

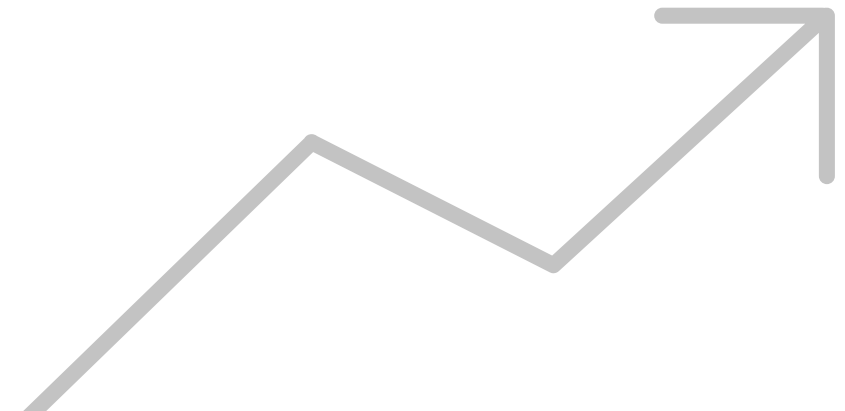
Mobile Network Data in Official Statistics

BERD Academy – Data Challenge: Mobile Phone Data

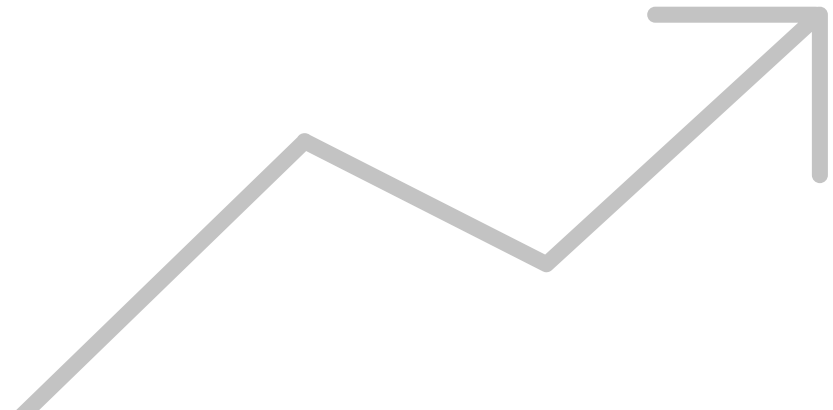
Munich, 12. November 2024

Lorenz Ade

Federal Statistical Office of Germany



Past Research at Destatis



EXSTAT

» Research into mobile network data since 2017

» Cooperations with Telekom (T-Systems) and Telefónica (Teralytics)

» Workshop reports:
https://www.destatis.de/EN/Service/EXSTAT/_node.html



Motivation

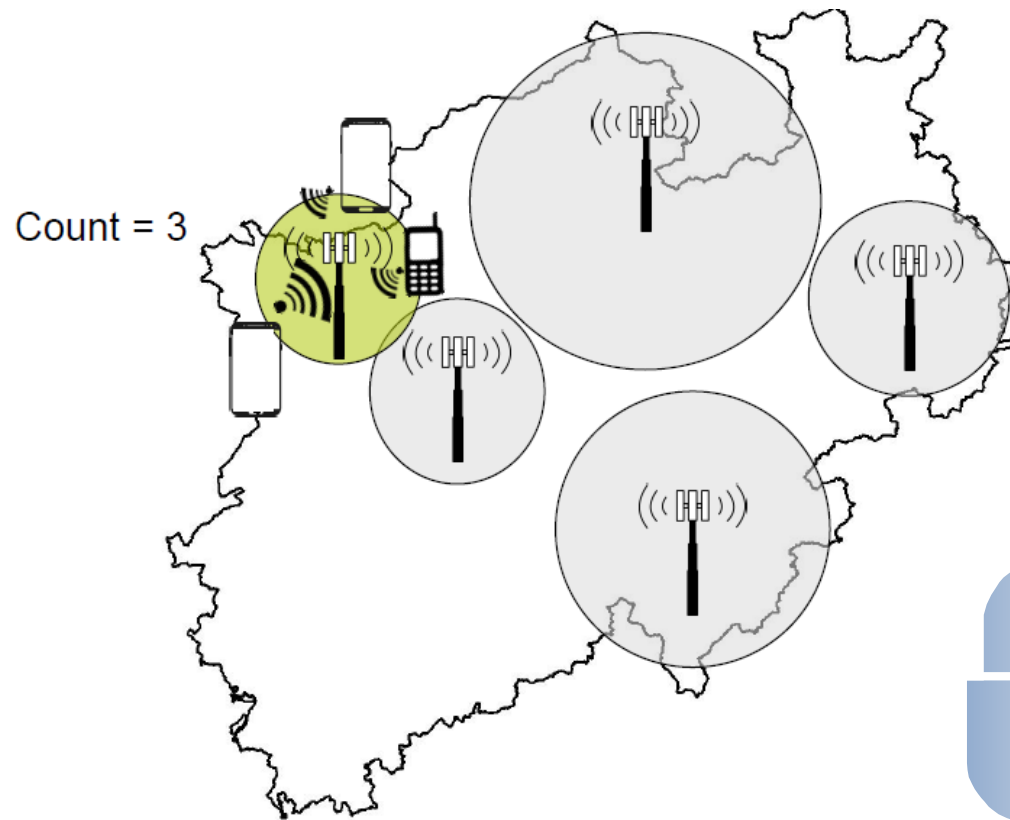
» Improvement of official statistics

- » Faster reporting, Better coverage, increased quality, (decreased burden on survey participants)
- » E.g. Population, Mobility, Tourism

» New statistical products

- » E.g. Day-Night Populations

Signal-Based Event Data



Mobile **network** data
≠ Mobile **phone** data

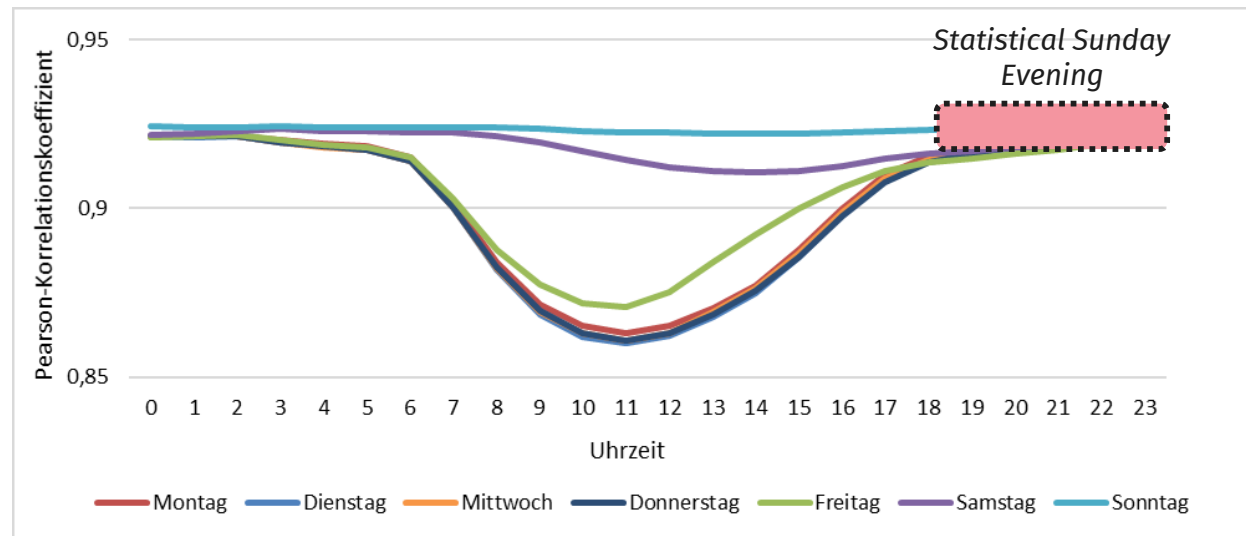


Anonymized & Aggregated

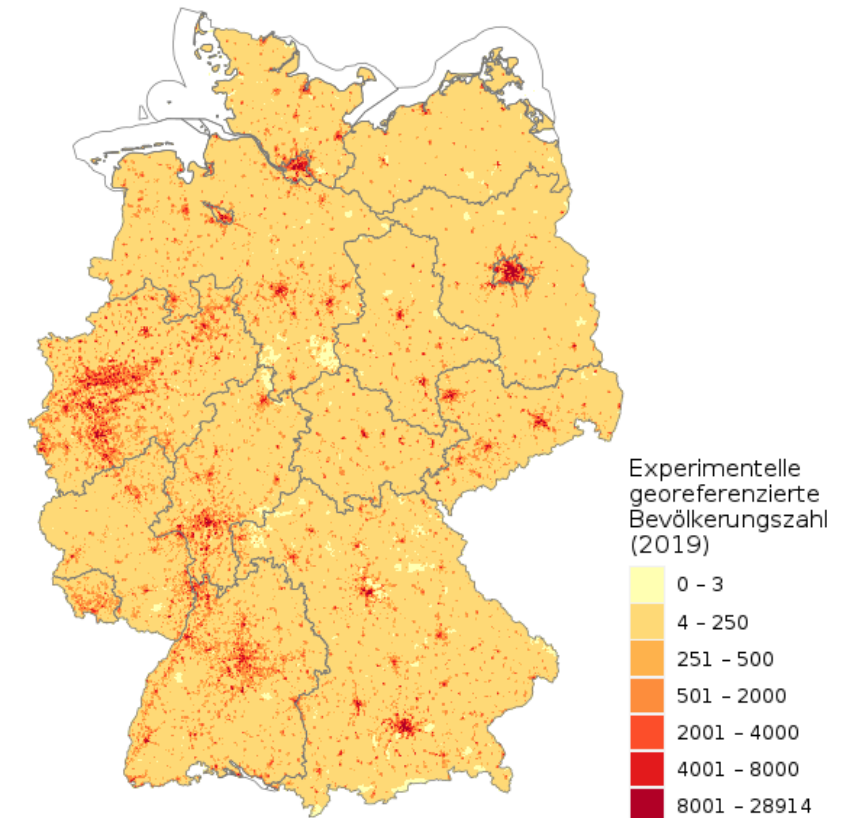
Quelle: IT.NRW

Population Statistics

Statistical correlation between census data and MNO data



Quelle: eigene Berechnung | © Teralytics



Quelle: eigene Berechnung | © Teralytics | © GeoBasis-DE / BKG, 2021
Siehe auch [Experimentelle georeferenzierte Bevölkerungszahl auf Basis der Bevölkerungsfortschreibung und Mobilfunkdaten](#) auf EXSTAT (aufgerufen am 23.05.2023)

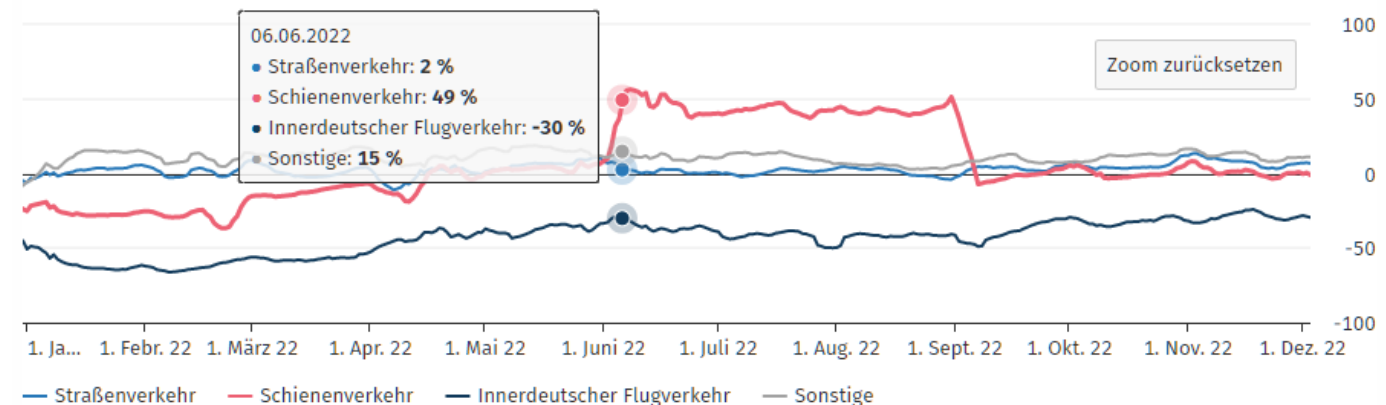
Population Mobility and „9-Euro“-Ticket

» **Goal and Result:** Findings on the impact of political measures on the mobility behavior of the population, e.g. corona pandemic or “9-euro ticket”

» **Method:** Changes in mobility behaviour in relation to a reference timeframe

Veränderung der Mobilität nach Verkehrsträgern gegenüber 2019

Ab 30 Kilometern Distanz, in %, 7-Tage-Durchschnitt



Hinweis: Für den 7-Tage-Durchschnitt werden nur die im 7-Tage-Fenster vorhandenen Daten zur Durchschnittsbildung berücksichtigt. Datenlücken entstehen i.d.R. aufgrund von technischen Problemen beim Mobilfunkanbieter. Quellen: eigene Berechnung | © Teralytics

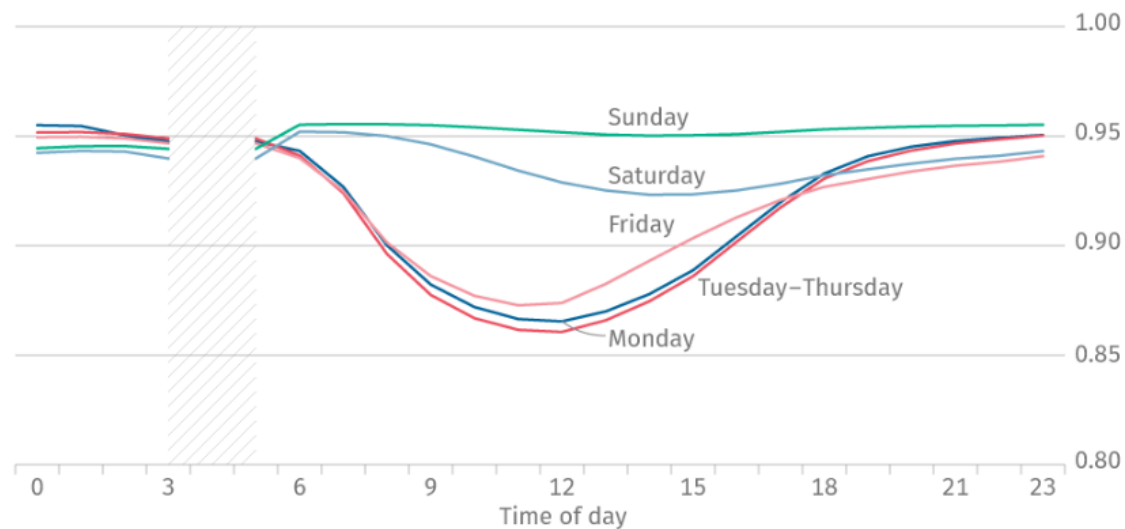
© Statistisches Bundesamt (Destatis), 2022

Siehe: <https://www.destatis.de/DE/Service/EXDAT/Datensaetze/mobilitaetsindikatoren-mobilfunkdaten.html>

Structural Comparison of two MNO's

Relationship between Telefónica + Telekom mobile network activities and 2011 Census population figures

As Pearson correlation coefficient by statistical day and time of day



Segment: value axis not starting at 0

*Please note: As the evaluation quality is not as good for hours of the night as for daytime hours, Telefónica Deutschland does not provide any mobile network data for the time between 4 and 5 a.m. Therefore, it is not possible to merge and show any mobile network data for that time span.

© Statistisches Bundesamt (Destatis), 2021

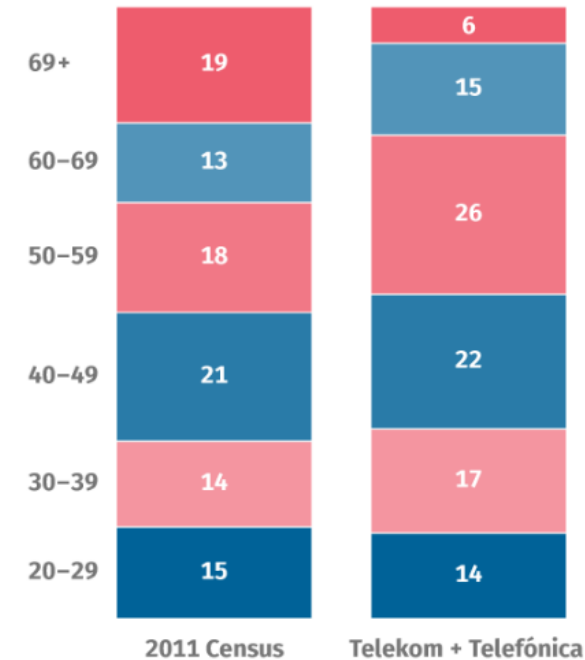
Source: 2011 Census, T-Systems/Telefónica Deutschland

© Statistisches Bundesamt (Destatis), 2021

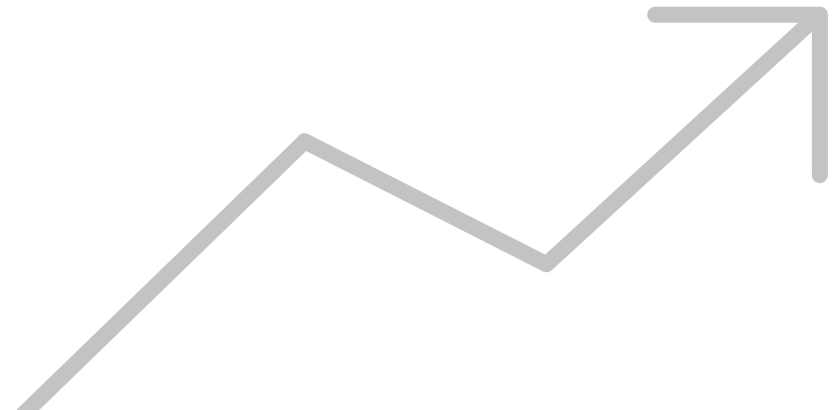
Source: 2011 Census, T-Systems/Telefónica Deutschland

Distribution by age group from the age of 20 according to the 2011 Census and contract customers of mobile network operators

Percent



AnigeD



Research Project „AnigeD“

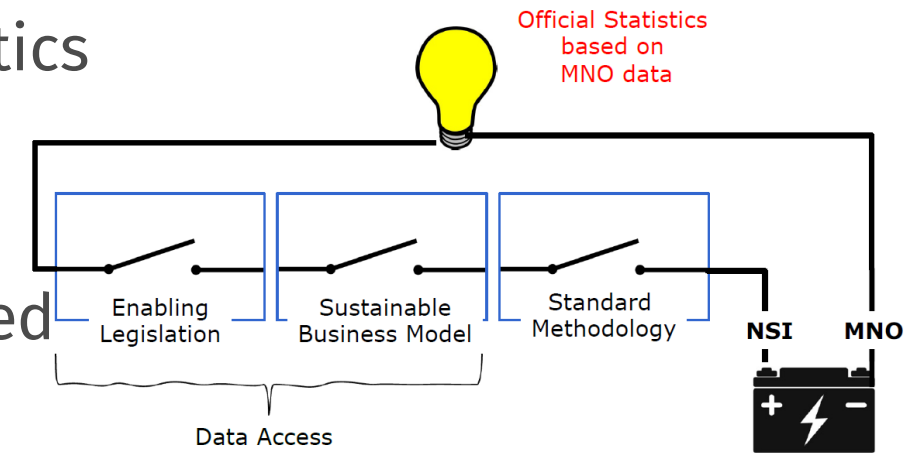
» **Big Picture:** Move MNO data into official statistics

» Open the black box of MNO data processing

» Experimental statistics are based on aggregated products

» Methods employed by MNO's are confidential

- E.g. local market shares, network maps

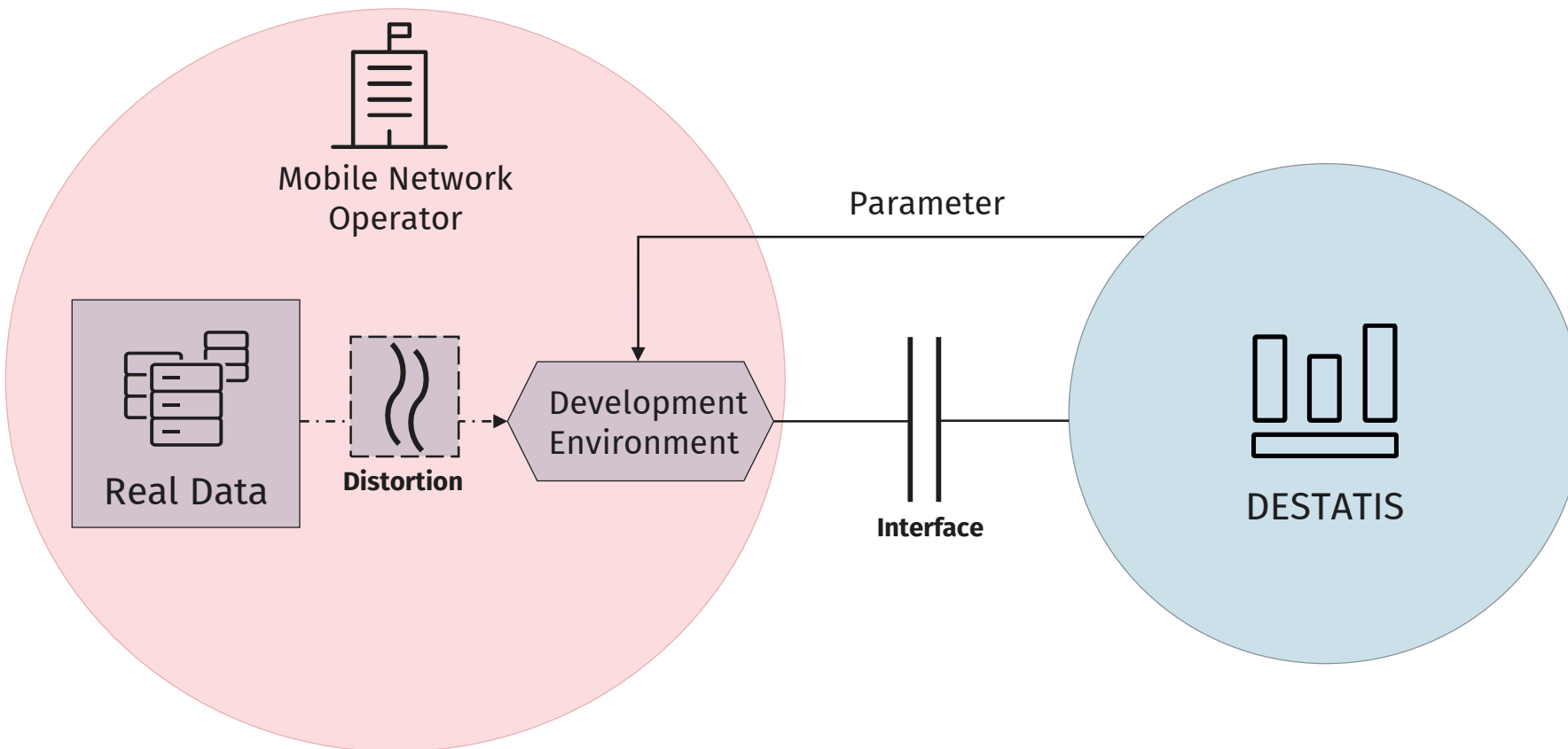


Source: Position paper of TF MNO

Goals of Research Project „AnigeD“

- » Implement reference pipeline developed in ESSnet projects:
 - » Big Data II, Multi-MNO (ongoing), MNO Minds (ongoing)
- » Identify anonymization needs
 - » Especially in relation to German data protection laws
 - » Maintain public trust in official statistics
- » Derive infrastructure for future production
 - » Interfacing two protected IT-systems is not trivial

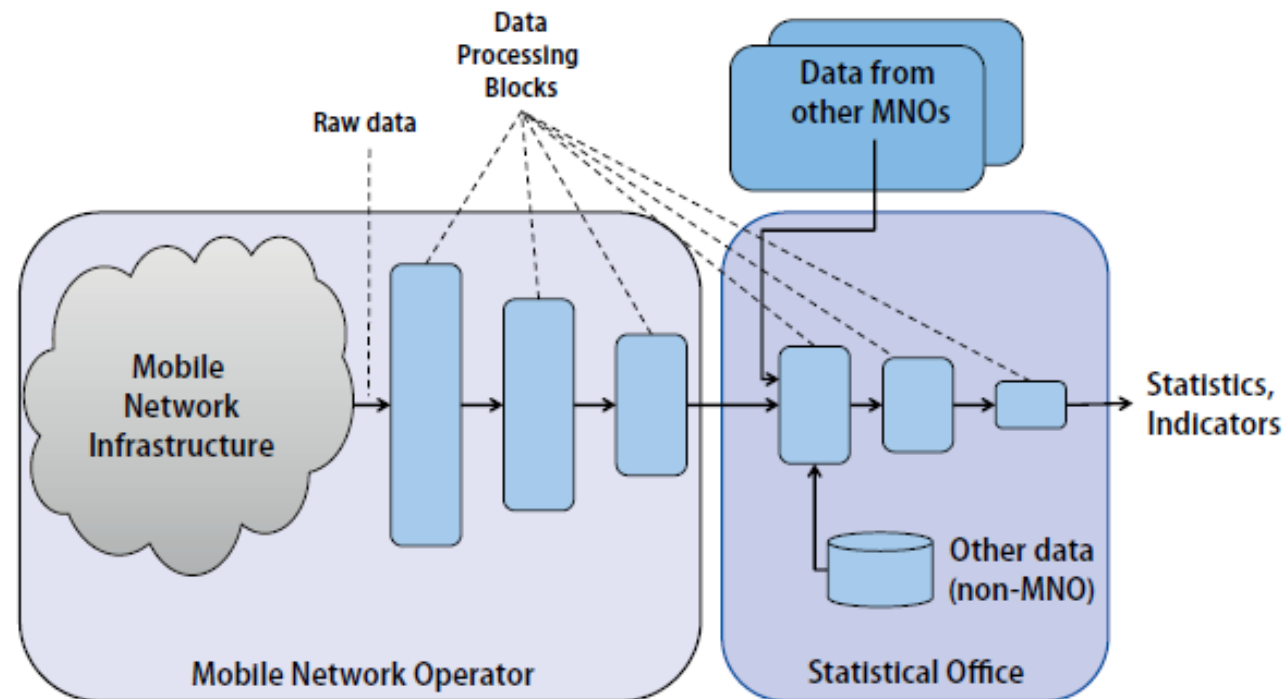
Technical Concept



» Data processing at the provider

» Development and documentation of own methods

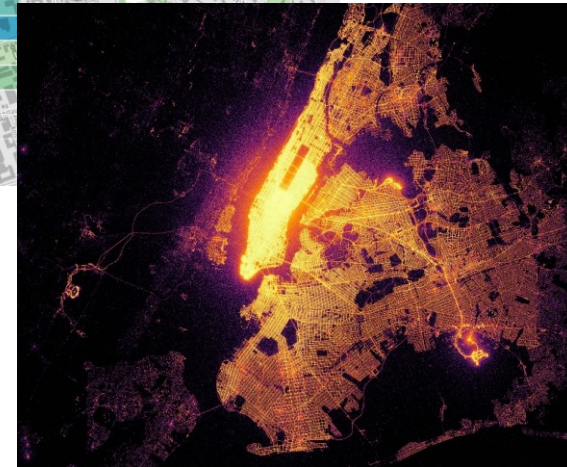
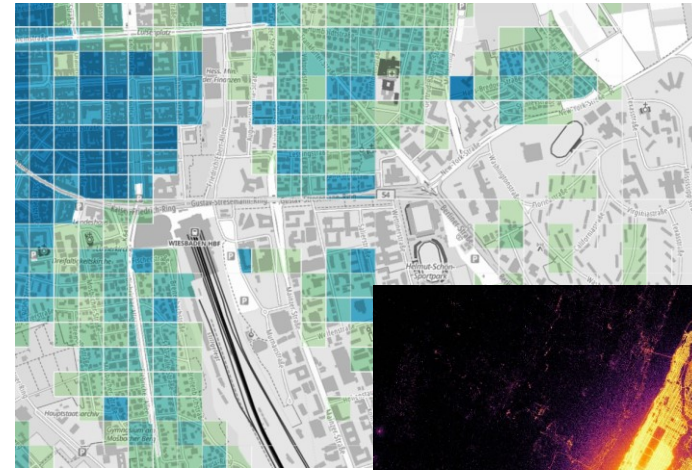
Reference Pipeline



Eurostat (2023). Reusing Mobile Network Operator data for Official Statistics: the case for a common methodological framework for the European Statistical System

Anonymization

- » MNO data is potentially sensitive
 - » Destatis needs to ensure robust statistical confidentiality to maintain trust in official statistics
- » Challenges:
 - » Scalability of anonymization methods
 - » Dynamic data is new for statistical confidentiality



Grafik: Ravi Shekar (2017). If Taxi Trips were Fireflies: 1.3 Billion NYC Taxi Trips Plotted. <https://towardsdatascience.com/if-taxi-trips-were-fireflies-1-3-billion-nyc-taxi-trips-plotted-b34e89f96cfa>. Abgerufen am 09.09.2024

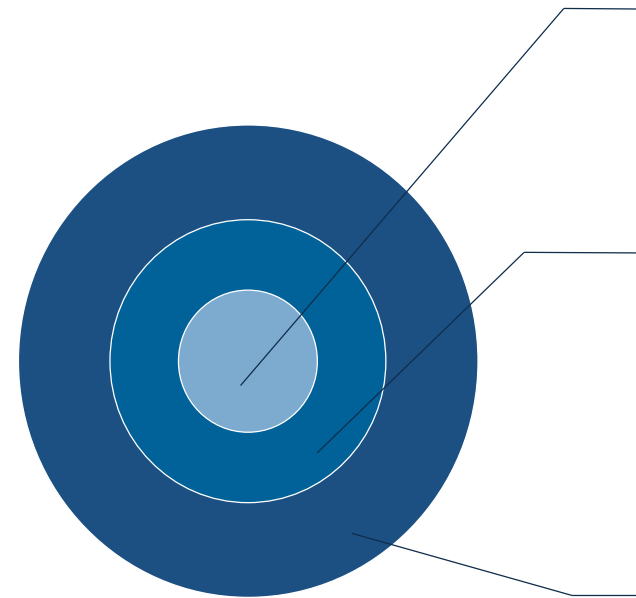
Screenshot. Zensusatlas (2024). <https://atlas.zensus2022.de/>. Abgerufen am 09.09.2024

Research Cluster AnigeD



Anonymization for Integrated and Georeferenced Data

- » Evaluation of anonymization methods using legal criteria
- » Potential of anonymization by synthetic data
- » Anonymization of georeferenced data
- » Testing of software tools for the efficient analysis and provision of anonymized georeferenced data



5 Project partners:

Destatis as cluster coordinator

6 Associated partners:

e.g. T-Systems International GmbH

Other research projects:

AnGer, DARIA, GANGES

Contact

Lorenz Ade

lorenz.ade@destatis.de

www.destatis.de

www.destatis.de/kontakt

