

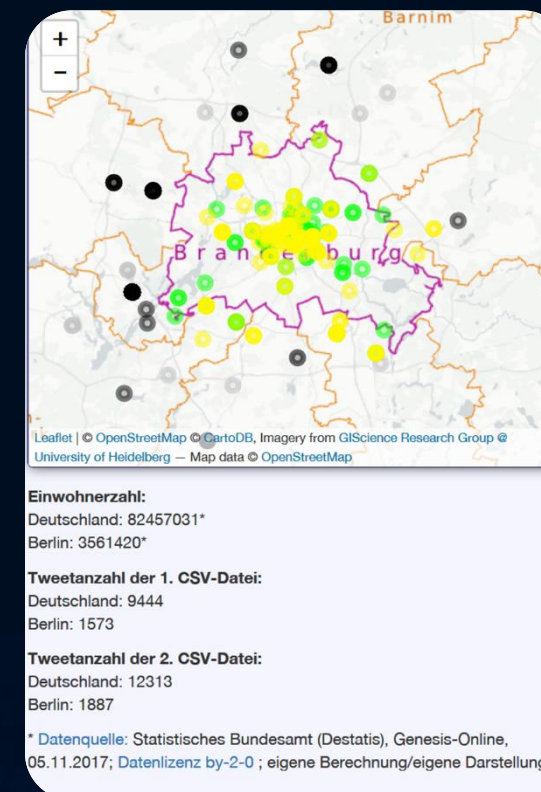
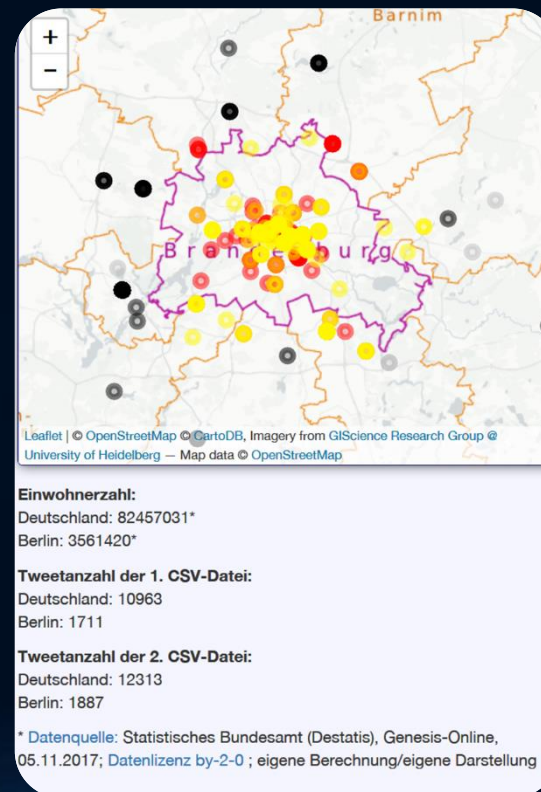
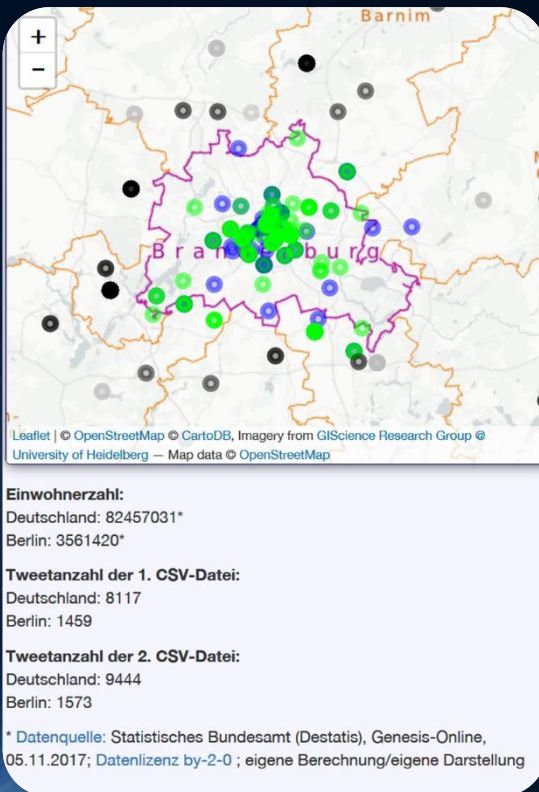
# Die (gesellschaftliche) Digitalisierung in Deutschland am Beispiel von Twitter-Nachrichten

VISUALISIERUNG IN R UND SHINY

Präsentation von Kristina-Susann Baudach  
am 7. E-Government-Hochschultag an der Hochschule Rhein-Waal

# Zu welcher Tageszeit ist die Digitalisierung in Berlin am stärksten?

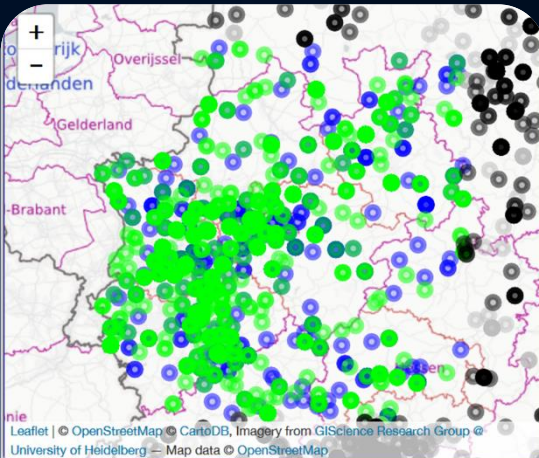
- Farbe: Morgen = blau, Mittag = grün, Nachmittag = rot, Abend = gelb





# Zu welcher Tageszeit ist die Digitalisierung in NRW am stärksten?

- Farbe: Morgen = blau, Mittag = grün, Nachmittag = rot, Abend = gelb

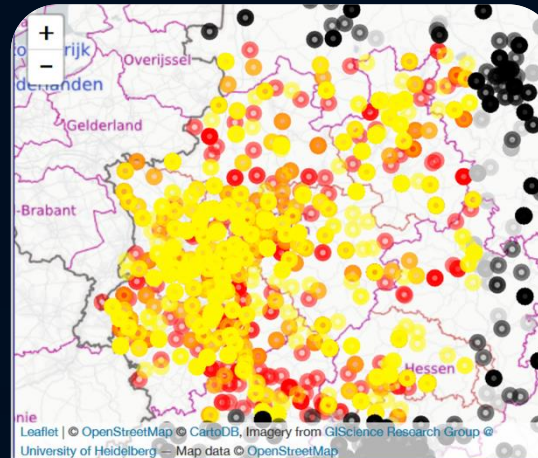


**Einwohnerzahl:**  
Deutschland: 82457031\*  
Nordrhein-Westfalen: 17890267\*

**Tweetanzahl der 1. CSV-Datei:**  
Deutschland: 8117  
Nordrhein-Westfalen: 1486

**Tweetanzahl der 2. CSV-Datei:**  
Deutschland: 9444  
Nordrhein-Westfalen: 2042

\* **Datenquelle:** Statistisches Bundesamt (Destatis), Genesis-Online, 05.11.2017; **Datenlizenz** by-2-0 ; eigene Berechnung/eigene Darstellung

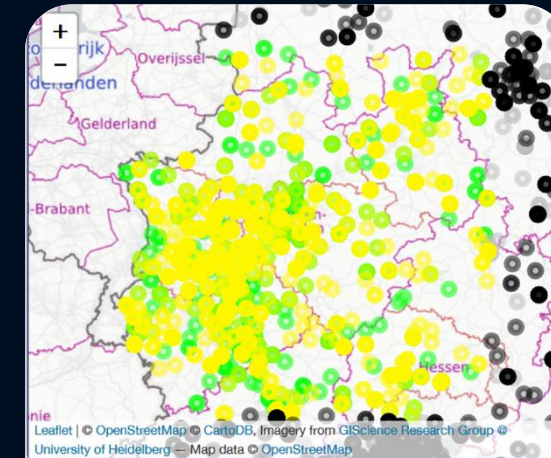


**Einwohnerzahl:**  
Deutschland: 82457031\*  
Nordrhein-Westfalen: 17890267\*

**Tweetanzahl der 1. CSV-Datei:**  
Deutschland: 10963  
Nordrhein-Westfalen: 2396

**Tweetanzahl der 2. CSV-Datei:**  
Deutschland: 12313  
Nordrhein-Westfalen: 2948

\* **Datenquelle:** Statistisches Bundesamt (Destatis), Genesis-Online, 05.11.2017; **Datenlizenz** by-2-0 ; eigene Berechnung/eigene Darstellung



**Einwohnerzahl:**  
Deutschland: 82457031\*  
Nordrhein-Westfalen: 17890267\*

**Tweetanzahl der 1. CSV-Datei:**  
Deutschland: 9444  
Nordrhein-Westfalen: 2042

**Tweetanzahl der 2. CSV-Datei:**  
Deutschland: 12313  
Nordrhein-Westfalen: 2948

\* **Datenquelle:** Statistisches Bundesamt (Destatis), Genesis-Online, 05.11.2017; **Datenlizenz** by-2-0 ; eigene Berechnung/eigene Darstellung

# Kritik

- Bounding-Box ist quadratisch, Grenzen nicht
- Farben & Reihenfolge der eingeladenen Daten beeinflussen Wahrnehmung
- Daten können nur aktiv, live empfangen werden
- Standortdaten
  - müssen vom Tweet-Ersteller explizit freigegeben werden
  - abgerufene Standortdaten können auch Ortsnamen statt Koordinaten sein
- Nur Twitter betrachten reicht nicht aus → zu wenig repräsentativ
- keine Überprüfung, ob die empfangenen Daten vollständig und korrekt sind
  - es muss der Twitter-API vertraut werden, da keine eigene Erhebung möglich ist

# Ausblick für die Forschung

- Mit dem Werkzeug lassen sich viele Fragen untersuchen
  - Untersuchungen über einen längeren Zeitraum vornehmen
- Weiterentwicklung der App
  - andere soziale Medien, wie Facebook oder Instagram, optional mit einbeziehen
  - mehr als zwei Dateien vergleichbar machen
  - Möglichkeit zum direkten Vergleich von 2 verschiedenen Bundesländern bieten
  - Speicher- & Download-Funktion der Visualisierung (Karte mit Markern) entwickeln
- Untersuchung:
  - Ist die Umsetzung mit anderen Bibliotheken besser geeignet, als mit R und Shiny?
  - Wie viel Prozent der Twitter-Nutzer in Deutschland geben ihre Standortdaten tatsächlich frei?
  - Erbringt die kostenpflichtige Version der Streaming-API mehr / exaktere Daten als die kostenfreie API?

Vielen Dank für die Aufmerksamkeit!