

INTERNETBASIERTE TECHNOLOGIEN UND DIGITALE GESCHÄFTSMODELLE

Analysen zum Innovationsverhalten kleiner und
mittelständischer Unternehmen in Deutschland

München, 02. November 2016

PROF. DIETMAR HARHOFF

PROF. WOLFGANG KELLERER

PROF. MONIKA SCHNITZER

ZHAOXIN PU

MATTHIAS WILHELM

MICHAEL CHROMIK

Eine zweite Datenerhebung wurde im Oktober auf Basis der neuen technische Architektur erfolgreich durchgeführt

Aktueller Stand

- Die **technische Architektur** wurde erfolgreich überarbeitet und deren Implementierung erfolgreich validiert (*interne Validierung*)
- Eine **neue Datenerhebung** der 30.000 URLs wurde im Oktober durchgeführt

Projektleitung

- Prof. Dietmar Harhoff, Ph.D. (MPI)
- Prof. Dr.-Ing. Wolfgang Kellerer (TUM)
- Prof. Dr. Monika Schnitzer (LMU)

Projektteam

- Zhaoxin Pu (MPI)
- Matthias Wilhelm (LMU)
- Michael Chromik (CDTM)

Nächste Schritte

- **Wiederholte Erhebung** und statistische Auswertung im Vergleich zu der aktuellen Erhebung
- **Anreicherung der Datenbasis** (z.B. um Lokationsdaten und Stellenanzeigen)
- **Externe Validierung**

Die überarbeitete technische Architektur gewährleistet eine wiederholbare und nachvollziehbare Datenerhebung sowie -Nutzung

Charakteristika der technischen Architektur

– Erweiterbarkeit durch modularen Aufbau:

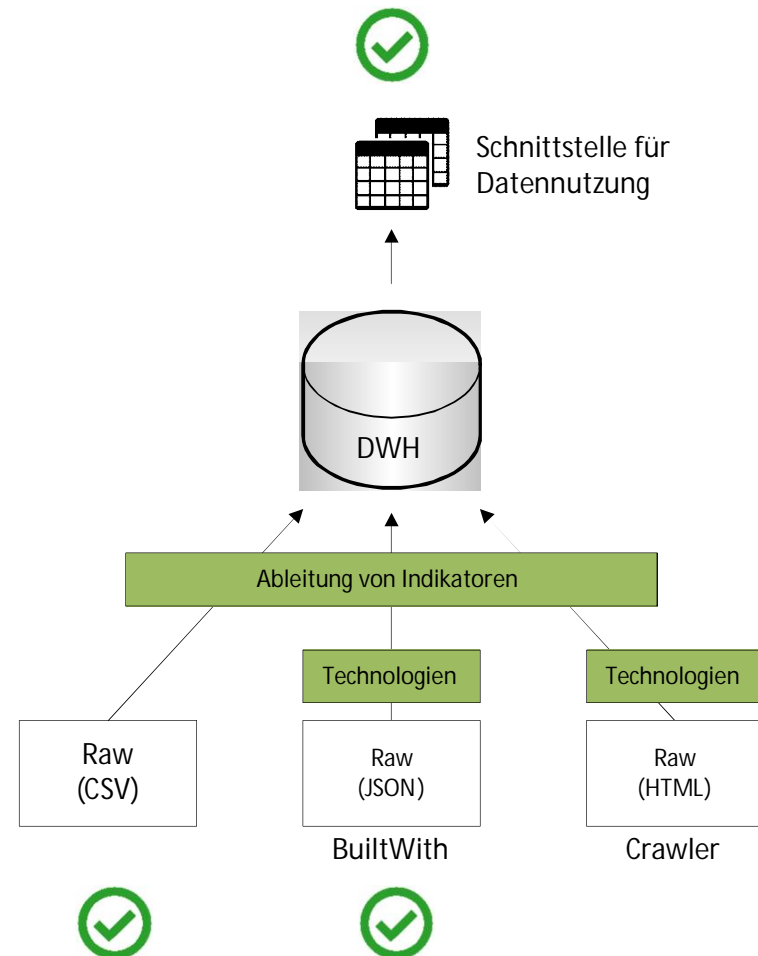
- Jede Datenquelle stellt ein eigenes unabhängiges Modul dar.

– Aufbau eines zentralen Data Warehouses:

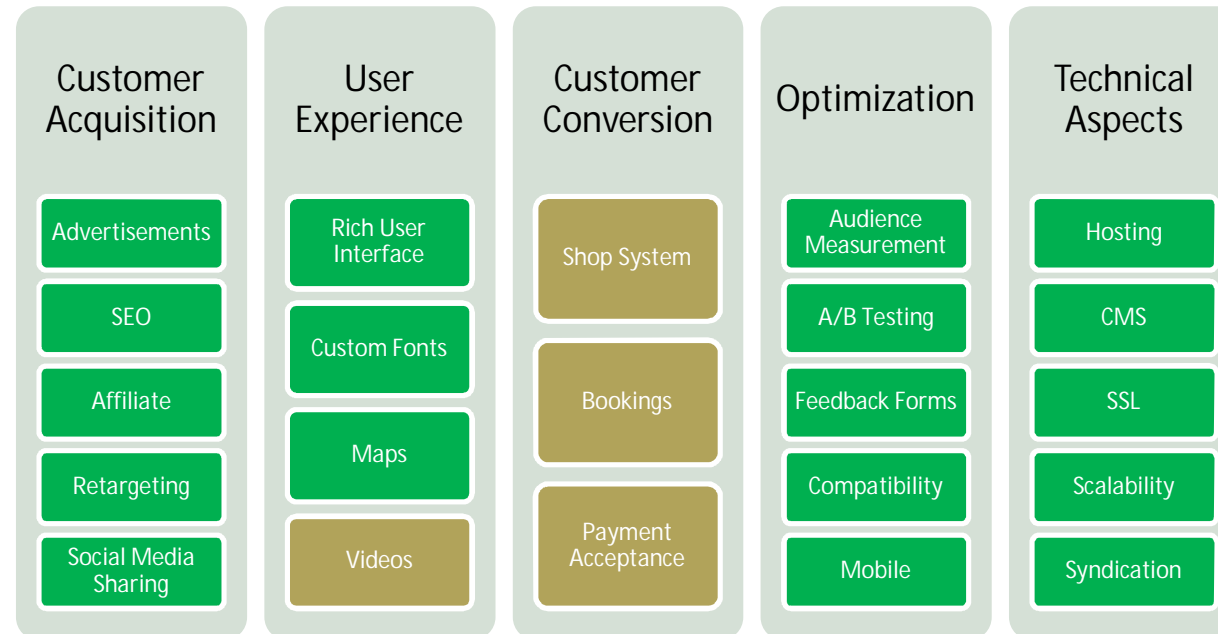
- Alle in den Modulen erhobenen (Roh)Daten und Technologien werden strukturiert in einem zentralen Data Warehouse dokumentiert.
- Dadurch soll eine verbesserte Nachvollziehbarkeit der abgeleiteten Indikatoren erreicht werden (Drilldowns).

– Data Warehouse als „Self-Service“-Schnittstelle für Datennutzung während inhaltlicher Auswertung:

- Für wiederkehrende Fragestellungen werden vordefinierte Sichten (Views) bereitgestellt
- Ad-hoc Abfragen über SQL
- Davor Dokumentation und Einweisung in Datenstrukturen notwendig



Für die zweite Datenerhebung wurden zusätzliche Indikatoren entlang des (digitalen) Kunden-Lebenszyklus erhoben



Limitationen der Datenerhebung

- Die Datenerhebung basiert derzeit lediglich auf den Startseiten der Webauftritte
 - Einige Indikatoren sind **seitenunabhängig** und lassen sich zuverlässig durch erhobene Technologien ermitteln (**grün** markiert)
 - Andere Indikatoren sind **seitenspezifisch** und erfordern ein tiefergehendes Crawling von Unterseiten (**gelb** markiert)

Beispiel Beherbergung (1/2): arcona.de nutzt Technologien aus 12 Indikatoren



Plattformübergreifender Internetauftritt, der potentiell auf die Bedürfnisse der Kunden optimiert als auch online vermarktet wird.

Indicator Group	Indicator	Example
Customer Acquisition	Advertisement	DoubleClick.Net
	Social Media Sharing	AddThis
Optimization	A/B Testing	Visual Website Optimizer
	Audience Measurement	Google Analytics
	Compatibility Optimization	Modernizr
	Mobile Optimization	Min Width
Technical Aspects	CMS	WordPress 3.5
	Scalability	Google Hosted jQuery
	SSL	Thawte SSL
User Experience	Custom Fonts	Font Awesome
	Map Technology	Google Maps API
	Rich User Interface	jQuery UI

Ein zweistufiges Validierungskonzept soll eine hohe Datenqualität der erhobenen Daten sicherstellen und deren Grenzen ermitteln

✔ Interne Validierung: „Ist unsere Implementierung korrekt?“

- Prüfung, ob die abgeleiteten Technologien und Indikatoren **konsistent mit Rohdaten** (z.B. BuiltWith) sind.
- Vorgehen:
 - Abgleich durch **Test-Skripte** und **manuelle Stichproben** durch Entwickler
 - Behebung von Bugs und Definitionslücken



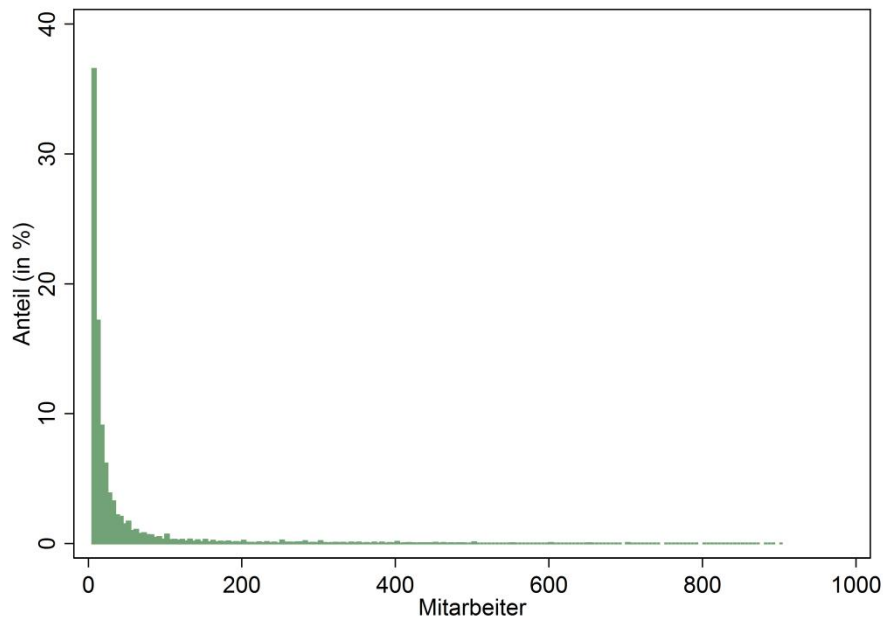
Externe Validierung: „Entsprechen die Datenquellen der Realität?“

- Prüfung, ob die Daten der angebundenen Datenquellen ein **wahrheitsgemäßes Abbild der Realität** darstellen.
- (Geplantes) Vorgehen:
 - **Manuelle Prüfung** der Internetauftritte von ca. 100 KMUs pro Indikator
 - **Sichtbare Technologien** werden vom Prüfer über **Prüfungstool** verifiziert (z.B. „Hat diese Webseite ein Kontaktformular?“)
 - **Nicht sichtbare Technologien** werden vom Prüfer über **direkte Kontaktaufnahme** mit dem KMU erfragt (Survey und ggfs. telefonischer Follow-up)

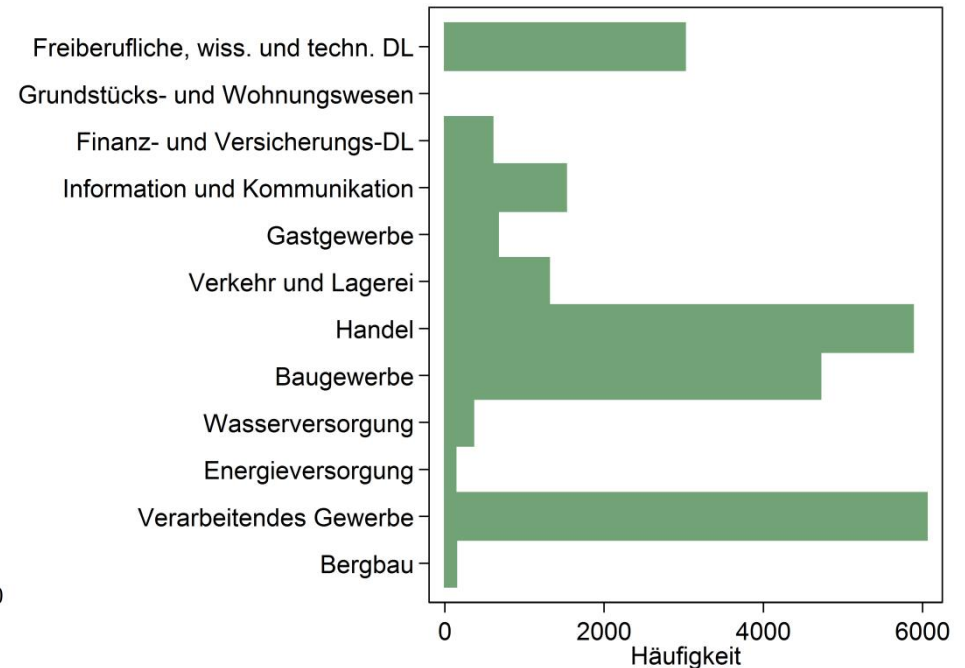
AUSWERTUNG DER WEBSEITEN AUS DER AKTUELLEN ERHEBUNG

Firmen mit <100 Mitarbeiter machen 90% unserer Stichprobe aus, über 50% der Stichprobe sind im Handel, Bau- oder verarbeitenden Gewerbe tätig

Verteilung der Firmen nach Größe



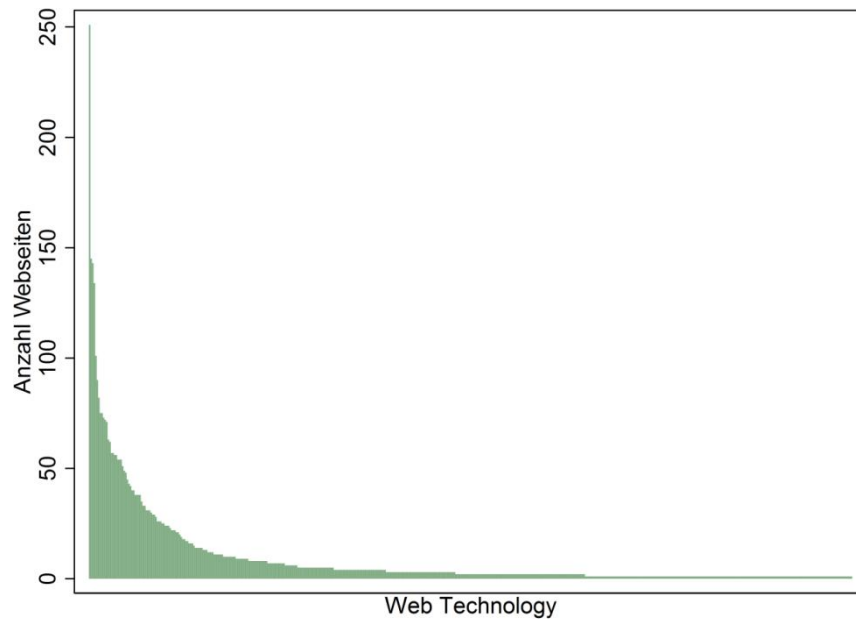
...und Wirtschaftsabschnitten (WZ1)



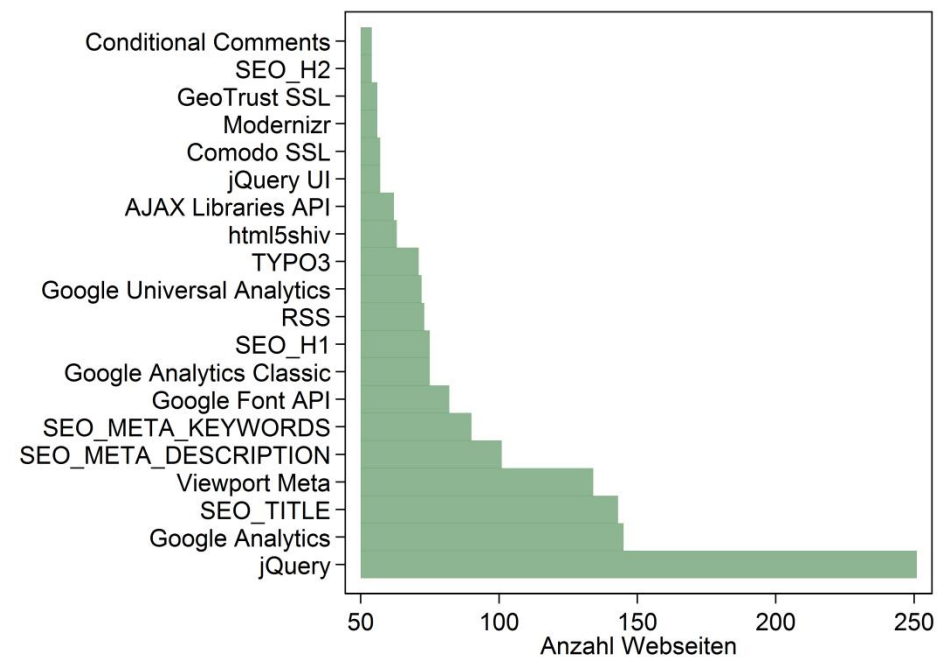
- Mehr als 90% der Firmen in unserer Stichprobe haben <100 Mitarbeiter – die **rechtsschiefe Verteilung** entspricht der Verteilung in der Population der deutschen Firmen.
- **Handel, Baugewerbe und verarbeitendes Gewerbe** sind die Wirtschaftsabschnitte mit den meisten Unternehmen in unserer Stichprobe.

Wir erfassen mit den 30k Webseiten etwa 1500 unterschiedliche indikator-relevante Technologien, die wir 23 Indikatoren zuordnen

Verteilung der erfassten Webtechnologien

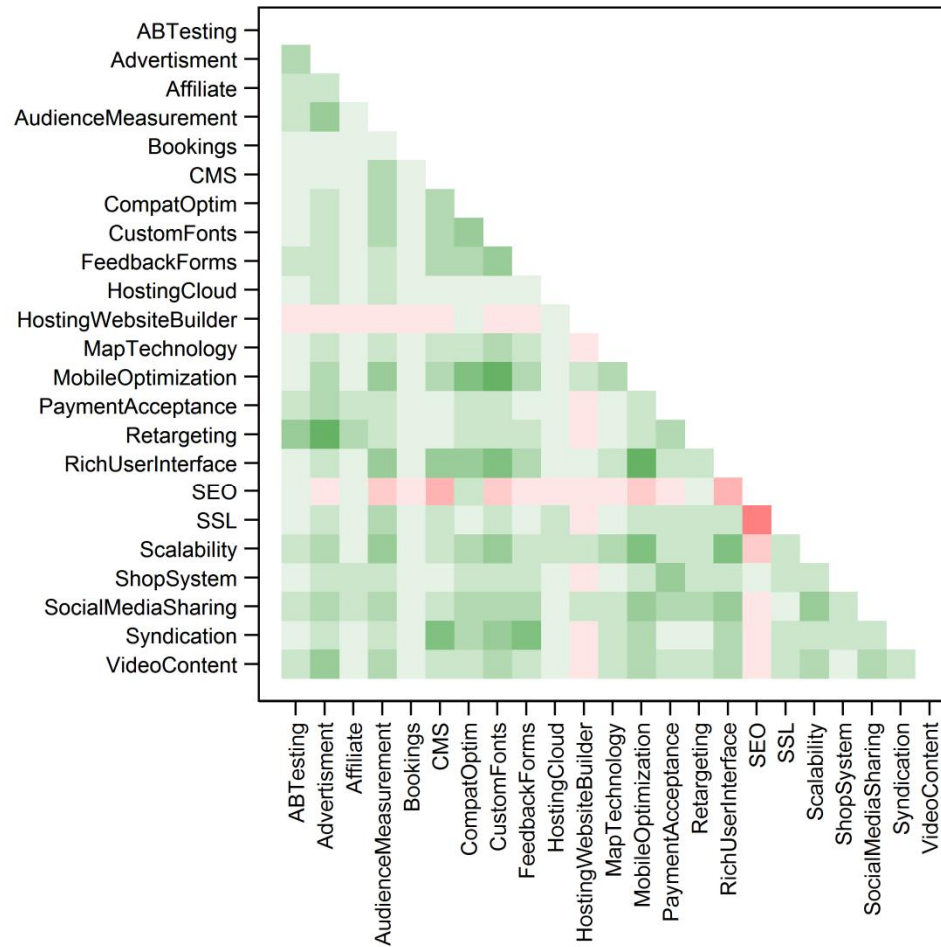


...und die Top 20



- Die Firmen verwenden eine enorme **Bandbreite an Webtechnologien**; diese lassen sich jedoch in verschiedene Kategorien einteilen.
- Nur die Top 5 Webtechnologien werden von mehr als 100 Firmen verwendet.

Unsere Analyse zeigt einzelne positive und negative Korrelationen zwischen gewissen Indikatoren auf

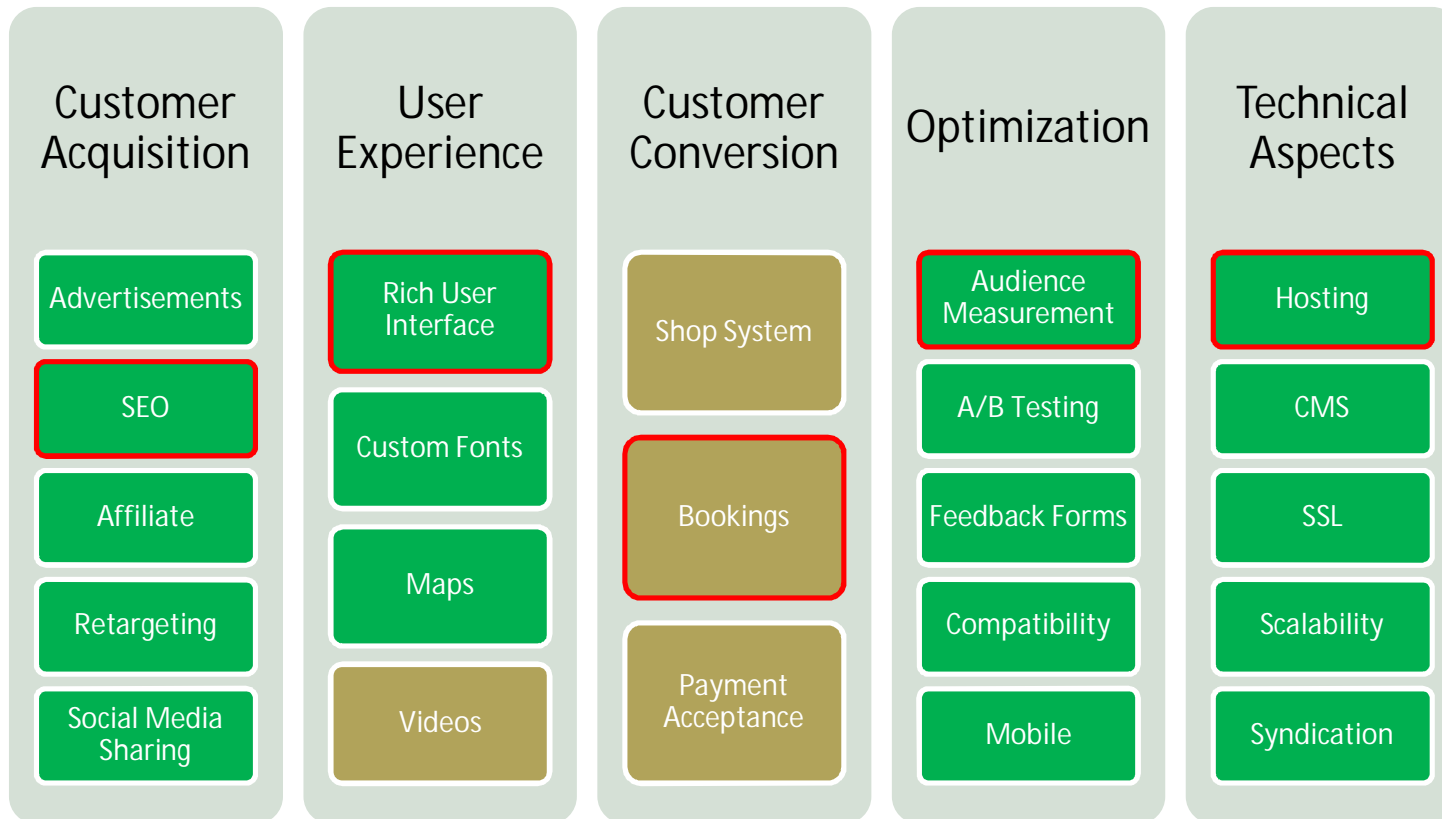


Korrelationstabelle zwischen den 23 Indikatoren

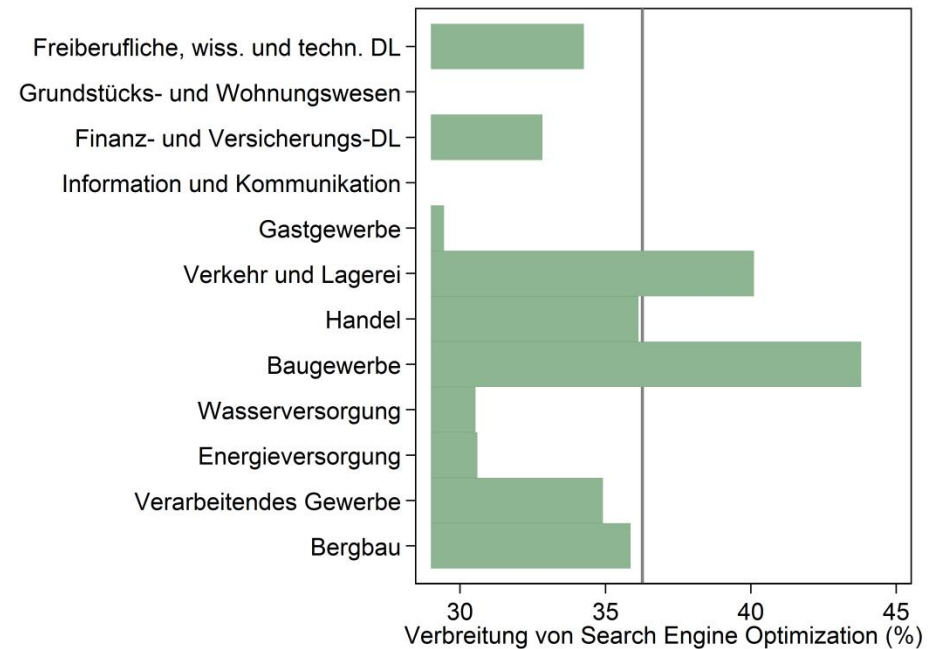
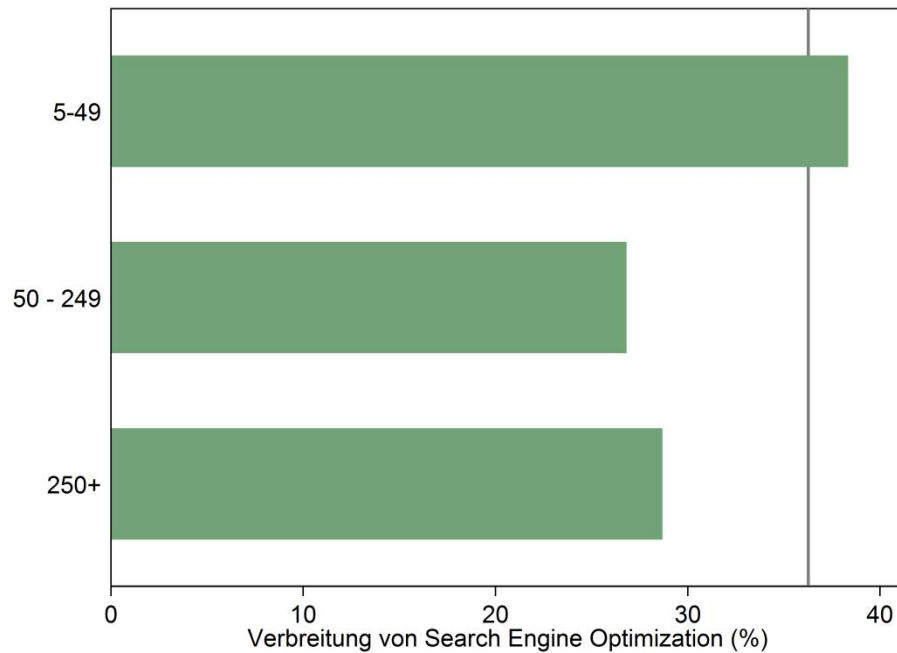


- Die Indikatoren „Search Engine Optimization“ und „Hosting Website Builder“ korrelieren mit fast allen anderen Indikatoren negativ
- Positive Korrelation zwischen ‚User Experience‘ Indikatoren und ‚Optimization‘ Indikatoren

Im Folgenden zeigen wir Details für ausgewählte Indikatoren entlang des (digitalen) Kunden-Lebenszyklus

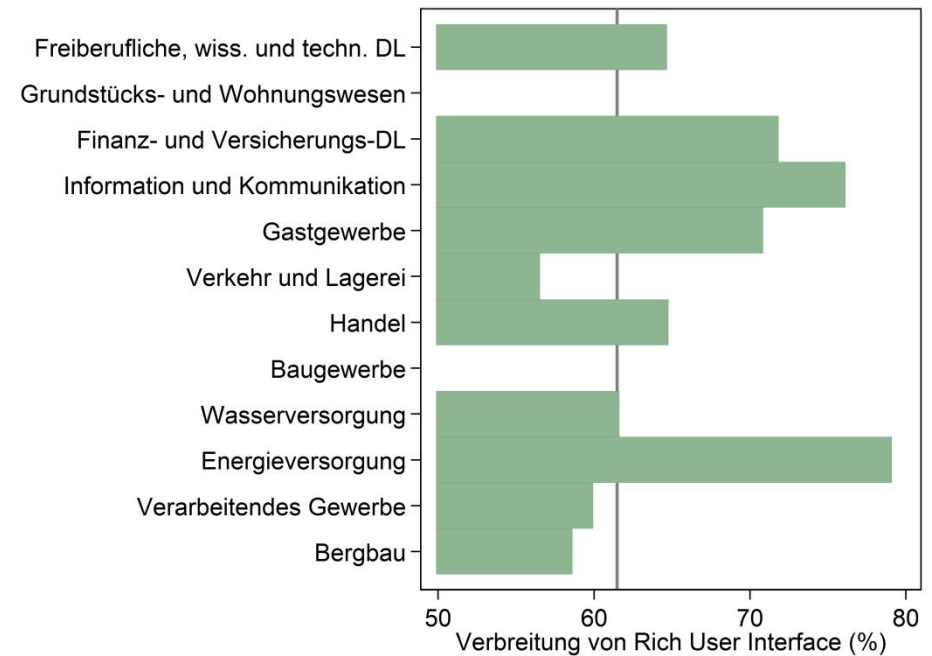
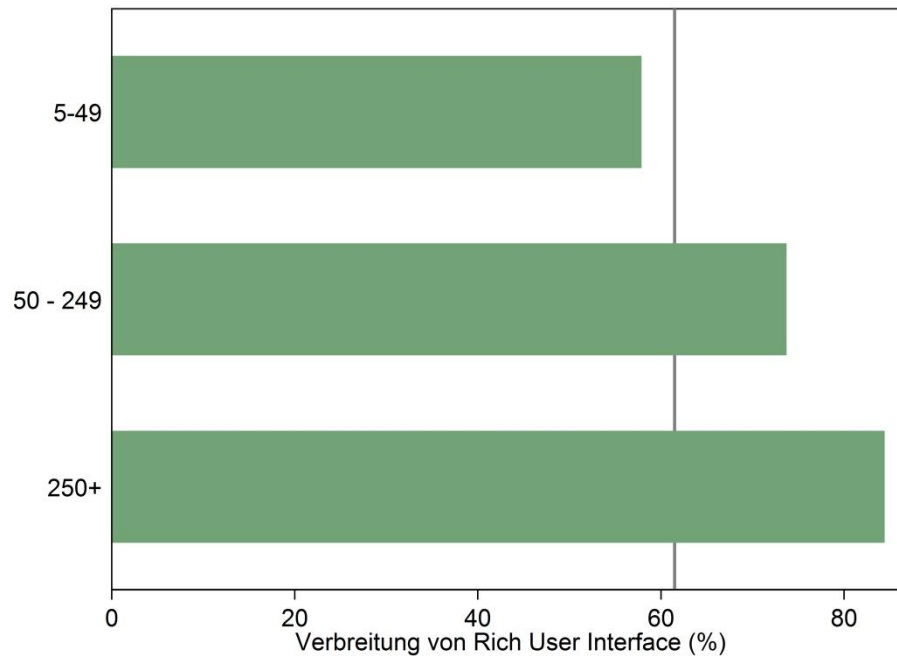


Beispiel im Bereich „Customer Acquisition“ Search Engine Optimization



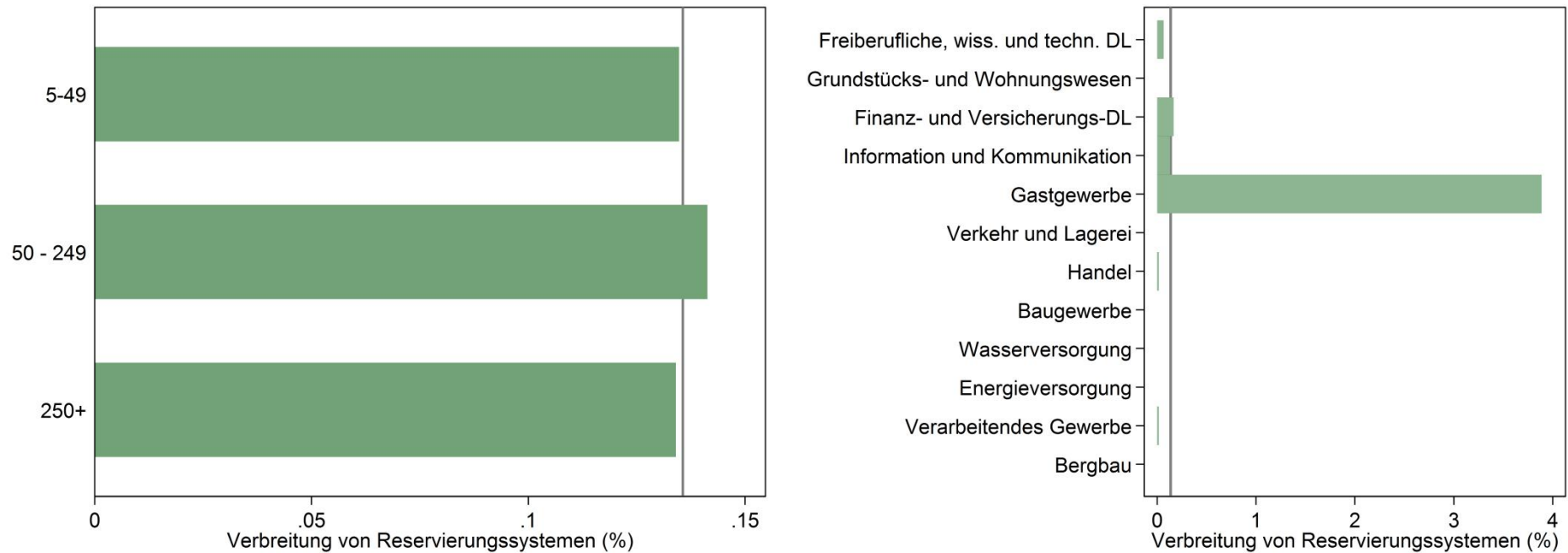
- SEO wird von knapp über einem Drittel der Firmen in unserer Stichprobe genutzt.
 - Die Verbreitung von SEO ist dabei bei kleinen Firmen größer als bei mittelgroßen Firmen.
 - Insbesondere wird SEO von Firmen im Baugewerbe und in Verkehr und Lagerei eingesetzt.
-

Beispiel im Bereich „User Experience“ *Rich User Interface*



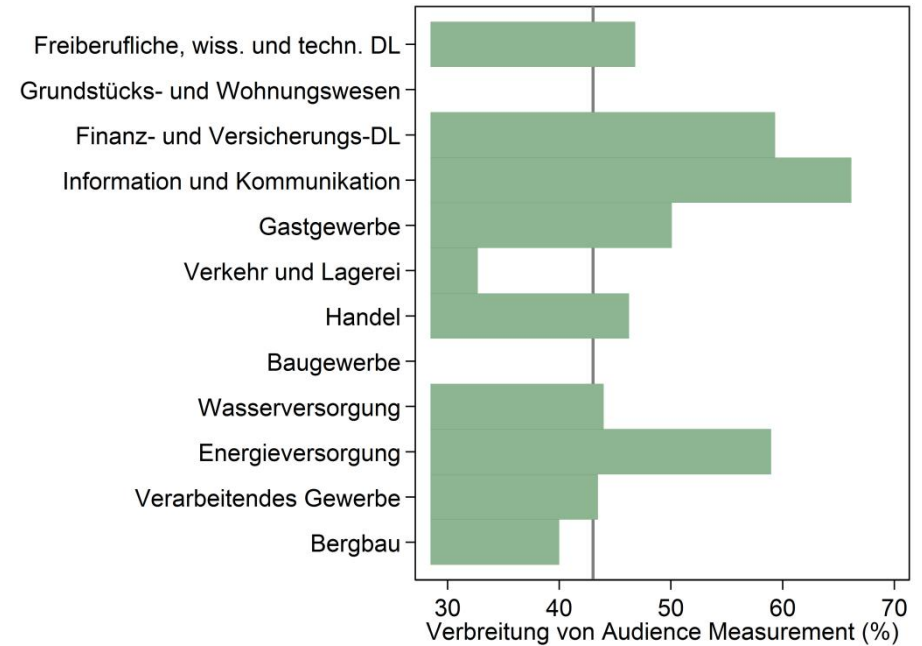
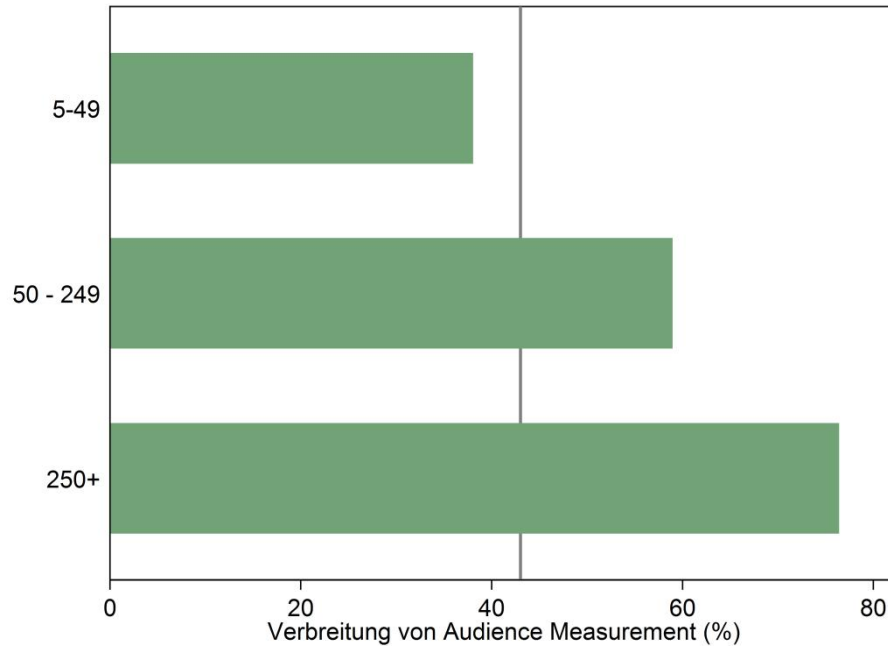
- Rich User Interface wird von mehr als 60% der Firmen in unserer Stichprobe verwendet.
 - Unter den Firmen mit >250 Mitarbeitern liegt die Verbreitung sogar bei mehr als 80%.
 - Die Nutzung ist jedoch sehr heterogen zwischen den Branchen – insbesondere im Vergleich zu SEO.
-

Beispiel im Bereich „Customer Conversion“ *Bookings*



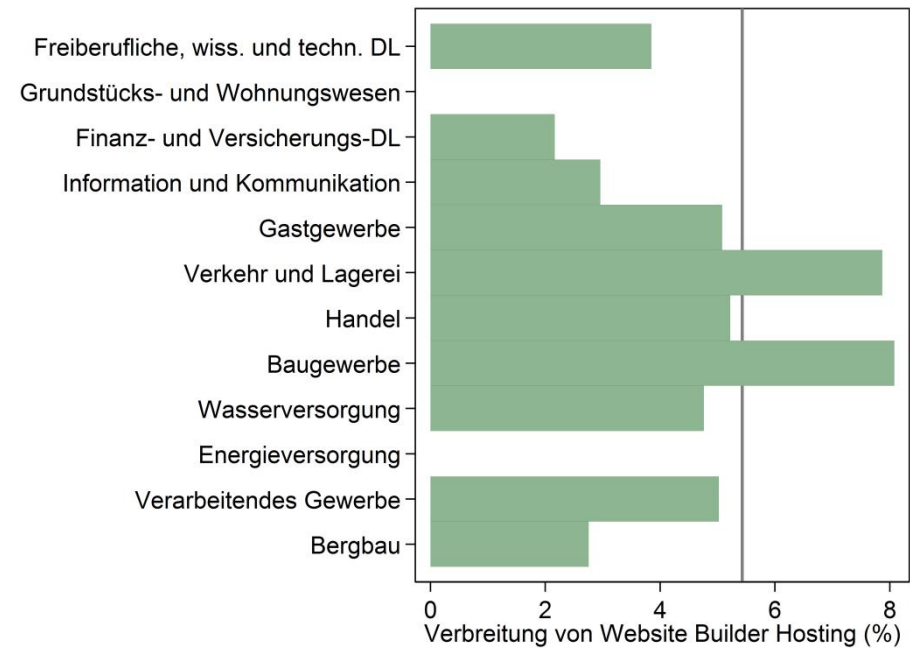
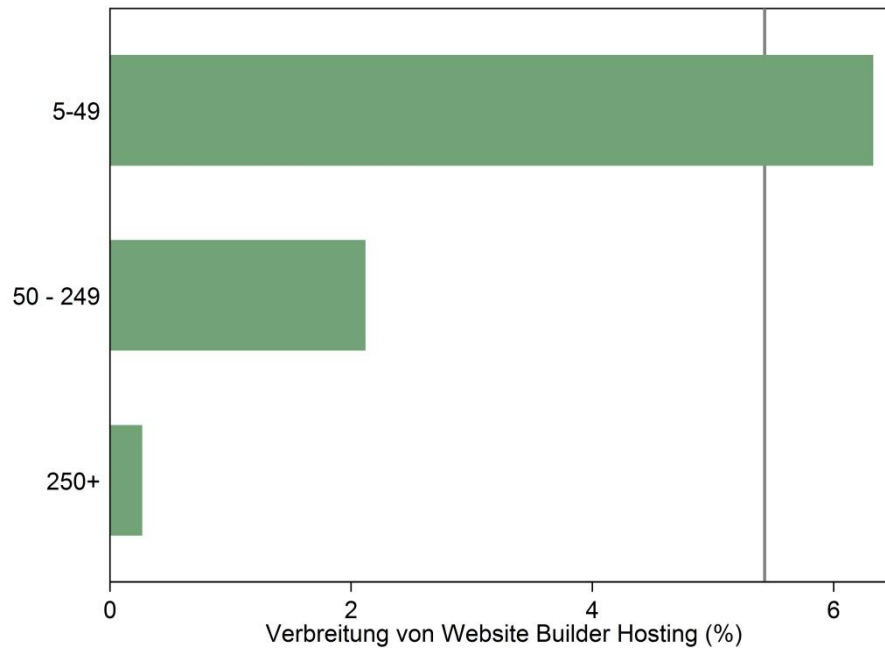
- Buchungssysteme verwenden weniger als 1% aller Firmen in unserer Stichprobe und ist eine sehr Branchen-spezifische Technologieklasse.
 - Wie zu erwarten sind es vornehmlich Firmen im Gastgewerbe, die diese Webtechnologie nutzen, allerdings auch nur zu 4%.
-

Beispiel im Bereich „Optimization“ *Audience Measurement*



- Audience Measurement ist eine beliebte Technologieklasse, insbesondere bei Firmen, die in direkten Kontakt mit dem Endkunden treten (Dienstleistung, Gastgewerbe, Handel)

Beispiel im Bereich „Technical Aspects“ *Hosting Website Builder*



- Die Verbreitung von Website Builder Hosting ist sehr heterogen nach Beschäftigten-Größenklassen.
 - Im Baugewerbe und in Verkehr und Lagerei sind Website Builder Hosting Lösungen am meisten verbreitet, wobei diese Wirtschaftsabschnitte sehr schwach bei den anderen Indikatoren abschneiden.
-

ERWEITERUNGEN UND NÄCHSTE SCHRITTE

Schritte bis zum nächsten Sprint-Review

Zweite Vollerhebung durchführen

- Zweite Erhebung durchführen um Panelstruktur aufzubauen
- Vergleich mit Ergebnissen aus der aktuellen Erhebung
- Externe Validierung

Anreicherung der Datenbasis

- Erhebung des Sitz des Unternehmens, um geographische Muster untersuchen zu können



Nächste Analyse-Schritte

- Dekomposition der Heterogenität: Wie stark unterscheiden sich Firmen innerhalb von Branchen und Größenklassen?
- Wie ändert sich die Verbreitung von Webtechnologien über die Zeit?