

Herausforderungen bei der Nutzung unstrukturierter Daten

Eine politikwissenschaftliche Perspektive

Prof. Dr. Theresa Gessler

Europa-Universität Viadrina Frankfurt (Oder)

Übersicht

Übersicht

- Unstrukturierte Daten in der Politikwissenschaft

Übersicht

- Unstrukturierte Daten in der Politikwissenschaft
- Herausforderungen

Übersicht

- Unstrukturierte Daten in der Politikwissenschaft
- Herausforderungen
- Sozialwissenschaftliche Methodenausbildung

Übersicht

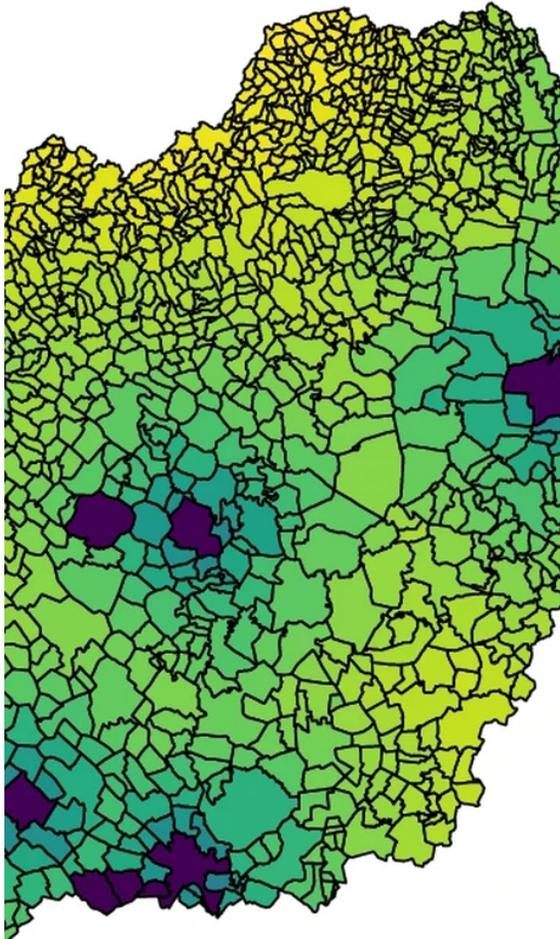
- Unstrukturierte Daten in der Politikwissenschaft
- Herausforderungen
- Sozialwissenschaftliche Methodenausbildung
- (Gleicher) Datenzugang

Übersicht

- Unstrukturierte Daten in der Politikwissenschaft
- Herausforderungen
- Sozialwissenschaftliche Methodenausbildung
- (Gleicher) Datenzugang
- Forschung

Unstrukturierte Daten in der Politikwissenschaft

Nutzung unstrukturierter Daten



Data Linkage



Digital Trace
Data

Kandidierende im Netz

Anteil der Kandidierenden pro Partei mit Accou

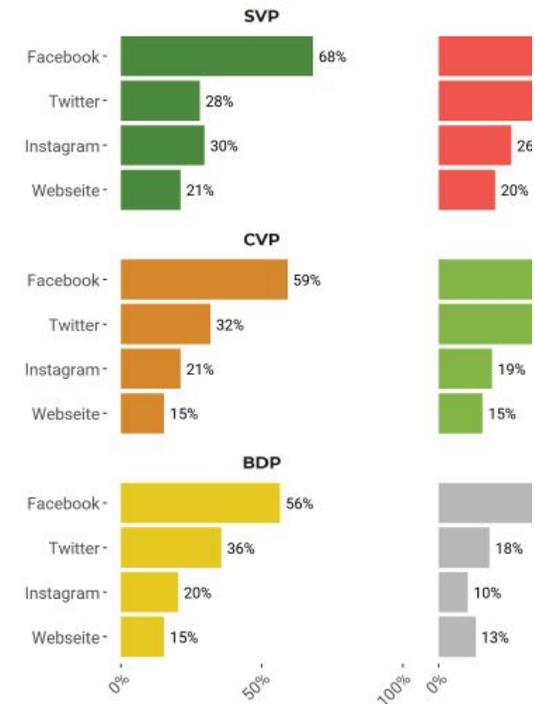
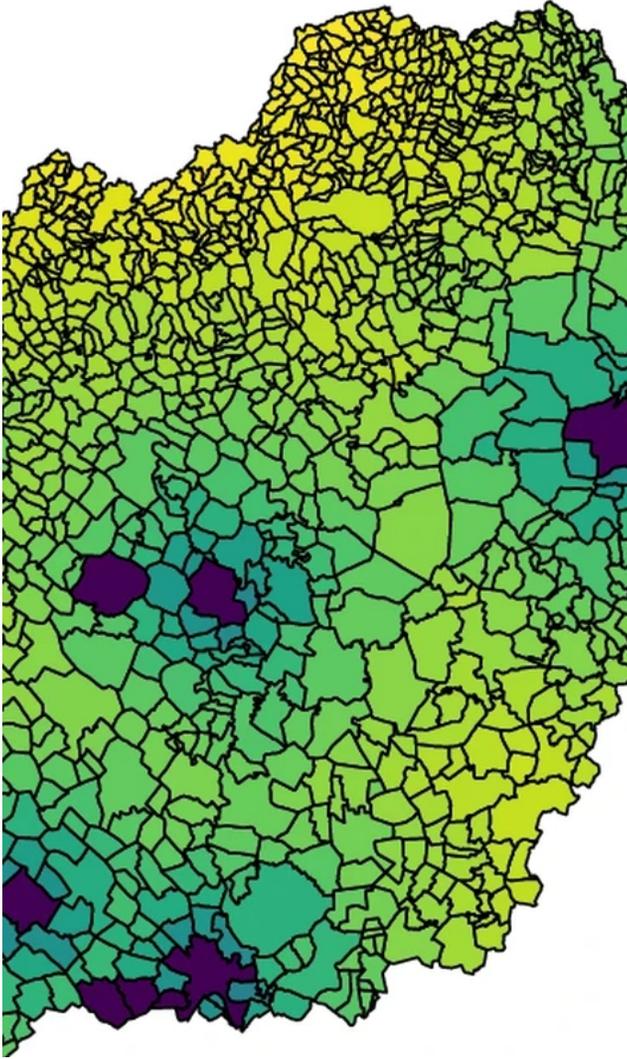


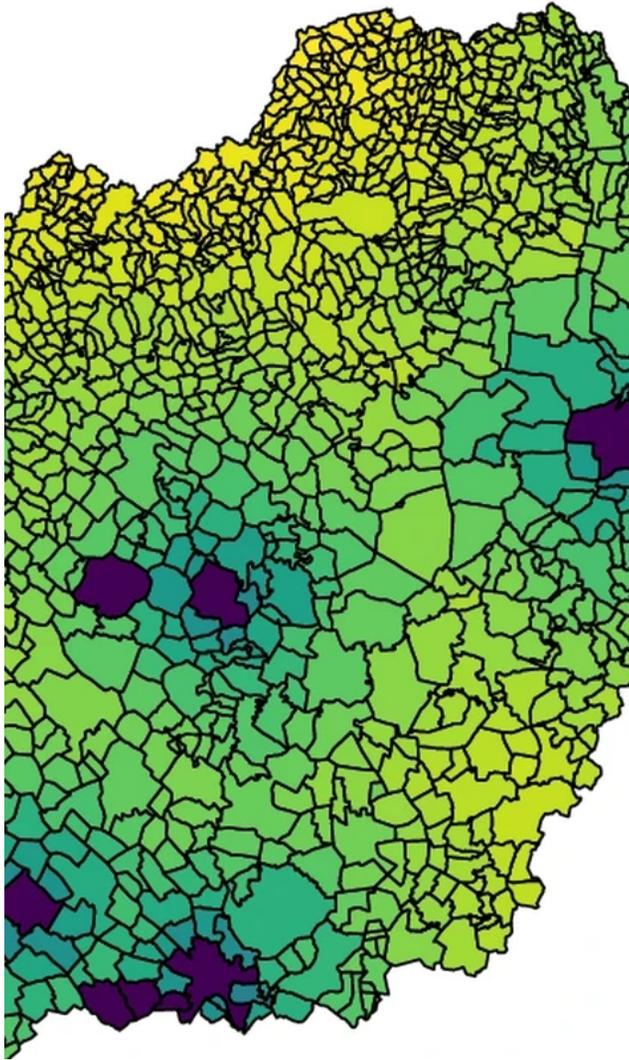
Abbildung 15: I

Digitalisierung
der Politik

Data Linkage

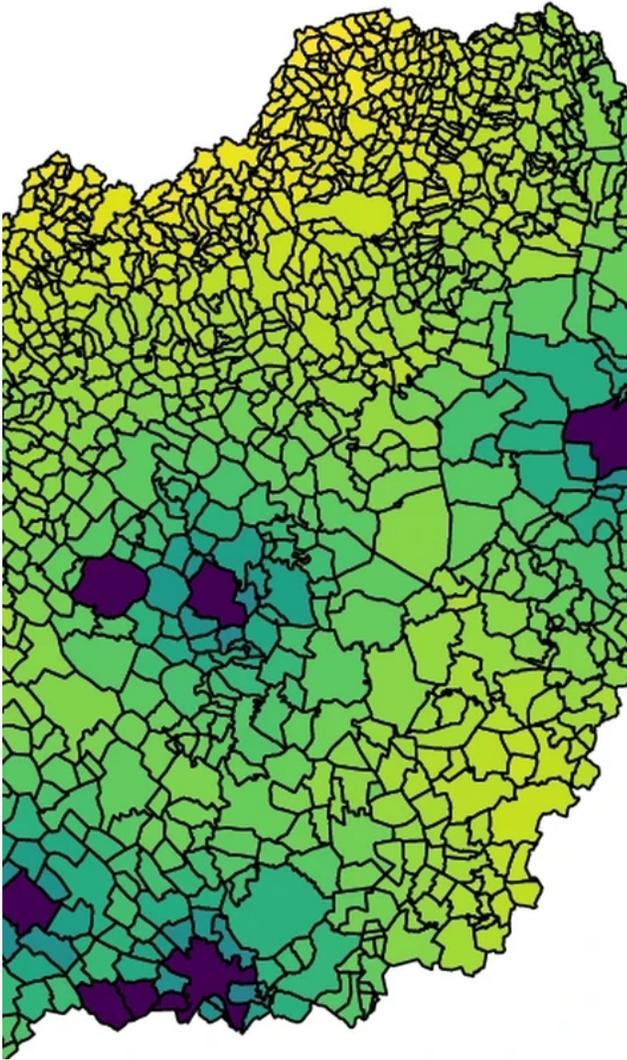


Data Linkage



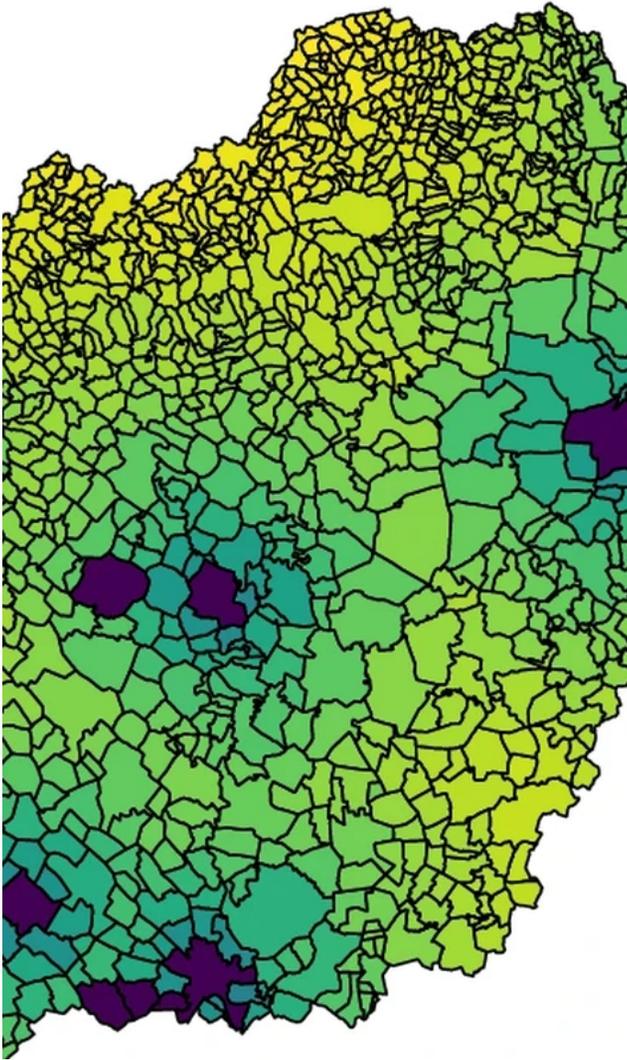
- Verknüpfung traditioneller Daten mit unstrukturierten Daten

Data Linkage



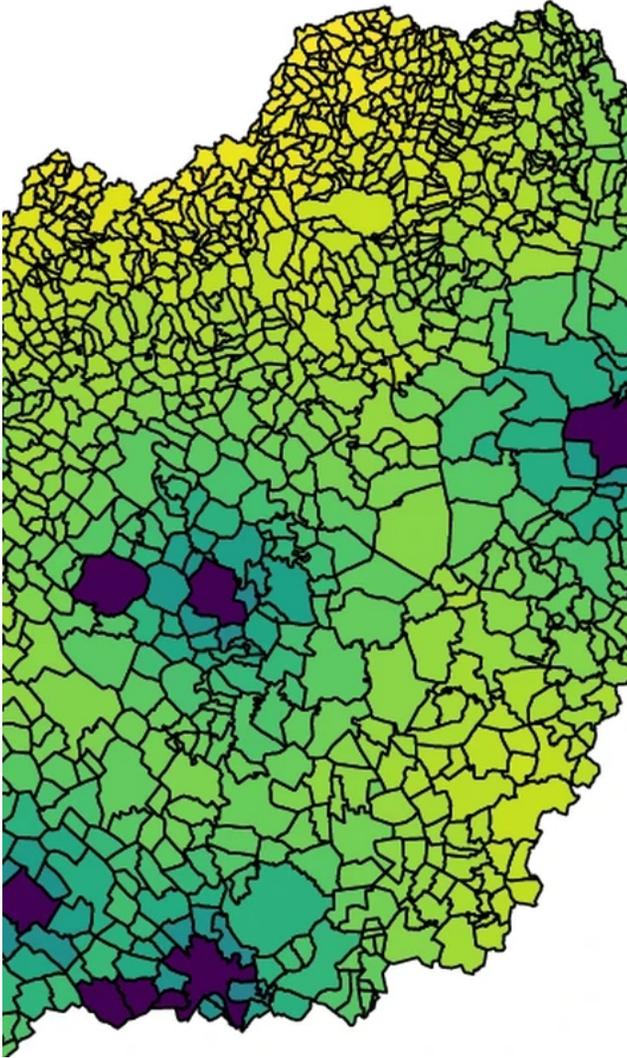
- Verknüpfung traditioneller Daten mit unstrukturierten Daten
 - z.B. Daten über geographische Region aus social media, Google street view, Suchanfragen, ...

Data Linkage



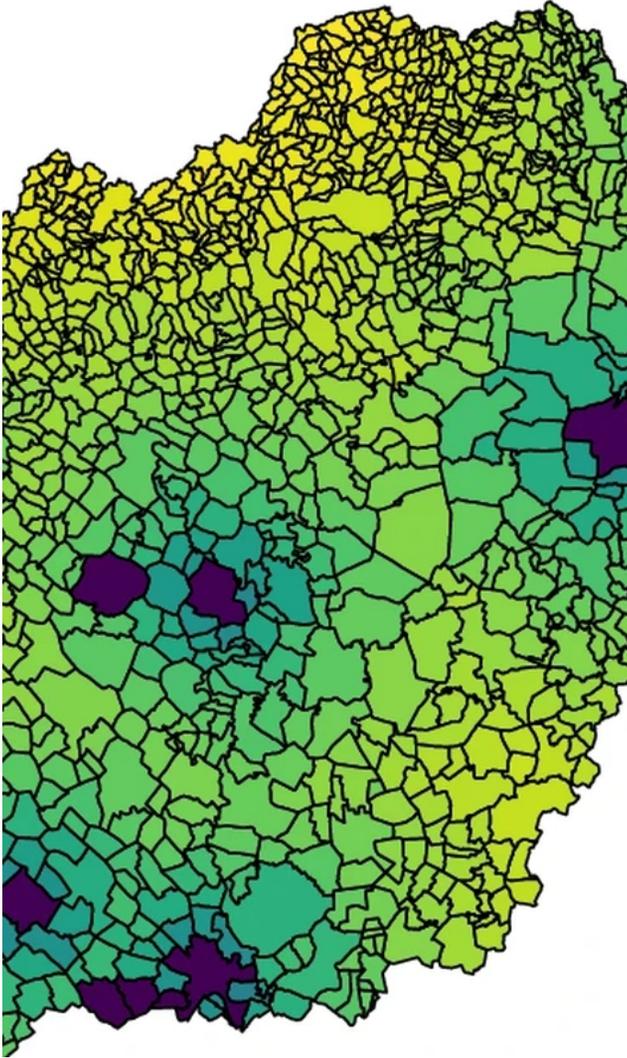
- **Verknüpfung traditioneller Daten mit unstrukturierten Daten**
 - z.B. Daten über geographische Region aus social media, Google street view, Suchanfragen, ...
 - Information über Individual- oder Registerdaten hinaus

Data Linkage



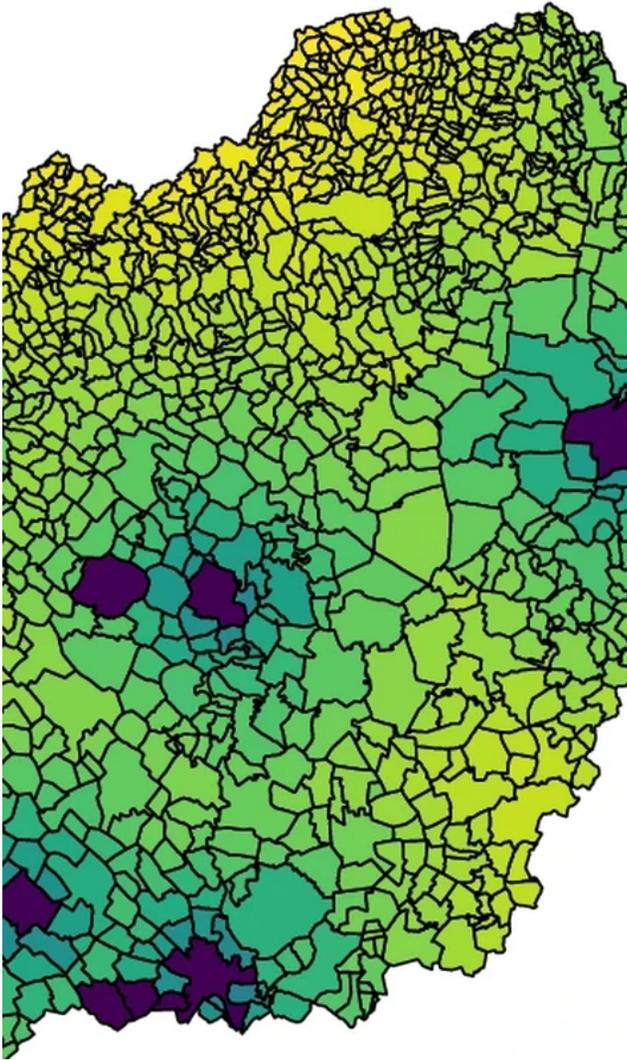
- **Verknüpfung traditioneller Daten mit unstrukturierten Daten**
 - z.B. Daten über geographische Region aus social media, Google street view, Suchanfragen, ...
 - Information über Individual- oder Registerdaten hinaus
 - immer häufiger: Data Linkage auf Ebene der Individuen

Data Linkage



- **Verknüpfung traditioneller Daten mit unstrukturierten Daten**
 - z.B. Daten über geographische Region aus social media, Google street view, Suchanfragen, ...
 - Information über Individual- oder Registerdaten hinaus
 - immer häufiger: Data Linkage auf Ebene der Individuen
- z.B. [Gessler, Theresa, Gergő Tóth, and Johannes Wachs. 'No Country for Asylum Seekers? How Short-Term Exposure to Refugees Influences Attitudes and Voting Behavior in Hungary'. Political Behavior 44, no. 4 \(December 2022\): 1813–41.](#)

Data Linkage



- **Verknüpfung traditioneller Daten mit unstrukturierten Daten**
 - z.B. Daten über geographische Region aus social media, Google street view, Suchanfragen, ...
 - Information über Individual- oder Registerdaten hinaus
 - immer häufiger: Data Linkage auf Ebene der Individuen
- z.B. [Gessler, Theresa, Gergő Tóth, and Johannes Wachs. 'No Country for Asylum Seekers? How Short-Term Exposure to Refugees Influences Attitudes and Voting Behavior in Hungary'. Political Behavior 44, no. 4 \(December 2022\): 1813–41.](#)
 - Nutzung von social media posts & online news für Pfade von Geflüchteten

Digital Trace Data



Digital Trace Data



- Nutzung von **Digital Trace Daten** für Messung des Verhalten von Bürger:innen für unterschiedliche Formen von politischer Beteiligung

Digital Trace Data



- Nutzung von **Digital Trace Daten** für Messung des Verhalten von Bürger:innen für unterschiedliche Formen von politischer Beteiligung
 - (politische) Informationssuche

Digital Trace Data



- Nutzung von **Digital Trace Daten** für Messung des Verhalten von Bürger:innen für unterschiedliche Formen von politischer Beteiligung
 - (politische) Informationssuche
 - Meinungsäußerung

Digital Trace Data



- Nutzung von **Digital Trace Daten** für Messung des Verhalten von Bürger:innen für unterschiedliche Formen von politischer Beteiligung
 - (politische) Informationssuche
 - Meinungsäußerung
 - politischer Aktivismus

Digital Trace Data



- Nutzung von **Digital Trace Daten** für Messung des Verhalten von Bürger:innen für unterschiedliche Formen von politischer Beteiligung
 - (politische) Informationssuche
 - Meinungsäußerung
 - politischer Aktivismus
- z.B. Gessler, Theresa. 'But Is She Married? Gender Bias and Users' Gendered Interest in Politicians on Wikipedia'. Working Paper.

Digital Trace Data



- Nutzung von **Digital Trace Daten** für Messung des Verhalten von Bürger:innen für unterschiedliche Formen von politischer Beteiligung
 - (politische) Informationssuche
 - Meinungsäußerung
 - politischer Aktivismus
- z.B. Gessler, Theresa. 'But Is She Married? Gender Bias and Users' Gendered Interest in Politicians on Wikipedia'. Working Paper.
 - politischer bias in Informationssuche zu politischen Themen

Digitalisierung der Politik

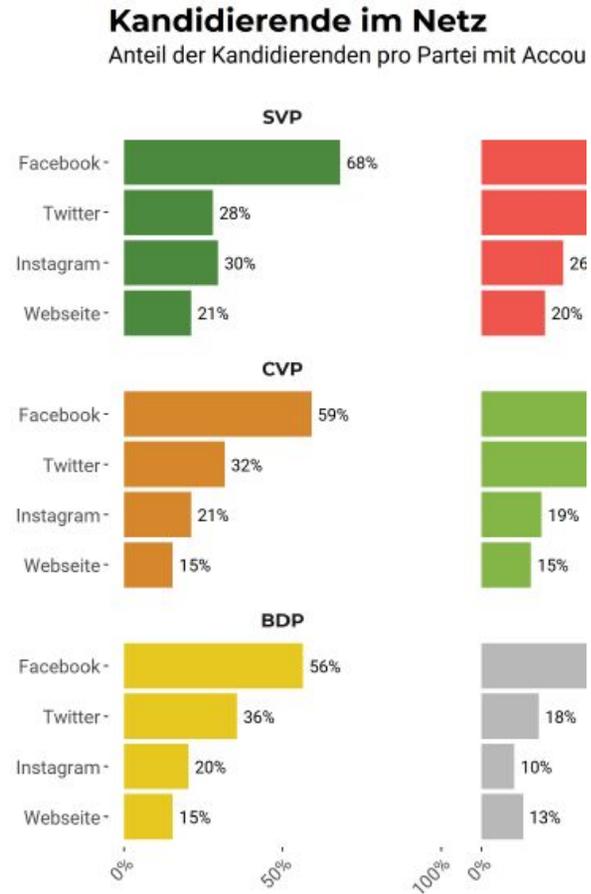


Abbildung 15: I

Digitalisierung der Politik

- Nutzung des Internets durch politische Akteure

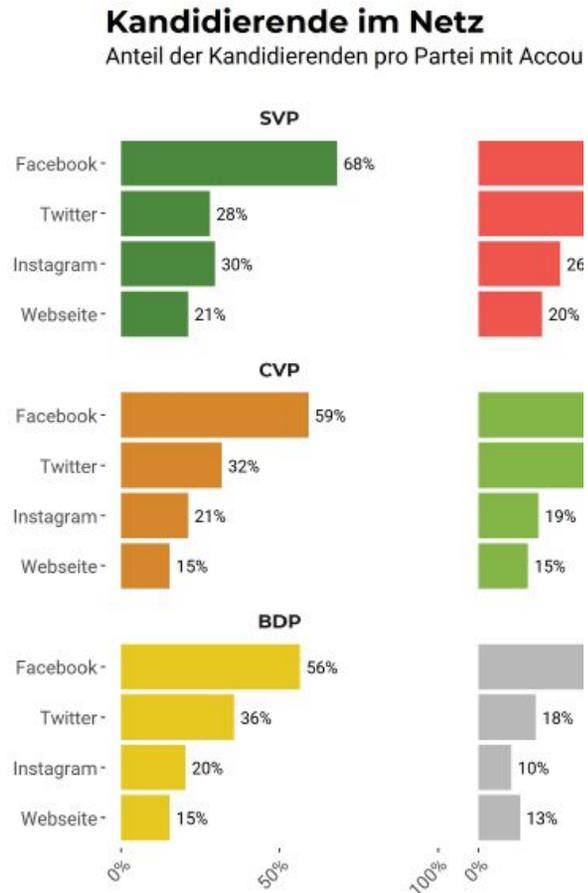


Abbildung 15: I

Digitalisierung der Politik

- Nutzung des Internets durch politische Akteure
 - Parteien

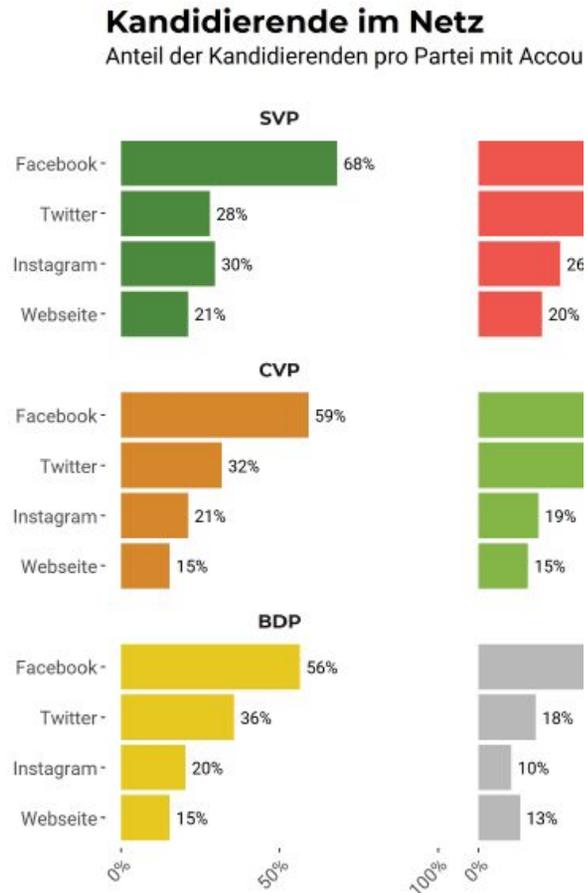


Abbildung 15: I

Digitalisierung der Politik

- Nutzung des Internets durch politische Akteure

- Parteien
- politische Kampagnen & Bewegungen

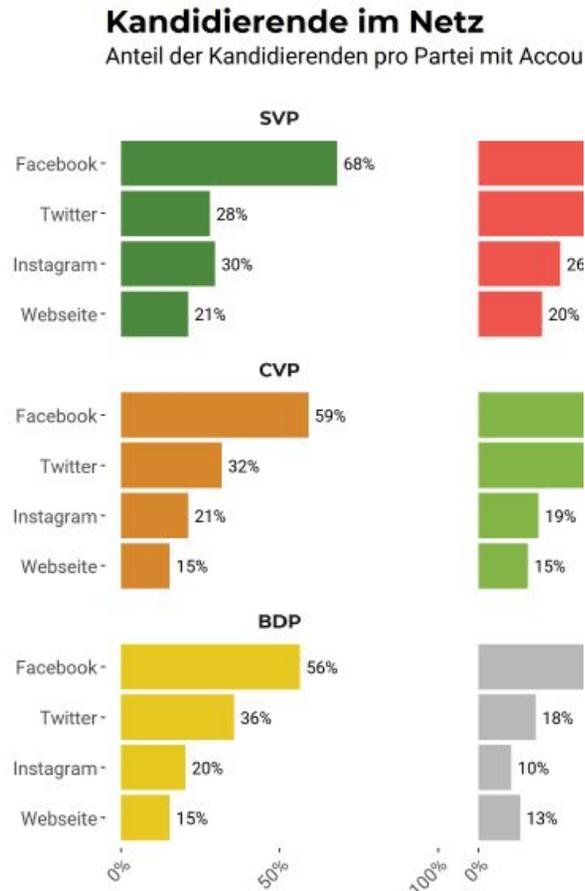


Abbildung 15: I

Digitalisierung der Politik

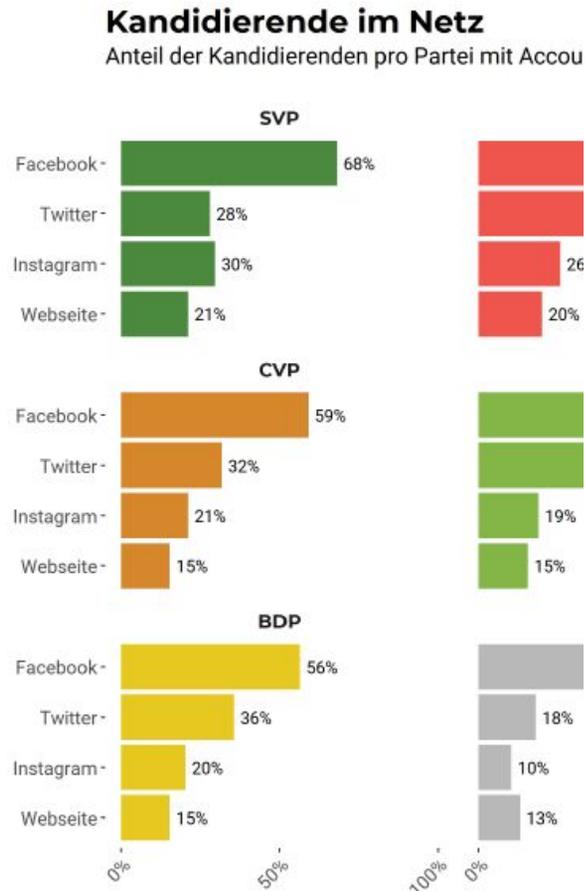


Abbildung 15: I

- Nutzung des Internets durch politische Akteure
 - Parteien
 - politische Kampagnen & Bewegungen
- Digitalisierung politischer Prozesse

Digitalisierung der Politik

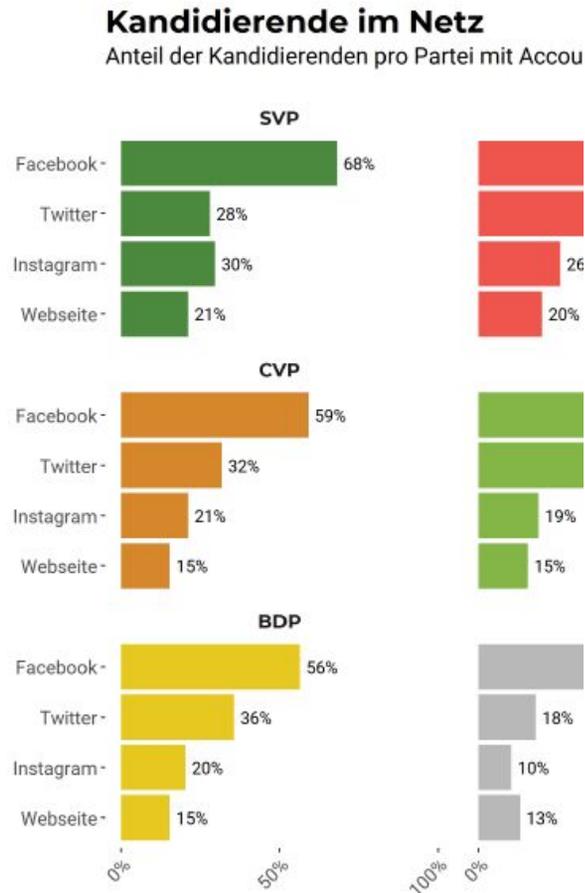


Abbildung 15: I

- Nutzung des Internets durch politische Akteure
 - Parteien
 - politische Kampagnen & Bewegungen
- Digitalisierung politischer Prozesse
 - z.B. Abstimmungen - in Deutschland aber noch weniger stark entwickelt

Digitalisierung der Politik

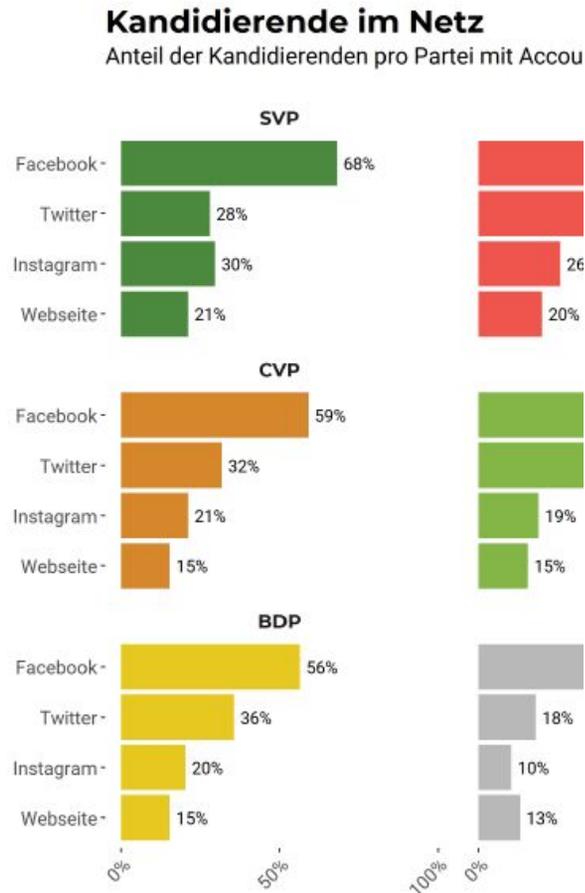


Abbildung 15: I

- Nutzung des Internets durch politische Akteure
 - Parteien
 - politische Kampagnen & Bewegungen
- Digitalisierung politischer Prozesse
 - z.B. Abstimmungen - in Deutschland aber noch weniger stark entwickelt
- z.B. [Gilardi, Fabrizio, Theresa Gessler, Maël Kubli, and Stefan Müller. 'Social Media and Political Agenda Setting'. Political Communication 39, no. 1 \(2 January 2022\): 39–60.](#)

Digitalisierung der Politik

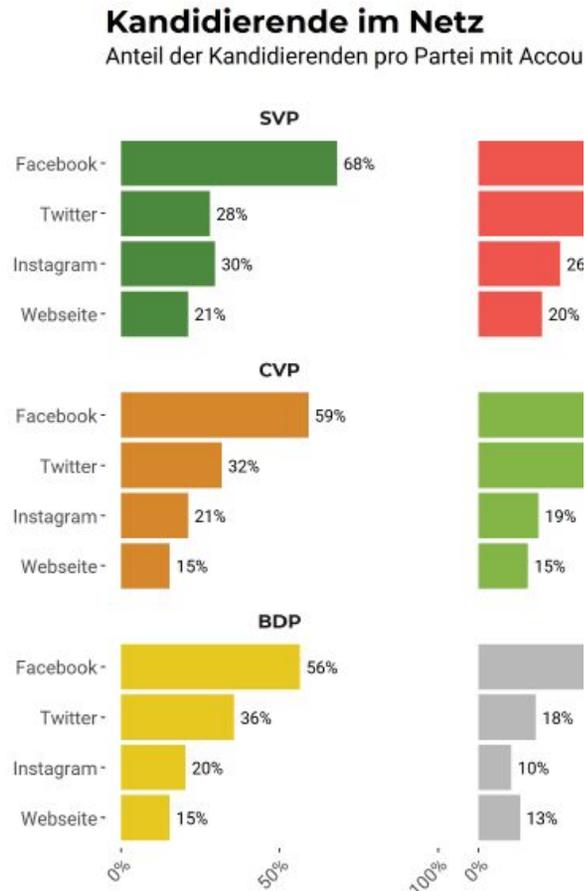


Abbildung 15: I

- Nutzung des Internets durch politische Akteure
 - Parteien
 - politische Kampagnen & Bewegungen
- Digitalisierung politischer Prozesse
 - z.B. Abstimmungen - in Deutschland aber noch weniger stark entwickelt
- z.B. [Gilardi, Fabrizio, Theresa Gessler, Maël Kubli, and Stefan Müller. 'Social Media and Political Agenda Setting'. Political Communication 39, no. 1 \(2 January 2022\): 39–60.](#)
 - weitere Forschung zu agenda-setting in der Schweiz

Digitalisierung der Politik

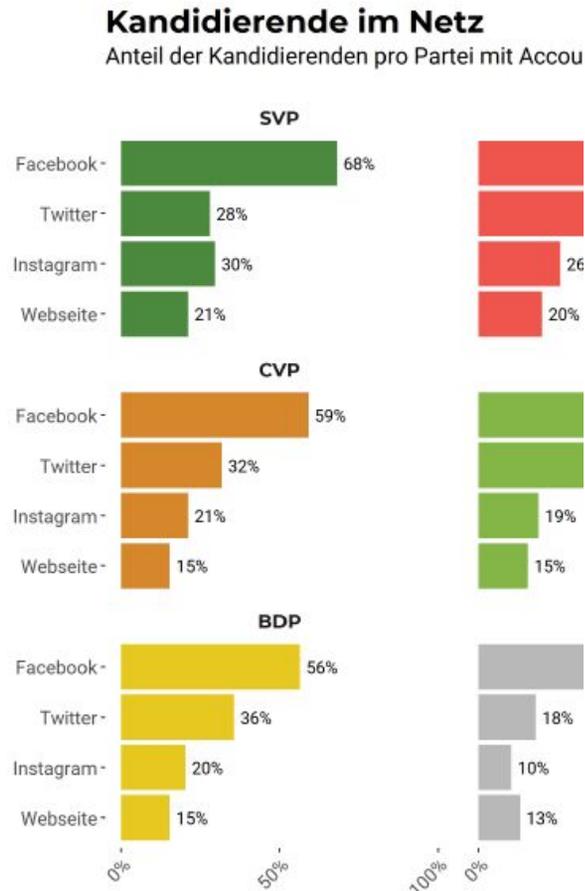


Abbildung 15: I

- Nutzung des Internets durch politische Akteure
 - Parteien
 - politische Kampagnen & Bewegungen
- Digitalisierung politischer Prozesse
 - z.B. Abstimmungen - in Deutschland aber noch weniger stark entwickelt
- z.B. [Gilardi, Fabrizio, Theresa Gessler, Maël Kubli, and Stefan Müller. 'Social Media and Political Agenda Setting'. Political Communication 39, no. 1 \(2 January 2022\): 39–60.](#)
 - weitere Forschung zu agenda-setting in der Schweiz
 - darüberhinaus: Forschung zu #MeToo & Frauenstreik

Digitalisierung der Politik: Lab Strukturen



DigDemLab

Digitalisierung der Politik: Lab Strukturen

- z.B. Digital Democracy Lab: Auswirkungen digitaler Technologie auf Politik und Demokratie



DigDemLab

Digitalisierung der Politik: Lab Strukturen



DigDemLab

- z.B. Digital Democracy Lab: **Auswirkungen digitaler Technologie auf Politik und Demokratie**
- Monitoring von (digitaler) Politik

Digitalisierung der Politik: Lab Strukturen



DigDemLab

- z.B. Digital Democracy Lab: **Auswirkungen digitaler Technologie auf Politik und Demokratie**
- Monitoring von (digitaler) Politik
 - soziale Medien: Twitter, Facebook

Digitalisierung der Politik: Lab Strukturen



DigDemLab

- z.B. Digital Democracy Lab: **Auswirkungen digitaler Technologie auf Politik und Demokratie**
- Monitoring von (digitaler) Politik
 - soziale Medien: Twitter, Facebook
 - Zeitungsartikel

Digitalisierung der Politik: Lab Strukturen



DigDemLab

- z.B. Digital Democracy Lab: **Auswirkungen digitaler Technologie auf Politik und Demokratie**
- Monitoring von (digitaler) Politik
 - soziale Medien: Twitter, Facebook
 - Zeitungsartikel
 - parlamentarische Debatten

Digitalisierung der Politik: Lab Strukturen



DigDemLab

- z.B. Digital Democracy Lab: **Auswirkungen digitaler Technologie auf Politik und Demokratie**
- Monitoring von (digitaler) Politik
 - soziale Medien: Twitter, Facebook
 - Zeitungsartikel
 - parlamentarische Debatten
 - Pressemitteilungen von Parteien und Verbänden

Digitalisierung der Politik: Lab Strukturen



DigDemLab

- z.B. Digital Democracy Lab: **Auswirkungen digitaler Technologie auf Politik und Demokratie**
- Monitoring von (digitaler) Politik
 - soziale Medien: Twitter, Facebook
 - Zeitungsartikel
 - parlamentarische Debatten
 - Pressemitteilungen von Parteien und Verbänden
- Forschung u.a. zu

Digitalisierung der Politik: Lab Strukturen



DigDemLab

- z.B. Digital Democracy Lab: **Auswirkungen digitaler Technologie auf Politik und Demokratie**
- Monitoring von (digitaler) Politik
 - soziale Medien: Twitter, Facebook
 - Zeitungsartikel
 - parlamentarische Debatten
 - Pressemitteilungen von Parteien und Verbänden
- Forschung u.a. zu
 - politische Kommunikation und öffentliche Meinung

Digitalisierung der Politik: Lab Strukturen



DigDemLab

- z.B. Digital Democracy Lab: **Auswirkungen digitaler Technologie auf Politik und Demokratie**
- Monitoring von (digitaler) Politik
 - soziale Medien: Twitter, Facebook
 - Zeitungsartikel
 - parlamentarische Debatten
 - Pressemitteilungen von Parteien und Verbänden
- Forschung u.a. zu
 - politische Kommunikation und öffentliche Meinung
 - e-government und öffentliche Verwaltung

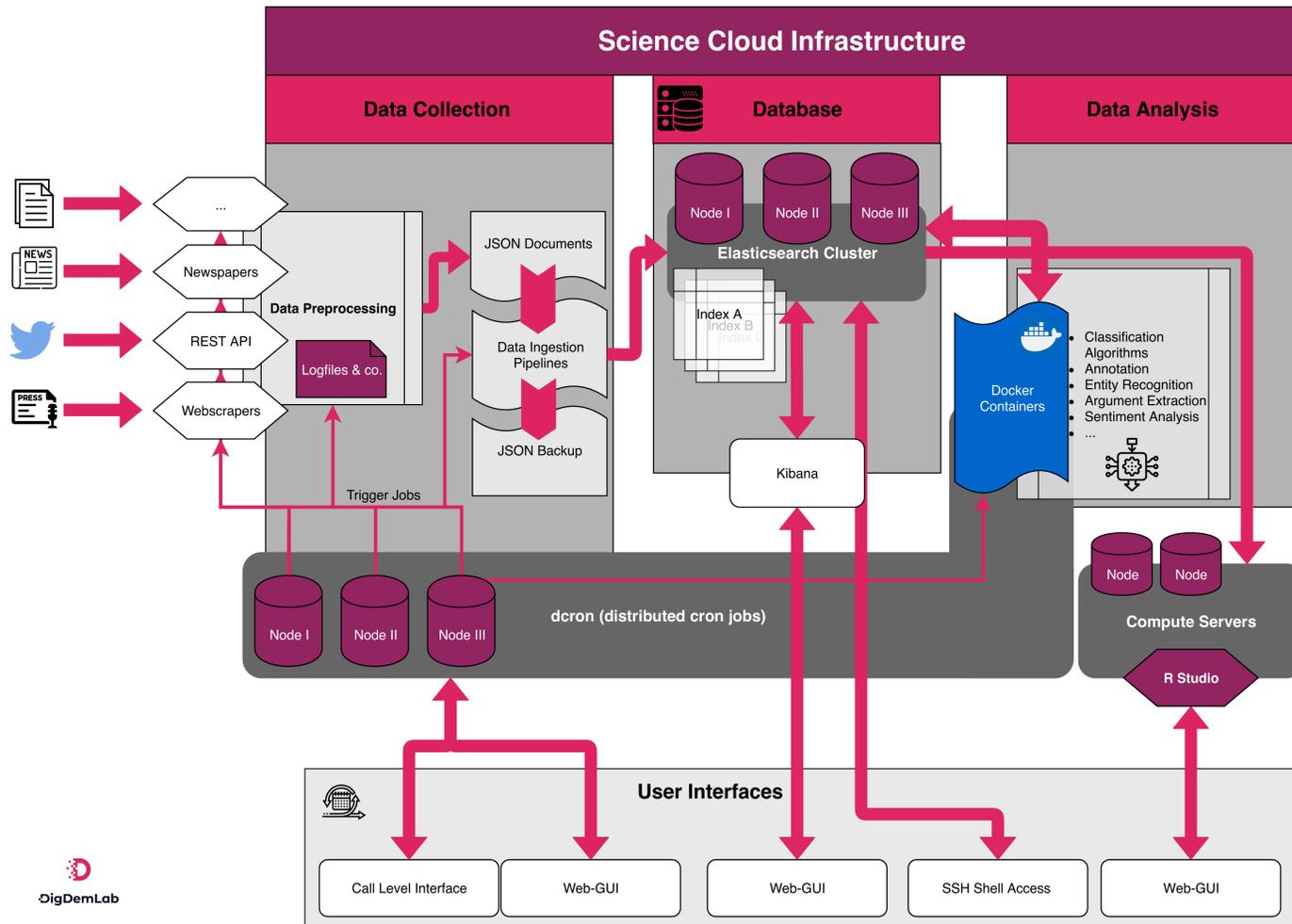
Digitalisierung der Politik: Lab Strukturen



DigDemLab

- z.B. Digital Democracy Lab: **Auswirkungen digitaler Technologie auf Politik und Demokratie**
- Monitoring von (digitaler) Politik
 - soziale Medien: Twitter, Facebook
 - Zeitungsartikel
 - parlamentarische Debatten
 - Pressemitteilungen von Parteien und Verbänden
- Forschung u.a. zu
 - politische Kommunikation und öffentliche Meinung
 - e-government und öffentliche Verwaltung
 - 'civic tech' und politische Beteiligung

Digital Democracy Lab



Herausforderungen

Herausforderungen

Methodenausbildung



Herausforderungen

Methodenausbildung



Datenzugang



Herausforderungen

Methodenausbildung



Datenzugang



Forschung

Sozialwissenschaftliche Methodenausbildung

Sozialwiss. Methodenausbildung

Sozialwiss. Methodenausbildung

- Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen
 - schnellere Innovationszyklen

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen
 - schnellere Innovationszyklen
- **Methodenkompetenz unter Lehrenden**

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen
 - schnellere Innovationszyklen
- **Methodenkompetenz unter Lehrenden**
 - keine systematische Programmierausbildung in den Sozialwissenschaften

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen
 - schnellere Innovationszyklen
- **Methodenkompetenz unter Lehrenden**
 - keine systematische Programmierausbildung in den Sozialwissenschaften
 - kaum Anreize zur Weiterbildung

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen
 - schnellere Innovationszyklen
- **Methodenkompetenz unter Lehrenden**
 - keine systematische Programmierausbildung in den Sozialwissenschaften
 - kaum Anreize zur Weiterbildung
- **keine einheitliche Methodenausbildung**

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen
 - schnellere Innovationszyklen
- **Methodenkompetenz unter Lehrenden**
 - keine systematische Programmierausbildung in den Sozialwissenschaften
 - kaum Anreize zur Weiterbildung
- **keine einheitliche Methodenausbildung**
 - Unterschiede vor & auf Doktoratsebene

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen
 - schnellere Innovationszyklen
- **Methodenkompetenz unter Lehrenden**
 - keine systematische Programmierausbildung in den Sozialwissenschaften
 - kaum Anreize zur Weiterbildung
- **keine einheitliche Methodenausbildung**
 - Unterschiede vor & auf Doktoratsebene
 - Stückwerk-Charakter des Angebots

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen
 - schnellere Innovationszyklen
- **Methodenkompetenz unter Lehrenden**
 - keine systematische Programmierausbildung in den Sozialwissenschaften
 - kaum Anreize zur Weiterbildung
- **keine einheitliche Methodenausbildung**
 - Unterschiede vor & auf Doktoratsebene
 - Stückwerk-Charakter des Angebots
- **Schwierigkeiten insbesondere bei der Betreuung**

Sozialwiss. Methodenausbildung

- **Ausgangslage: erhöhte & weiter wachsende Anforderungen an Methodenkenntnisse**
 - breiteres Spektrum an Methoden & Sprachen /Plattformen
 - schnellere Innovationszyklen
- **Methodenkompetenz unter Lehrenden**
 - keine systematische Programmierausbildung in den Sozialwissenschaften
 - kaum Anreize zur Weiterbildung
- **keine einheitliche Methodenausbildung**
 - Unterschiede vor & auf Doktoratsebene
 - Stückwerk-Charakter des Angebots
- **Schwierigkeiten insbesondere bei der Betreuung**

→ Herausforderungen bei der **Bildung von Kompetenzen im Umgang mit unstrukturierte Daten**

(Gleicher) Datenzugang

(Gleicher) Datenzugang

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden
 - exklusive Zusammenarbeiten

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden
 - exklusive Zusammenarbeiten
 - Vertragsvoraussetzungen & Bürokratie

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden
 - exklusive Zusammenarbeiten
 - Vertragsvoraussetzungen & Bürokratie
 - Status-Voraussetzungen

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden
 - exklusive Zusammenarbeiten
 - Vertragsvoraussetzungen & Bürokratie
 - Status-Voraussetzungen
- **technische Voraussetzungen**

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden
 - exklusive Zusammenarbeiten
 - Vertragsvoraussetzungen & Bürokratie
 - Status-Voraussetzungen
- **technische Voraussetzungen**
 - mangelnder Zugang zu high performance computing in den Sozialwissenschaften an vielen Standorten

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden
 - exklusive Zusammenarbeiten
 - Vertragsvoraussetzungen & Bürokratie
 - Status-Voraussetzungen
- **technische Voraussetzungen**
 - mangelnder Zugang zu high performance computing in den Sozialwissenschaften an vielen Standorten
 - Größe der Datensätze

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden
 - exklusive Zusammenarbeiten
 - Vertragsvoraussetzungen & Bürokratie
 - Status-Voraussetzungen
- **technische Voraussetzungen**
 - mangelnder Zugang zu high performance computing in den Sozialwissenschaften an vielen Standorten
 - Größe der Datensätze
 - Struktur der Datensätze

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden
 - exklusive Zusammenarbeiten
 - Vertragsvoraussetzungen & Bürokratie
 - Status-Voraussetzungen
- **technische Voraussetzungen**
 - mangelnder Zugang zu high performance computing in den Sozialwissenschaften an vielen Standorten
 - Größe der Datensätze
 - Struktur der Datensätze
- **Kosten** für API Zugänge, high performance computing etc.

(Gleicher) Datenzugang

- Zugang ist häufig an (größere) **Institutionen** gebunden
 - exklusive Zusammenarbeiten
 - Vertragsvoraussetzungen & Bürokratie
 - Status-Voraussetzungen
- **technische Voraussetzungen**
 - mangelnder Zugang zu high performance computing in den Sozialwissenschaften an vielen Standorten
 - Größe der Datensätze
 - Struktur der Datensätze
- **Kosten** für API Zugänge, high performance computing etc.

→ Herausforderung: **Einschränkungen für Wissenschaftler:innen in der Qualifizierungsphase und an kleineren Universitäten**

Forschung

Forschung

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung
 - skalierbare Erhebung & Analyse

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung
 - skalierbare Erhebung & Analyse
 - (scheinbare) Einheitlichkeit

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung
 - skalierbare Erhebung & Analyse
 - (scheinbare) Einheitlichkeit
- **Probleme mit Messung & Repräsentativität** (→ TSE-Framework)

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung
 - skalierbare Erhebung & Analyse
 - (scheinbare) Einheitlichkeit
- **Probleme mit Messung & Repräsentativität** (→ TSE-Framework)
 - Repräsentativität als Voraussetzung in vielen Bereichen der Politikwissenschaft

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung
 - skalierbare Erhebung & Analyse
 - (scheinbare) Einheitlichkeit
- **Probleme mit Messung & Repräsentativität** (→ TSE-Framework)
 - Repräsentativität als Voraussetzung in vielen Bereichen der Politikwissenschaft
 - Messung und Konzeptionalisierung: in-progress

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung
 - skalierbare Erhebung & Analyse
 - (scheinbare) Einheitlichkeit
- **Probleme mit Messung & Repräsentativität** (→ TSE-Framework)
 - Repräsentativität als Voraussetzung in vielen Bereichen der Politikwissenschaft
 - Messung und Konzeptionalisierung: in-progress
- **Versteckte Herausforderungen durch unklaren Prozess der Datengenerierung**

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung
 - skalierbare Erhebung & Analyse
 - (scheinbare) Einheitlichkeit
- **Probleme mit Messung & Repräsentativität** (→ TSE-Framework)
 - Repräsentativität als Voraussetzung in vielen Bereichen der Politikwissenschaft
 - Messung und Konzeptionalisierung: in-progress
- **Versteckte Herausforderungen durch unklaren Prozess der Datengenerierung**
 - ungleiche Adoptionsraten & schwierige Datengrundlage (für Politik & Gesellschaft)

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung
 - skalierbare Erhebung & Analyse
 - (scheinbare) Einheitlichkeit
- **Probleme mit Messung & Repräsentativität** (→ TSE-Framework)
 - Repräsentativität als Voraussetzung in vielen Bereichen der Politikwissenschaft
 - Messung und Konzeptionalisierung: in-progress
- **Versteckte Herausforderungen durch unklaren Prozess der Datengenerierung**
 - ungleiche Adoptionsraten & schwierige Datengrundlage (für Politik & Gesellschaft)
 - Einfluss von Algorithmen auf Daten (z.B. targeting)

Forschung

- **Attraktivität unstrukturierter Daten** (z.B. Webdaten) für vergleichende Forschung
 - skalierbare Erhebung & Analyse
 - (scheinbare) Einheitlichkeit
- **Probleme mit Messung & Repräsentativität** (→ TSE-Framework)
 - Repräsentativität als Voraussetzung in vielen Bereichen der Politikwissenschaft
 - Messung und Konzeptionalisierung: in-progress
- **Versteckte Herausforderungen durch unklaren Prozess der Datengenerierung**
 - ungleiche Adoptionsraten & schwierige Datengrundlage (für Politik & Gesellschaft)
 - Einfluss von Algorithmen auf Daten (z.B. targeting)

→ Herausforderung: **Entwicklung der Forschung von Enthusiasmus zum Forschungsprogramm**

**Danke für die
Aufmerksamkeit!**

gessler@europa-uni.de | www.theresagessler.eu

