



School of  
Management and Law

# Willingness to share data: Wann sind Konsumenten bereit, ihre persönlichen Daten mit einem Unternehmen zu teilen?



Dr. Kurt A. Ackermann, Dr. Thoralf Mildenerger, Dr. Martin Frey

Dr. Linda Miesler, Dr. Angela Bearth

Präsentation Statistikertreffen, 25.06.2019

# Projekthintergrund und Eckwerte

# Kontext und Projektpartner



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Innosuisse - Schweizerische Agentur  
für Innovationsförderung

Forschungspartner

Zürcher Hochschule  
für Angewandte Wissenschaften

**zh  
aw** School of  
Management and Law

Hauptumsetzungs-  
partner

**AZ Direct**

+  
Weitere  
Praxispartner

- BSI
- Post
- Sanitas

***Entwicklung eines statistischen Modells zur Vorhersage der Datenteilbereitschaft von Konsumentinnen und Konsumenten (WTS-Index)***

***für eine Anwendung in den Bereichen...***

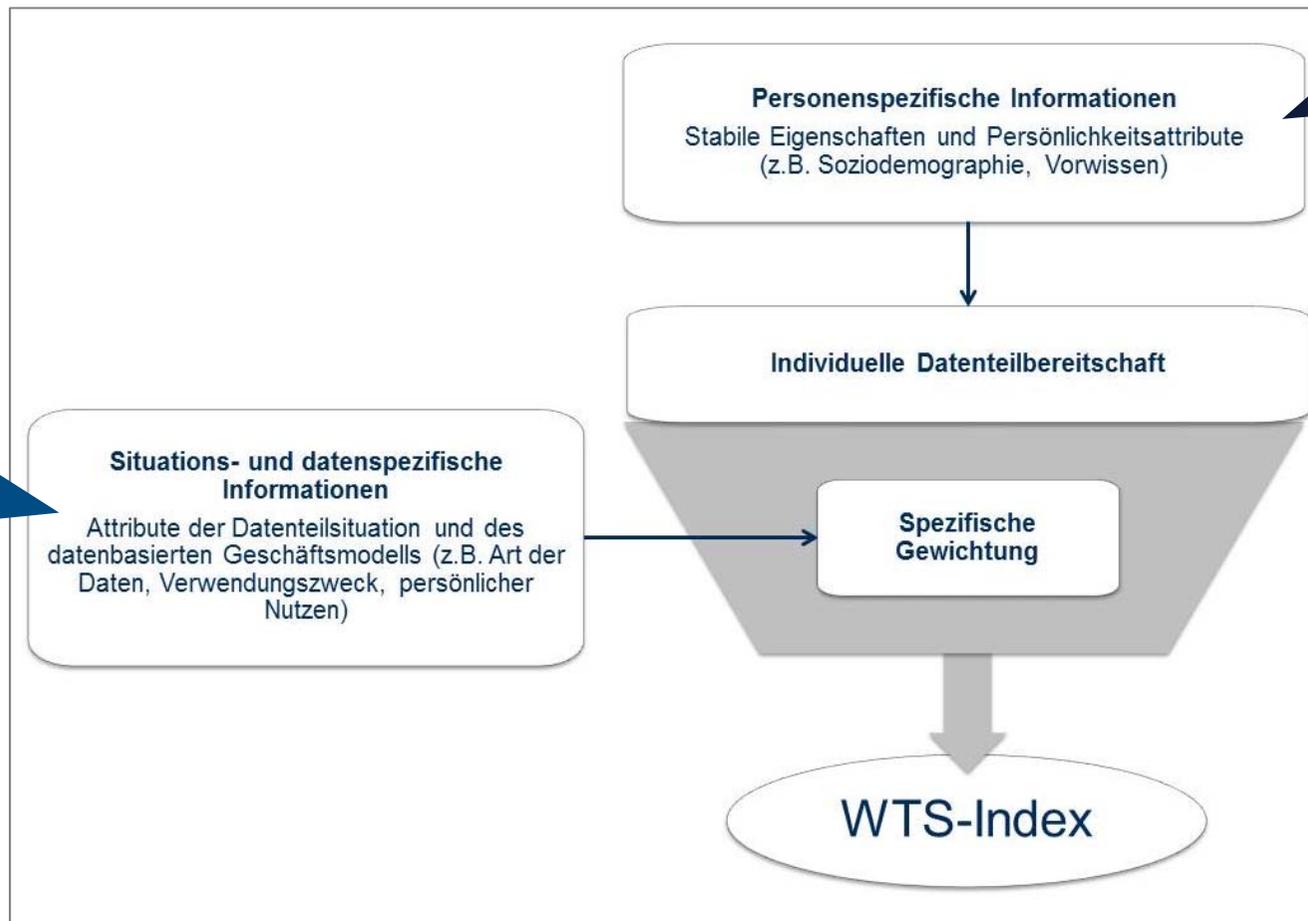
***(1) WTS Customer Scoring (Targeting)***

***(2) WTS Consulting (Beratung)***

# Konzeptioneller Hintergrund

Ziel:

Prognose der Wahrscheinlichkeit, mit welcher eine bestimmte Person ihre persönlichen Daten in einer bestimmten Situation mit einem Unternehmen teilt.



AZ Direct Data World

Zum Beispiel:

- Alter
- Geschlecht
- Ausbildung
- etc.

Zum Beispiel:

- Branche
- Art der Daten
- Gegenleistung
- etc.

Online-Experiment  
mit systematischer  
Variation von  
Szenarien

# Experimentelles Design

# Untersuchung der Datenteilbereitschaft mittels Online-Experiment

**Branche**

**Datentyp**

**Verwendungszweck**

**Gegenleistung / Nutzen**

**Anonymität**

**Beispiel**

Ein Detailhändler, bei dem Sie regelmässig einkaufen, möchte Daten zu den Routen, die Sie im Laden laufen nutzen, um Muster im Einkaufsverhalten der Kunden zu erkennen, aussagekräftige Profile der Kunden zu erstellen und Zielgruppen für bestimmte Produktsegmente (z.B. elektronische Produkte, Pflegeprodukte) zu identifizieren. Wenn Sie zustimmen, dass Ihre Informationen für diesen Zweck genutzt werden, erhalten Sie virtuelle Punkte, welche Sie in der App als Spieleinsätze nutzen können, um Rabatte für den Laden zu gewinnen. Die Auswertungen Ihrer Daten erlauben keine Rückschlüsse auf Sie als Einzelperson, d.h. Sie bleiben anonym.

# Untersuchung der Datenteilbereitschaft mittels Online-Experiment



Ihre Bank möchte Daten zu Ihrem Verhalten Online (z.B. Suchverlauf oder -begriffe, Seitenaufrufe und Mausbewegungen) nutzen, um Muster im Verhalten ihrer Kunden zu erkennen und darauf aufbauend innovative Produkte und Dienstleistungen anzubieten. Wenn Sie zustimmen, dass Ihre Informationen für diesen Zweck genutzt werden, erhalten Sie virtuelle Punkte, welche Sie als Geldbetrag für wohltätige Zwecke spenden können. Die Auswertungen Ihrer Daten erlauben Rückschlüsse auf Sie als Einzelperson, d.h. Sie sind für das Unternehmen identifizierbar.

**Würden Sie Ihre persönlichen Daten mit dem Unternehmen teilen, wenn sie für diesen Zweck genutzt werden?**

Ja

Nein



# Untersuchung der Datenteilbereitschaft mittels Online-Experiment

5 Szenario-Variablen:

- Branche (4 Stufen)
- Datentyp (10 Stufen)
- Verwendungszweck (4 Stufen)
- Gegenleistung / Nutzen (5 Stufen)
- Anonymität (2 Stufen)

Insgesamt  $4 * 10 * 4 * 5 * 2 = 1600$  Kombination.

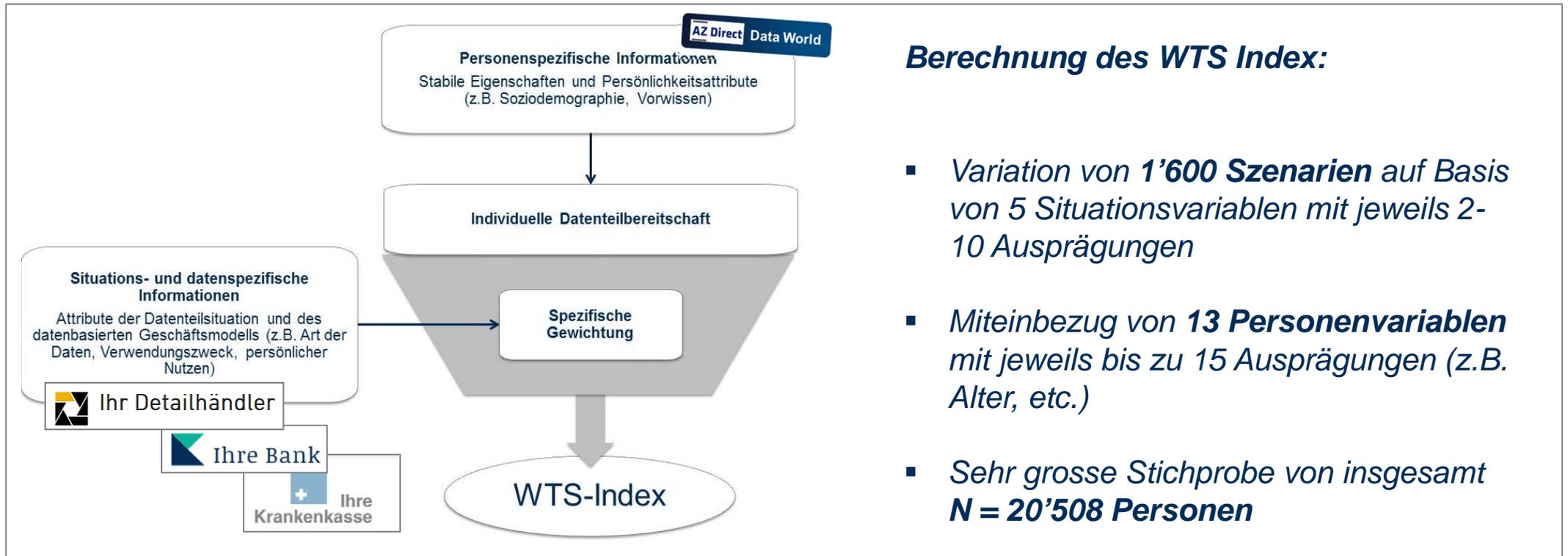
Anforderungen:

- Jede Testperson beantwortet ein Set von 10 Fragen
- Fragen pro Person sollen nicht zu bunt durchgewürfelt sein (nur 1 Verwendungszweck)
- Alle Haupteffekte, Zweifach-Interaktionen und ein ggf. vorhandener Set-Effekt sollen geschätzt werden können

**Incomplete Block Design** (R-Paket: AlgDesign, Funktion optBlock()):

- 160 Sets mit jeweils 10 Fragen zusammengestellt

# Untersuchung der Datenteilbereitschaft mittels Online-Experiment



## Berechnung des WTS Index:

- Variation von **1'600 Szenarien** auf Basis von 5 Situationsvariablen mit jeweils 2-10 Ausprägungen
- Miteinbezug von **13 Personenvariablen** mit jeweils bis zu 15 Ausprägungen (z.B. Alter, etc.)
- Sehr grosse Stichprobe von insgesamt **N = 20'508 Personen**

- **Individuelle Schätzung der Wahrscheinlichkeit**, mit welcher eine bestimmte Person ihre Daten im Kontext eines spezifischen Cases teilen würde.
- Bestimmung des **Einflusses einzelner Situationsvariablen** auf die Datenteilbereitschaft
- Bestimmung des **Einflusses einzelner Personenvariablen** auf die Datenteilbereitschaft

# Untersuchung der Datenteilbereitschaft mittels Online-Experiment

## Statistisches Modell: Logistische Regression

- Zielvariable: Datenteilbereitschaft (Wahrscheinlichkeit)
- Einflussvariablen: Szenariovariablen (Nutzen, Anonymität etc.)
- Persönliche Variablen: Alter, Demographie

## Problem:

- Vorhandene Persönlichkeitsvariablen erfassen vermutlich nicht alle individuellen Einflüsse auf die Datenteilbereitschaft
- Von jeder Testperson liegen Antworten auf 10 Fragen vor, diese sind nicht unabhängig

## Mixed-Effects-Modell:

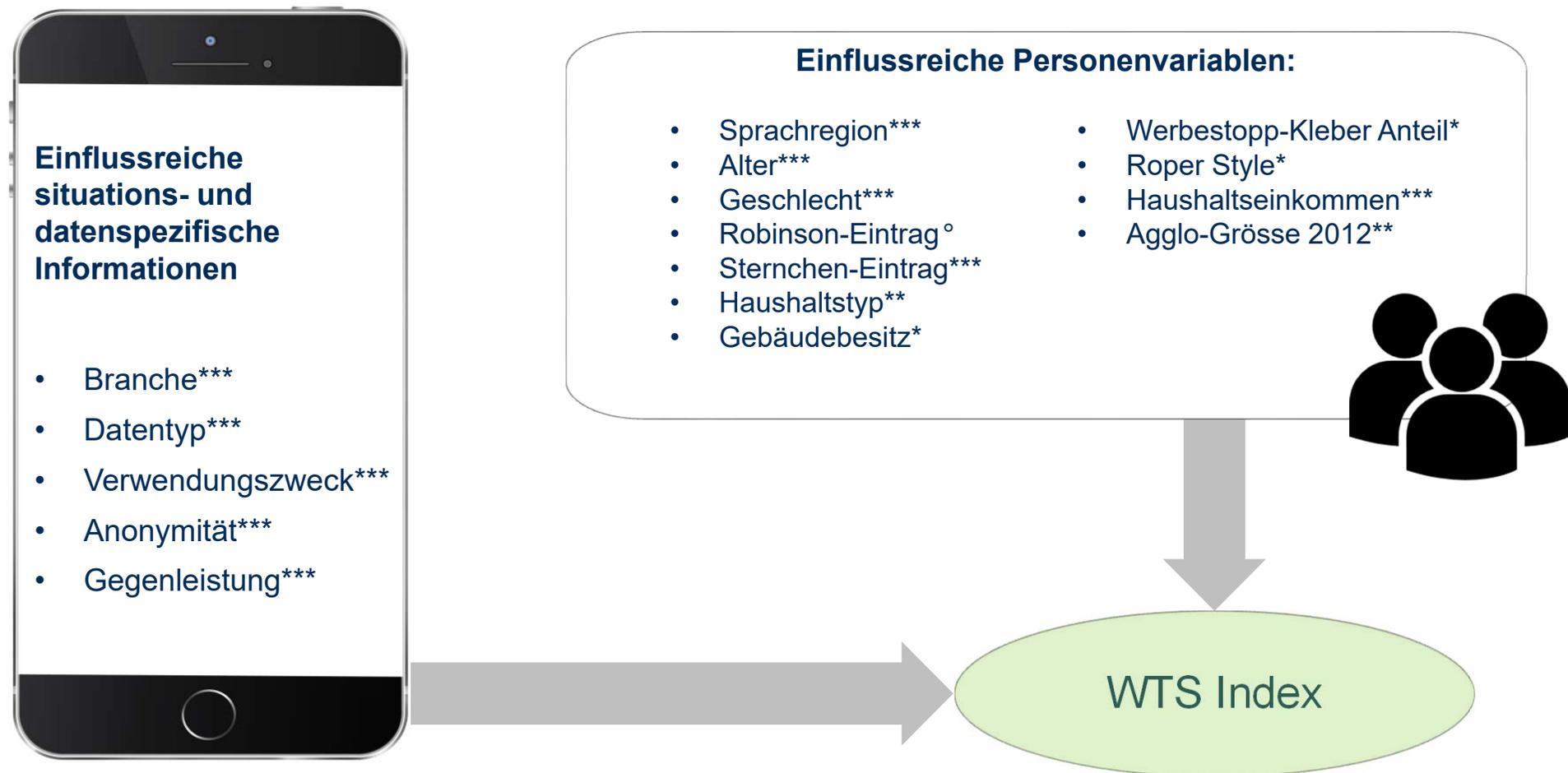
- Szenario- und Personenvariablen (fixed effects)
- Individueller Effekt pro Person (random effect): Jede Person hat noch eine individuelle Komponente, die die Datenteilbereitschaft erhöht oder senkt. Diesen Random Effect wollen wir quantifizieren und für die Vorhersage bei neuen Personen neutralisieren

Modell geschätzt mit Funktion `glmer()` in `lme4`. Rechnet sehr lange, einige Tricks nötig!

# Ergebnisse aus dem Online-Experiment

# Resultate

## Übersicht zu einflussreichen Variablen (WTS Scoring)

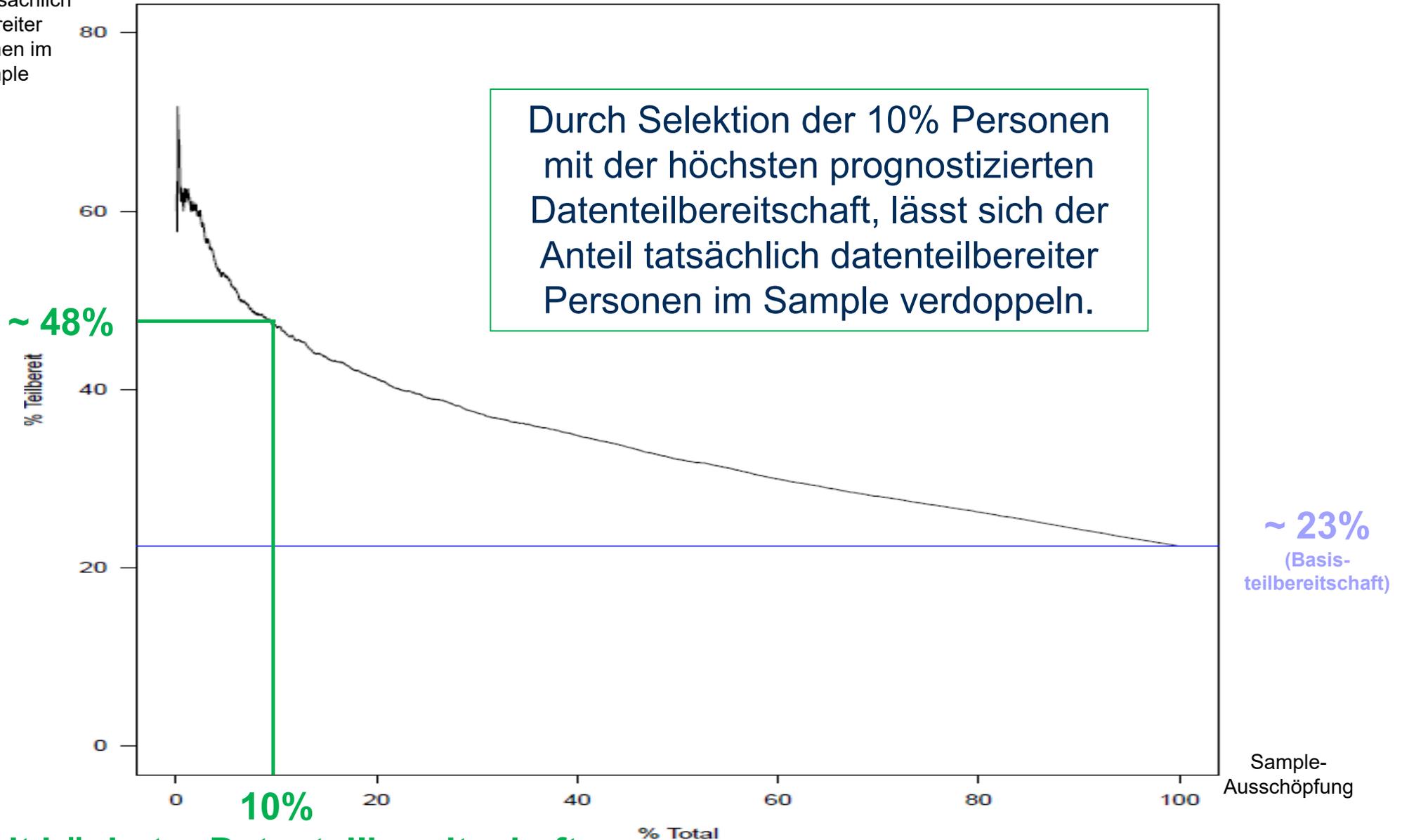


# Ausgewählte Resultate

- **Datenteilbereitschaft: ~20-40%** (*Annahme: Unterschätzung aufgrund Privacy Paradox*)
- **Branche:**
  - + *Retailer und Telekom-Unternehmen*
  - *Banken und Versicherungen*
- **Datentyp:**
  - + *Daten zum Zahlungs- oder Einkaufsverhalten, etc.*
  - *Medizinische Daten, Daten zur sozialen Kommunikation, etc.*
- **Verwendungszweck:**
  - + *Verbesserung des Angebots, besseres Kundenverständnis, etc.*
  - *Weitergabe / Verkauf der Daten, Werbezwecke*
- **Gegenleistung / Nutzen:**
  - + *Finanzielle Vergütung, Rabatte, etc.*
  - *Finanzielle Vergütung*
- **Anonymität:**
  - + *Anonymität gegeben*
  - *Rückschlüsse auf Person möglich*

# Prognosequalität des Modells (Lift chart)

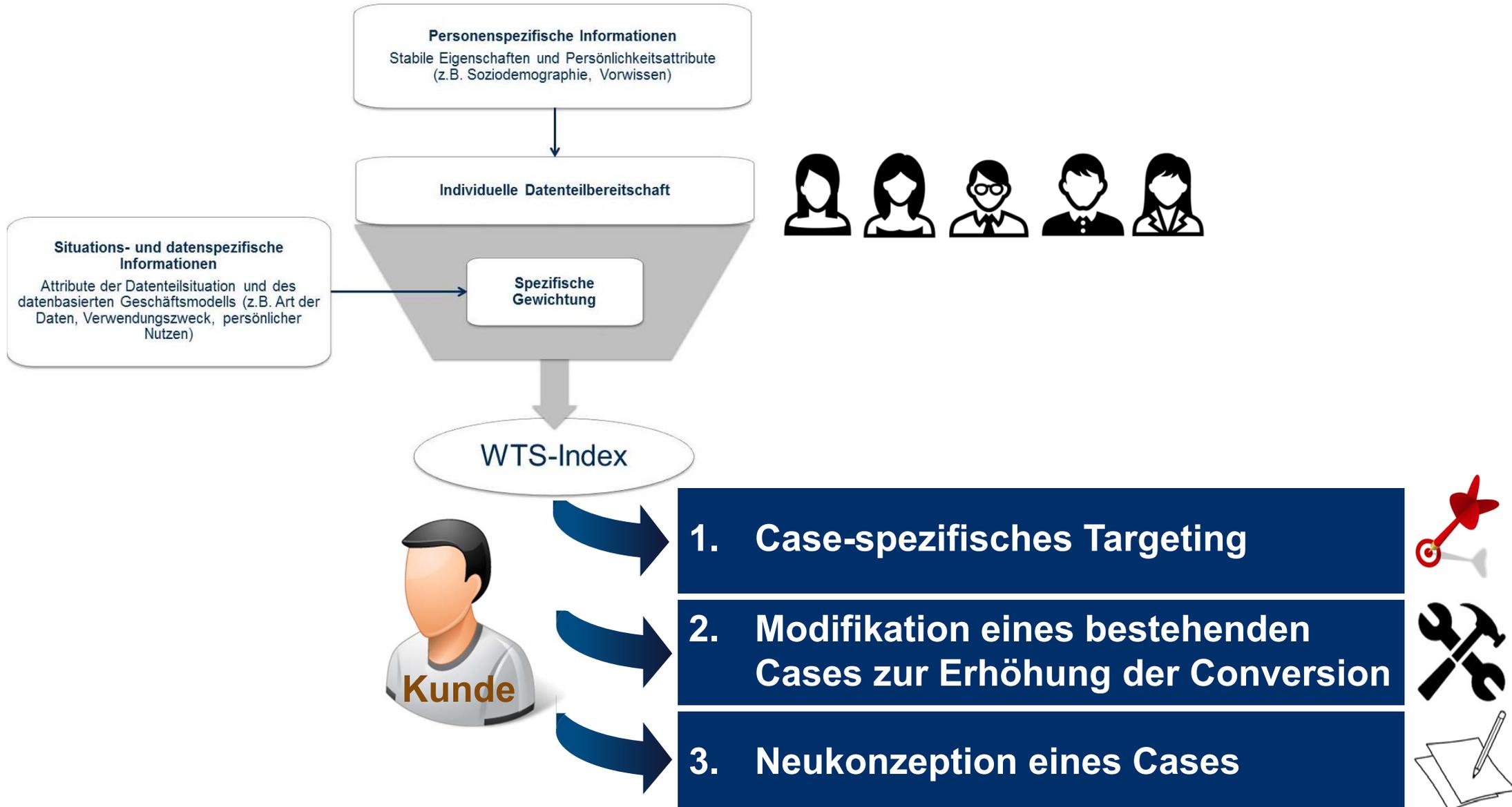
Anteil tatsächlich  
teillbereiter  
Personen im  
Sample



mit höchster Datenteilbereitschaft

# Wie kann der WTS-Index genutzt werden?

# Anwendungsfelder des WTS-Index



# Herzlichen Dank!

ZHAW Zürcher Hochschule für Angewandte Wissenschaften  
School of Management and Law  
Stadthausstrasse 14  
CH-8400 Winterthur

Web: [www.imm.zhaw.ch/sprc](http://www.imm.zhaw.ch/sprc)  
Blog: [blog.zhaw.ch/marketingmanagement](http://blog.zhaw.ch/marketingmanagement)  
Twitter: [www.twitter.com/imm\\_zhaw](http://www.twitter.com/imm_zhaw)

