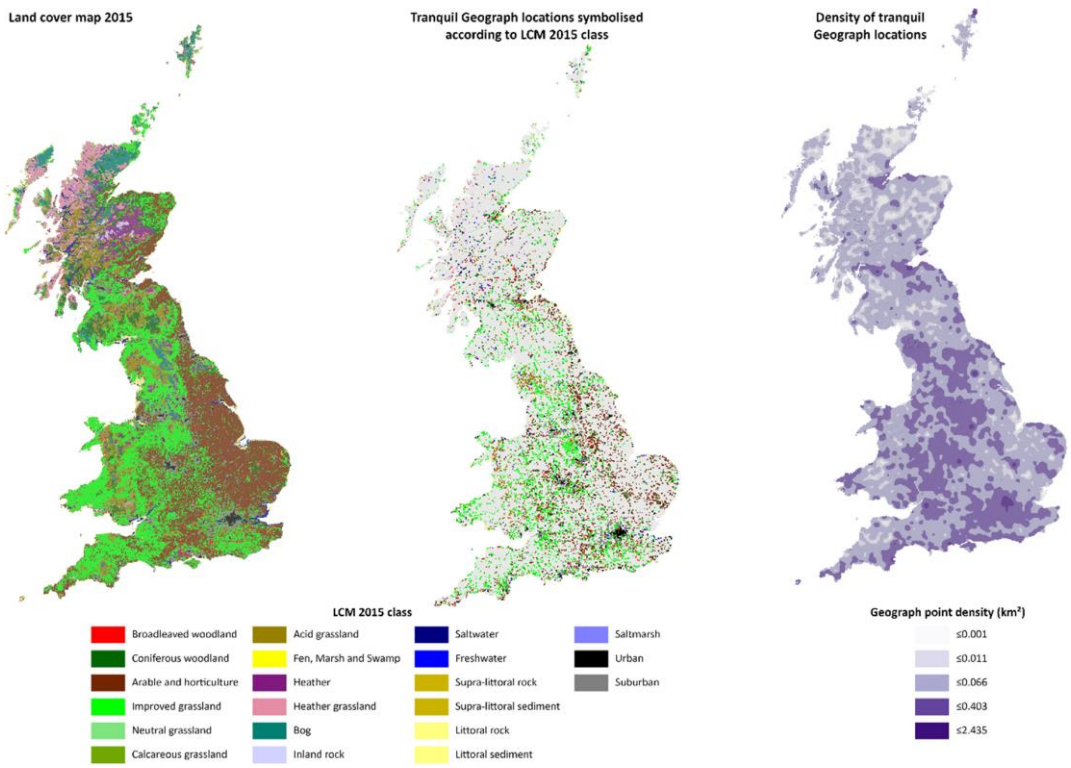
Where and what is tranquility?



- Identifying tranquil areas is important for landscape planning and policy-making
- discrepancies between modelled tranquil areas and where people experience tranquility in the field
- VGI platform: Geograph.UK (photos and descriptions per 1 km^2)
- 15’350 tranquility descriptions could be extracted
- analyses how tranquillity is described in different land cover classes

Wartmann, F. M., Koblet, O., & Purves, R. S. (2021). Assessing experienced tranquillity through natural language processing and landscape ecology measures. Landscape Ecology, 1-19.

Where and what is tranquility?

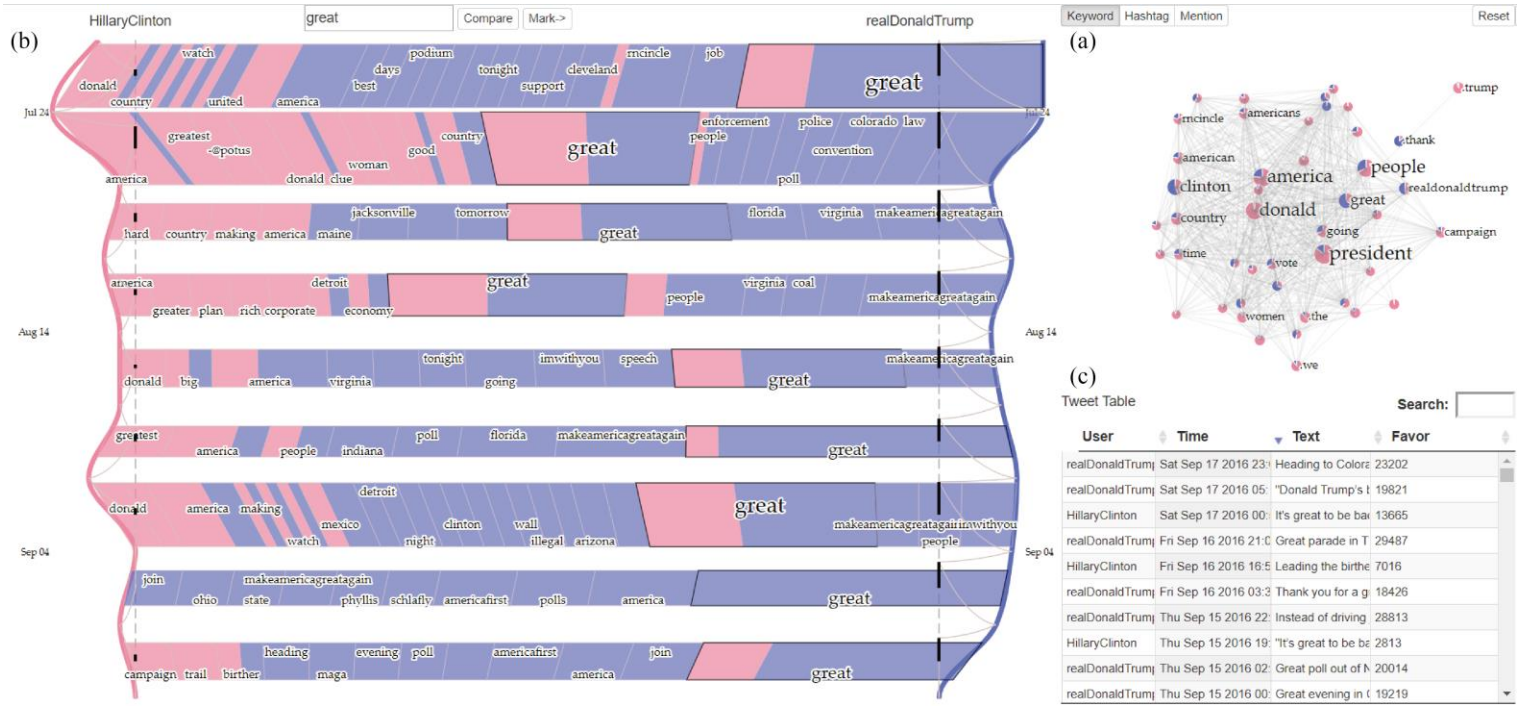


	Broadleaved woodland	Coniferous woodland	Arable & horticulture	Improved grassland	Urban	Suburban
road	26	26	194	263	30	148
lane	35		269	336	9	35
day	9	20	38	108	42	56
spot	26	14	75	191	37	54
street	4		27	42	74	114
place	21	11	50	96	34	39
morning	12	9	24	37	27	45
village			114	167		23
area	10	3	28	57	29	71
sac			27	29	14	71
corner	13	4	49	83	39	47
scene	15	7	68	105	22	36
afternoon	7	4	27	51	38	41
water		6	8	16	5	11
location	11	5	29	50	5	13
walk	4	5	13	22	3	9
countryside	5		60	66	8	
backwater	4		39	65	14	41
contemplation	9		4	8	5	
side	8	4	12	71		
valley		4	6	4	7	
time			12	18	22	10
stretch	8		47	72	8	10
moment	6	3	10	33	18	8
streets	4		25	23	18	28
evening				3	18	11
path	7		7	14		12
oasis			3	5	17	9
setting		3	16	36		6
conditions		3	5	10		3
waters	5		10	13	16	13
lanes			38	43		
part			15	29		27
garden			7	7	10	25
country	3		33	40		
roads	5		22	35	5	14
farmland			26	5		
enclave				3	10	8

1st quartile
2nd quartile
3rd quartile
4th quartile
No data

Wartmann, F. M., Koblet, O., & Purves, R. S. (2021). Assessing experienced tranquillity through natural language processing and landscape ecology measures. Landscape Ecology, 1-19.

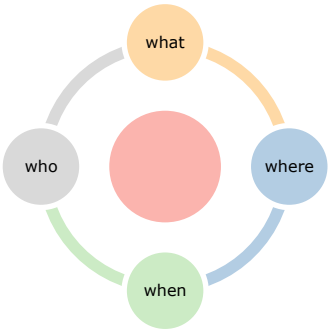
Comparison for Multi-item Data Streams



Siming Chen, Natalia Andrienko, Gennady Andrienko, Jie Li, Xiaoru Yuan. Co-Bridges: Pair-wise Visual Connection and Comparison for Multi-item Data Streams. IEEE Transactions on Visualization and Computer Graphics (VAST'20), 27(2): 1612-1622, 2021. doi: 10.1109/TVCG.2020.3030411

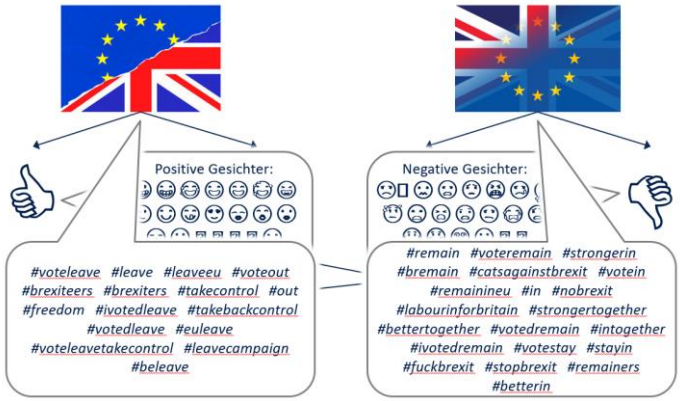
Analyzing Emotional Reactions Expressed by Emojis (Brexit)

- Case study and facets (where, when, what)
  - referendum regarding Brexit
    (287.206 georeferenced tweets from UK, June-July 2016)



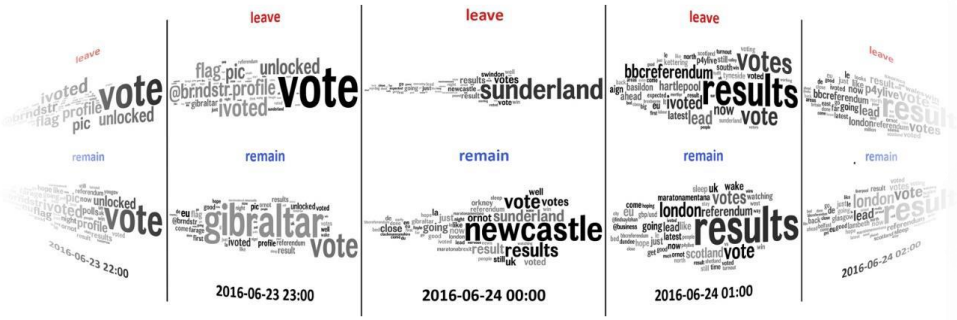
- Methods and visualisation metaphors
  - emojis as language-independent emotional signals combined with spatial-temporal analysis based on
    - word clouds
    - timelines
    - frequency diagrams
    - choropleth maps

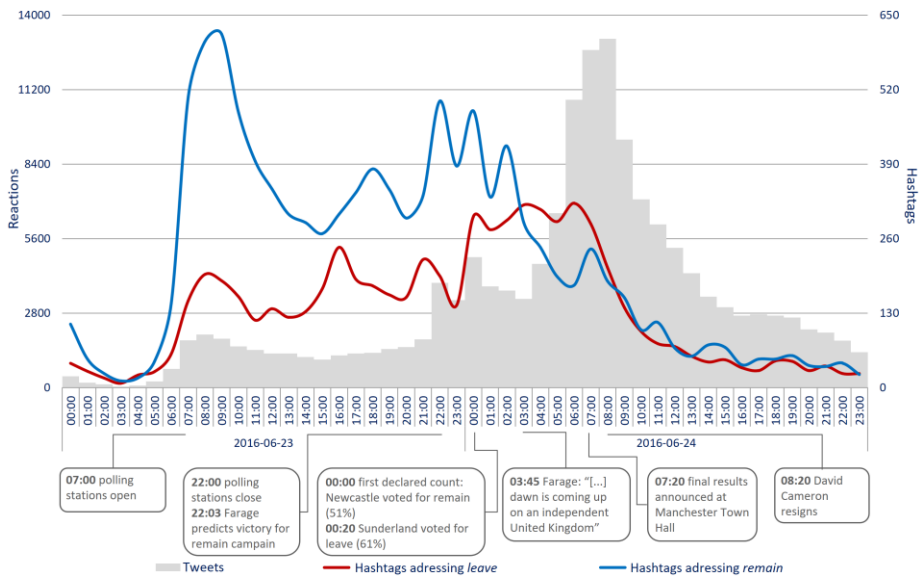
Unicod Emoji List v5.0,
face-negative

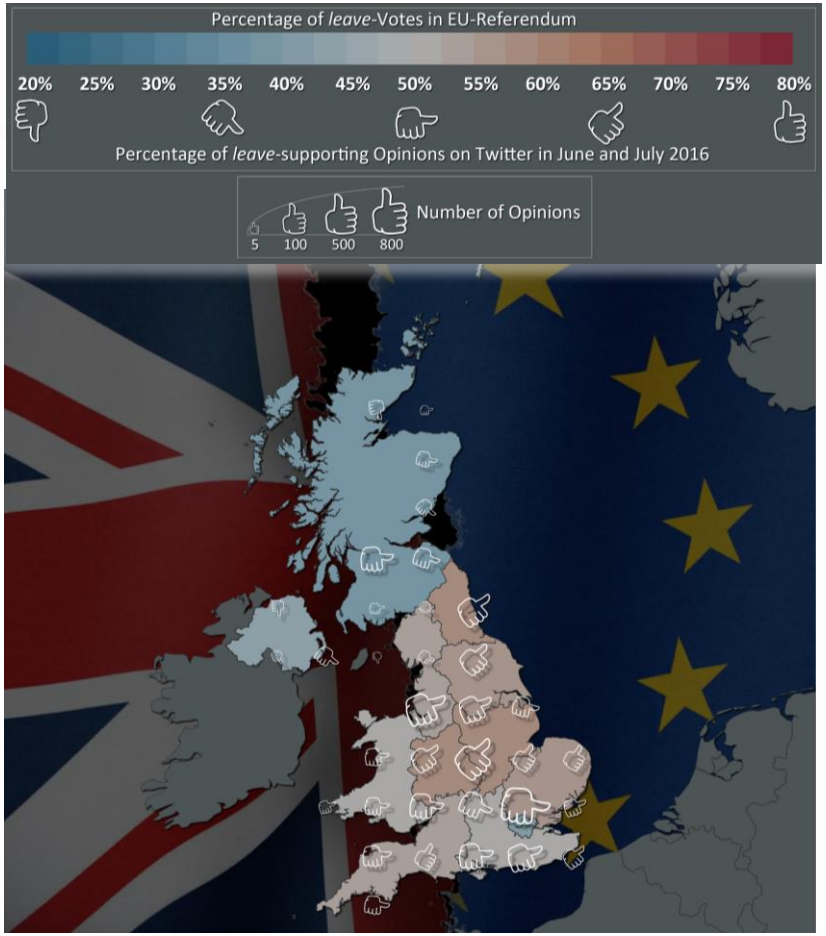


Hauthal, E.; Burghardt, D. and Dunkel, A. (2019). Analyzing and Visualizing Emotional Reactions Expressed by Emojis in Location-Based Social Media. ISPRS Int. J. Geo-Inf. 2019, 8(3), 113; https://doi.org/10.3390/ijgi8030113

Analyzing Emotional Reactions Expressed by Emojis (Brexit)







Challenges and solutions for geovisual analysis of LBSN

Challenges	Approaches and solutions
platform diversity and dynamics	Open LBSN schema https://lbsn.vgiscience.org/
theoretical foundations	theoretical model for pattern discovery and facet model
language processing	consideration of grammars (with negation and forms of increase), utilisation of emojis
influence of population distribution and individual users	metrics (postcount, usecount, userdays) and normalisation (e.g. tf-idf, chi-square)
resolution and granularity	interactive scale dependent visualisation (multiple representation)
privacy and ethical issues	privacy-aware data structures (HyperLogLog)

Selected publications

- **Andrienko, N., Andrienko, G.,** Miksch, S., Schumann, H. and **Wrobel, S.** (2020) A theoretical model for pattern discovery in visual analytics. Visual Informatics. https://doi.org/10.1016/j.visinf.2020.12.002
- Bahrehdar, A.H., Adams, B. and **Purves, R.S.** (2020). Streets of London: Using Flickr and OpenStreetMap to build an interactive image of the city, Computers, Environment and Urban Systems, 84, https://doi.org/10.1016/j.compenvurbsys.2020.101524
- **Chen, S., Andrienko, N., Andrienko, G.**, Li , J. and Yuan, X. (2021). Co-Bridges: Pair-wise Visual Connection and Comparison for Multi-item Data Streams, in IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 2, pp. 1612-1622, https://doi.org/10.1109/TVCG.2020.3030411
- **Das, R.D.** and **Purves, R.S.** (2020). Exploring the Potential of Twitter to Understand Traffic Events and Their Locations in Greater Mumbai, India, IEEE Transactions on Intelligent Transportation Systems, 21(12), 5213-5222, http://doi.org/10.1109/TITS.2019.2950782
- **Dunkel, A., Löchner, M.** and **Burghardt, D.** (2020). Privacy-aware visualization of volunteered geographic information (VGI) to analyze spatial activity: A benchmark implementation. Special Issue "Volunteered Geographic Information and Citizen Science", ISPRS International Journal of Geo-Information, 9(10), 607; https://doi.org/10.3390/ijgi9100607
- **Dunkel, A.; Andrienko, G.; Andrienko, N.; Burghardt, D.; Hauthal, E.** and **Purves, R.** (2018). A conceptual framework for studying collective reactions to events in location-based social media. International Journal of Geographical Information Science, 33:4, 780-804. https://doi.org/10.1080/13658816.2018.1546390
- Gröbe, M. and **Burghardt, D.** (2020). Micro diagrams: visualization of categorical point data from location-based social media. Cartography and Geographic Information Science, 47:4, 305-320, https://doi.org/10.1080/15230406.2020.1733438
- **Hauthal, E.; Burghardt, D.** and **Dunkel, A.** (2019). Analyzing and Visualizing Emotional Reactions Expressed by Emojis in Location-Based Social Media. ISPRS Int. J. Geo-Inf. 2019, 8(3), 113; https://doi.org/10.3390/ijgi8030113
- Wartmann, F.M., Koblet, O. and **Purves, R.S.** (2021). Assessing experienced tranquillity through natural language processing and landscape ecology measures. Landscape Ecol, https://doi.org/10.1007/s10980-020-01181-8