Data science meets sociology
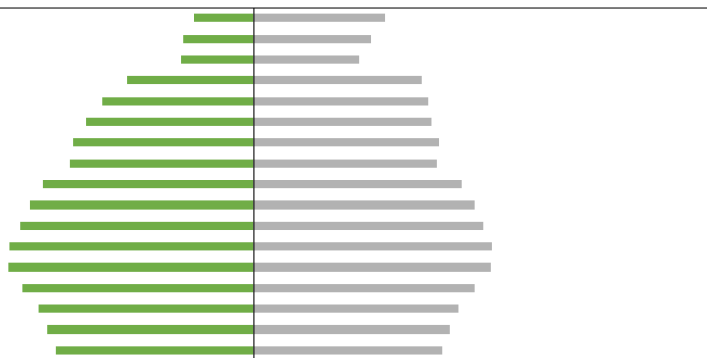# Analytical challenges from research on ageing

Martina Brandt
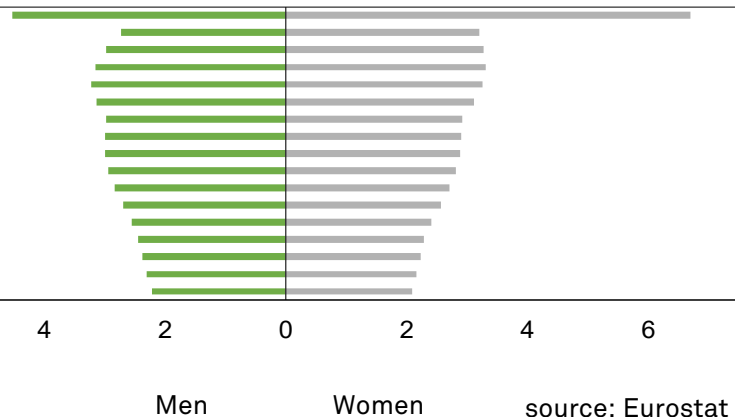
Sozialstruktur&Soziologie
alternder Gesellschaften

27.05.2021

EU27 1990

EU27 2050

4 2 0 2 4 6

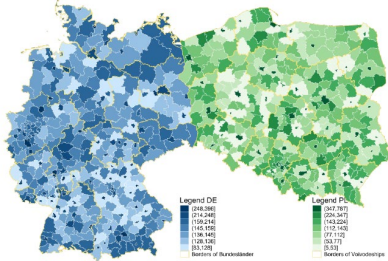Men   Women   source: Eurostat

- Rising support needs, decreasing potentials
- Rising inequalities (?)

- Health and social systems under pressure
- Discussion about efficient use of resources
  - Healthy, productive, active, successful…. ageing
  - social networks, communities, volunteering, informal care

- (How) can we influence this?
  - Assess contextual influences & life course developments
  - International/regional/temporal variation
  - (quasi-) experiments, «natural laboratory» and actual interventions
  - Main challenge: identify causalities!

Two examples from ageing research

- „broad" data
- Survey of Health, Ageing and Retirement in Europe (SHARE): cross-national panel & life history data with regional indicators
- Wellbeing of caregivers in regional context: complex multilevel interactions

- „thick" data
- Plug-in: participatory self-adaptive technology development
- Development process

SHARE: care & wellbeing

- 50+ in Europe, up to 29 countries (incl. IL), 160+ regions
- Wave 1 (2004), 2 (2006)/ 3 LIFE (2008), 4 (2010) … 8 (2020)
- 480,000 in-depth interviews with 140,000 individuals
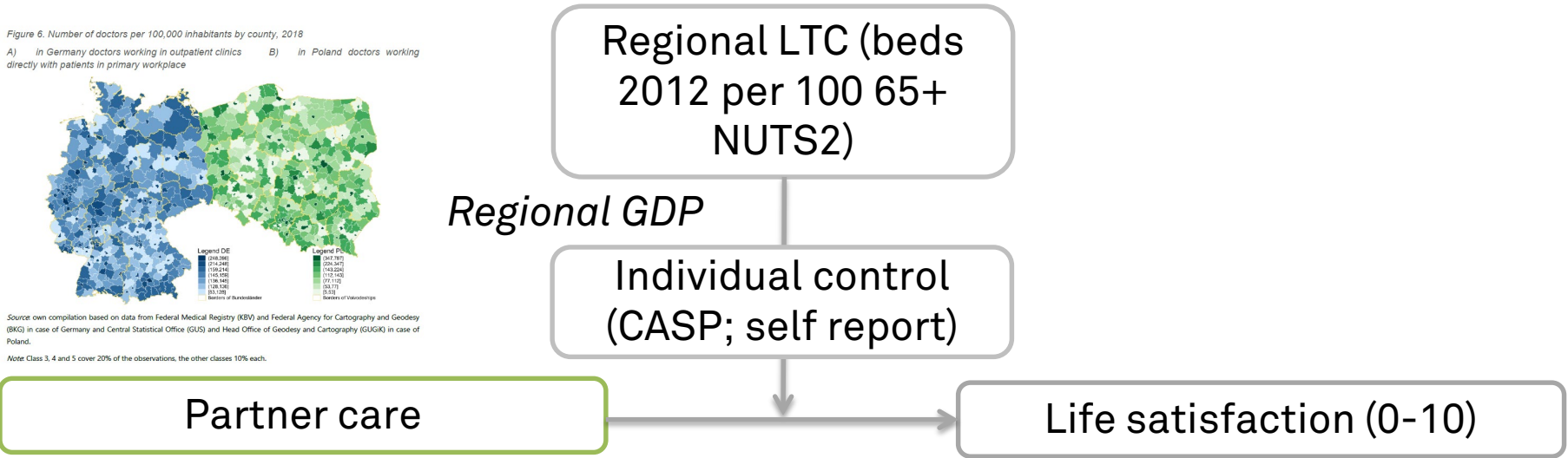- health, family & socio-economic situation, life histories

Care & wellbeing: set-up

Figure 6. Number of doctors per 100,000 inhabitants by county, 2018

A) in Germany doctors working in outpatient clinics   B) in Poland doctors working directly with patients in primary workplace

Source: own compilation based on data from Federal Medical Registry (KBV) and Federal Agency for Cartography and Geodesy (BKG) in case of Germany and Central Statistical Office (GUS) and Head Office of Geodesy and Cartography (GUGiK) in case of Poland.
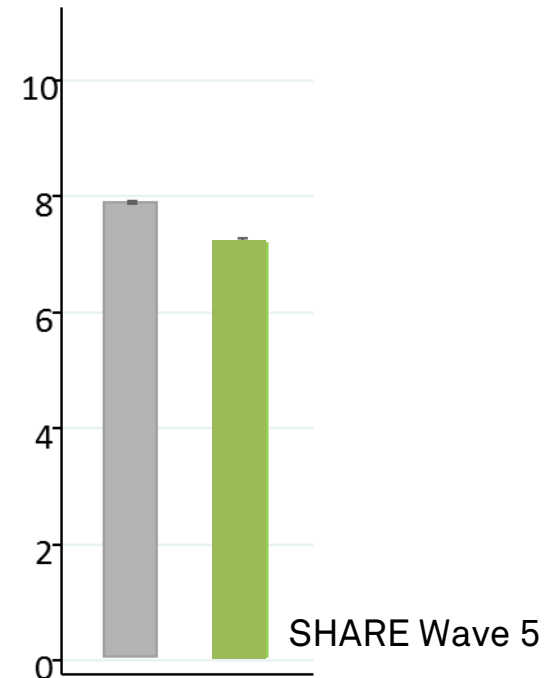
Note: Class 3, 4 and 5 cover 20% of the observations, the other classes 10% each.

Regional LTC (beds 2012 per 100 65+ NUTS2)

Regional GDP

Individual control (CASP; self report)

Partner care → Life satisfaction (0-10)

gender   children
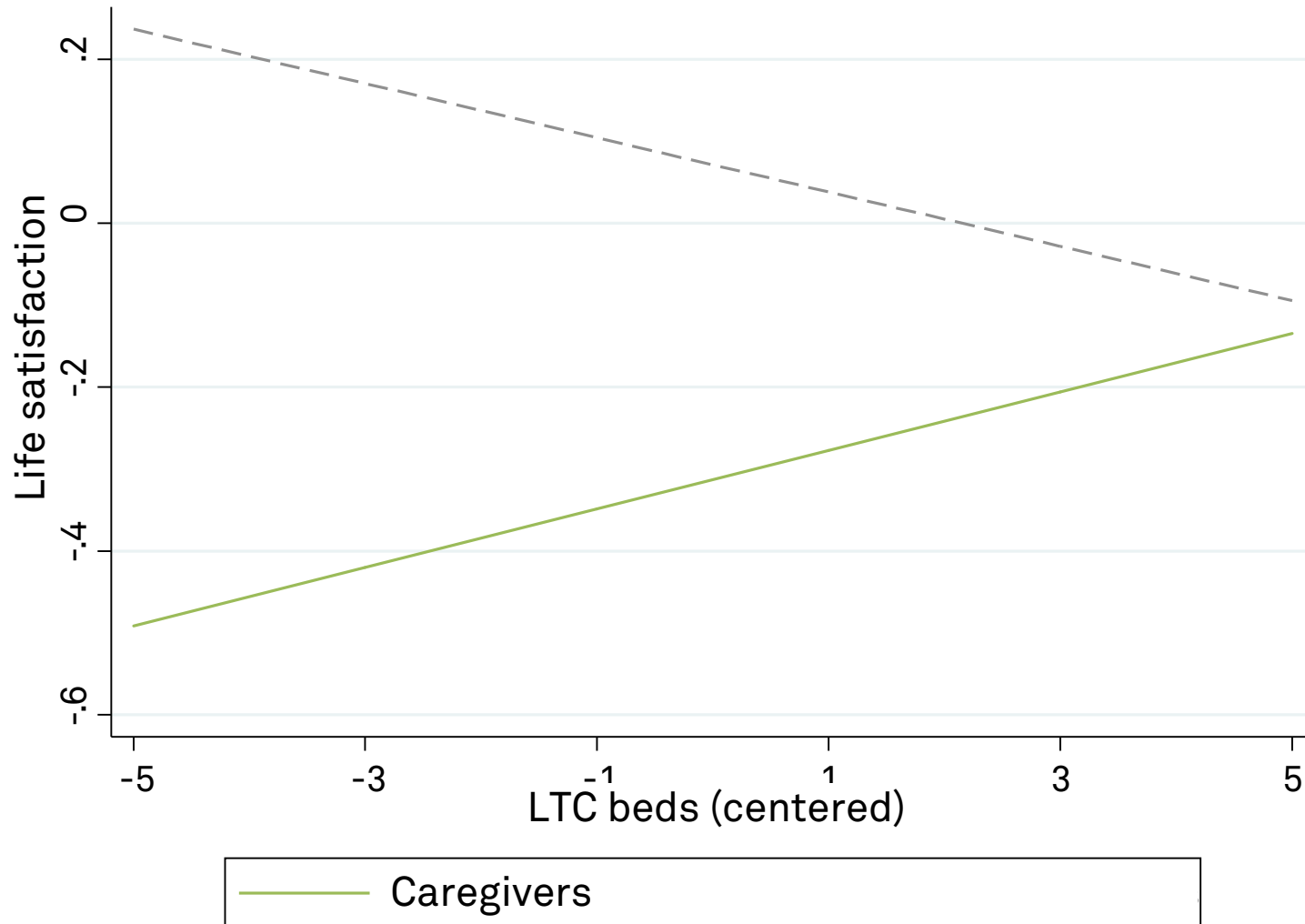age   hhsize
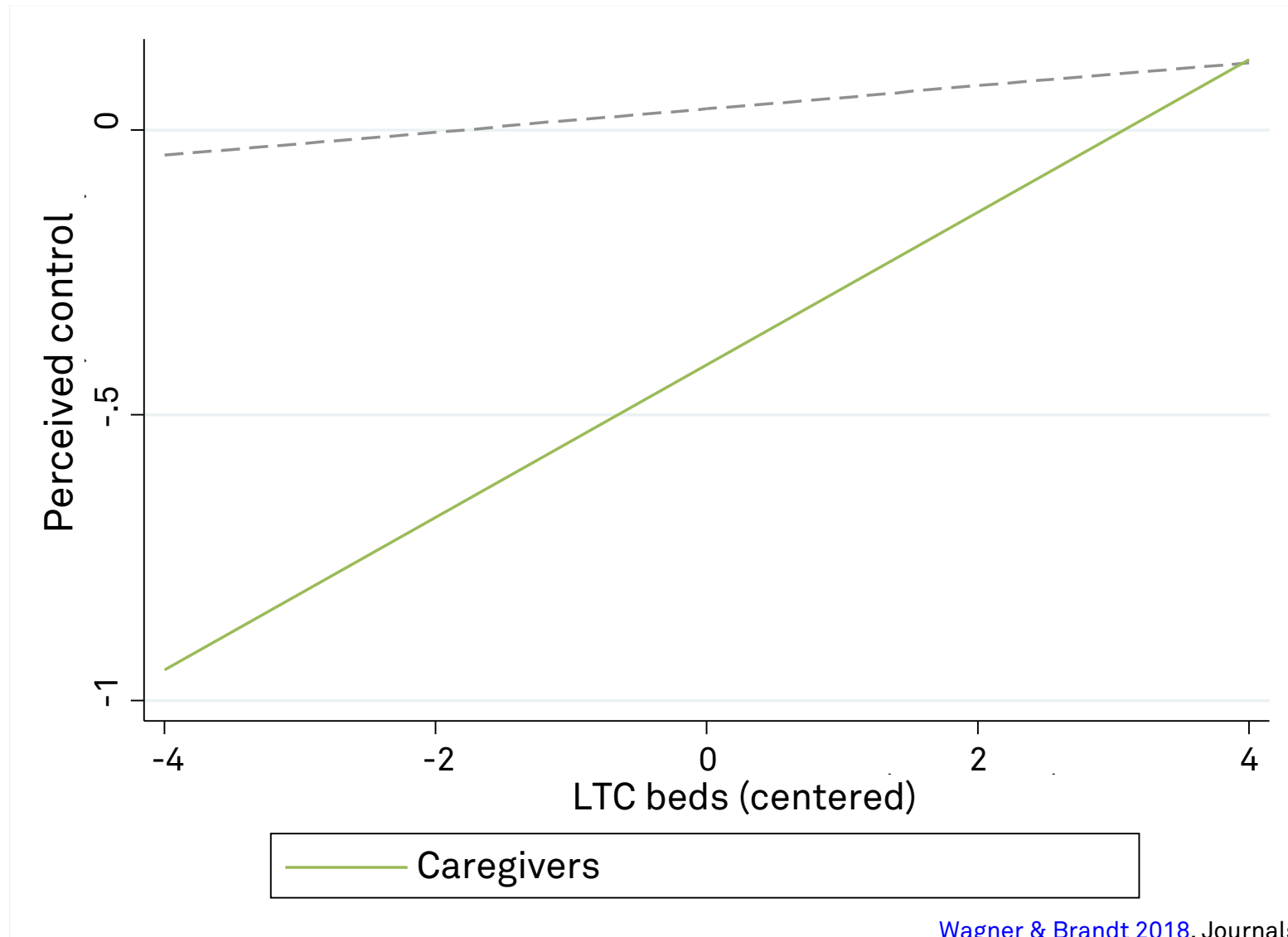education   health
hhincome   religiosity
employment   urban-rural

"On a scale from 0 to 10 where 0 means completely dissatisfied and 10 means completely satisfied, how satisfied are you with your life?"

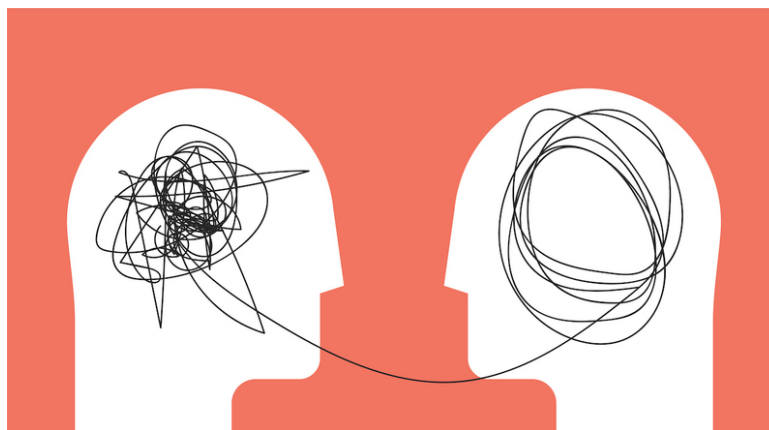SHARE Wave 5

Interaction life satisfaction*LTC

Based on a random-intercept-(ML) model with cross level interaction

Legend: Caregivers

Wagner & Brandt 2018, Journals of Gerontology Social Sciences

Interaction control*LTC

Wagner & Brandt 2018, Journals of Gerontology Social Sciences

Specific challenges

- ….this is still very very rough, and just a first indication for a correlation and a possible pathway….

- What we would need is a
    - longitudinal model (change in care infrastructure leading to change in wellbeing effects of care via different channels),
    - while accounting for the partner and care characteristics, life histories
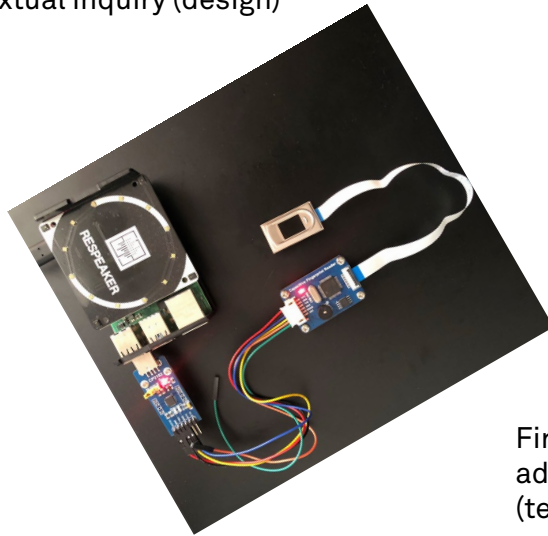   - and assessing heterogeneous effects between different groups
    - …

https://centralstationcrm.de/blog/komplexitaet-im-vertrieb-weniger-anbieten-
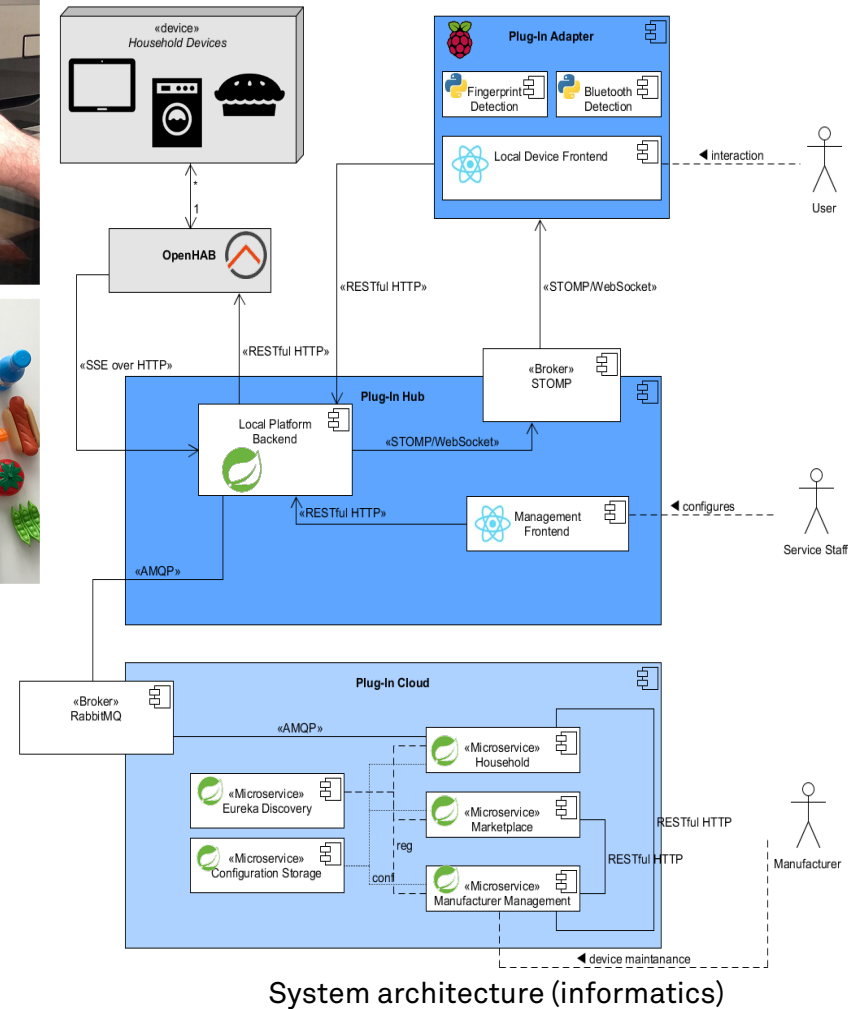-mehr-verkaufen

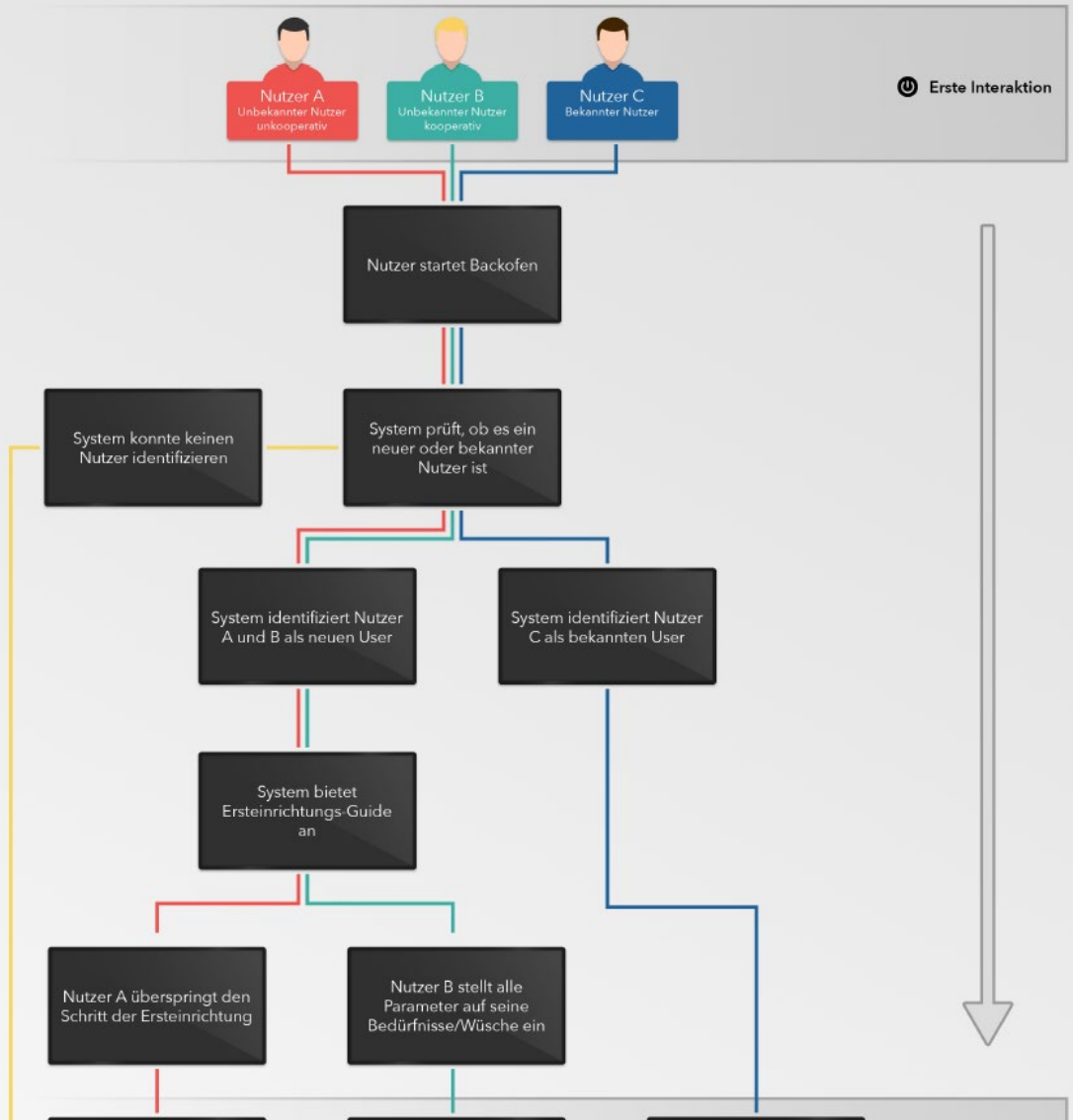… but now to something completely different:

Contextual inquiry (design)

First test adaptor (technology)

System architecture (informatics)

https://sag.sowi.tu-dortmund.de/forschung/projekte/plug-in

User adaption and participation

Flow chart (data science and engineering)

- Scenario 1: change font size

- Scenario 2: favorite functions

- Scenario 3: more than one user

- ....

- Our job at the moment:
User participation (via workshops,
expert interviews, quantitative
survey)

- How can we use the data from
the interface in future?

General challenges

- Complexity (and endogeneity)
    - high data intensity in sensoric data streams: what do we have to measure how in order to be able to learn and (make the machine) adapt efficiently? Data processing and feedback loops?
    - many life course streams within different contexts over time

- Selectivity, often central to research questions, e.g. participation, attrition/retention and response behaviour based on
    - health & cognition, socio-economic status, language skills
    - life situation (institutions, private hh sampling)
    - (panel) mortality
    - ….

- Comparability across time and space
    - Cultural differences (e.g., social desirability)
    - Methodological differences (sample selection, translation, measures)
    - Missing regional /longitudinal context information (weighting: gold standard?!)

Conclusion: a lot to do and find out

- Data & possibilities have become almost as complex as social reality
- Causalties hard to assess but knowlegde absolutely needed in order to understand and efficiently target social risks
- Sociology urgently needs statistics & data science support!

https://medium.com/on-philosophy/spiegel-im-spiegel-5b93a3079a73

THANK YOU!