

Deliverable 2.1.:

Personas und Abfrageszenarien

Erhebungsmethodik

Wien, Jänner 2020

Ein FEMtech-Forschungsprojekt gefördert vom BMVIT

 **Bundesministerium**
Verkehr, Innovation
und Technologie

Impressum

Projektleitung:

Mag.^a Louise Horvath

Österreichisches Institut für angewandte Telekommunikation

Ungargasse 64-66/3/404

1030 Wien

Konsortium:

Andrew Lindley, Austrian Institute of Technology

Michael Kreil, Ciuvo GmbH

Andrea Grisold, Wirtschaftsuniversität Wien

Inhalt

Impressum	2
Einleitung	4
I. Die Nutzung von Echt-User Daten	5
1.1. Einleitung	5
1.1.1. Andere Forschungsvorhaben mit Echt-User-Ansätzen	5
1.1.2. Bedeutung von Echt-User Daten	6
1.1.3. Ciuvo als begleitender Preisvergleich	6
1.2. Erhebung	10
1.2.1. Der preis.wert-Crawler	10
1.2.2. Vorgehensweise	11
1.2.3. Ausblick auf mögliche Ausbaustufen der Echt-User Datensammlung	12
2. Personas und Abfrageszenarien	13
2.1. Methodische Vorüberlegungen zu Personas	13
2.2. Modellierung von Personas	15
2.3. Methodische Vorüberlegungen zu Abfrageszenarien	16
2.4. Modellierung von Abfrageszenarien	17
2.5. Ausblick: Stufenweise Vorgehensweise	18
3. Quellenverzeichnis	20
3.1. Literatur	20

Einleitung

Die Messung von **personalisierter Preisgestaltung** ist ein schwieriges Unterfangen. Das Ziel des vorliegenden Berichts ist es einen Überblick über die Vorgangsweise im Projekt zu geben. Gewählt wurde ein Zugang, der das Ergebnis der algorithmischen Berechnung von Preisen untersucht („Output-Approach“). Im Gegensatz dazu stünde die Analyse der eingesetzten Algorithmen selbst („Input-Approach“), was jedoch aufgrund von fehlender Information über die eingesetzten Software-Programme kaum verfolgt werden kann (vgl. Oxera 2017).

Für die Untersuchung des Outputs von Preisalgorithmen gibt es zwei mögliche Zugänge, die in diesem Projekt beide ausgeschöpft werden sollen. Es kann über Echt-User gearbeitet werden – geläufig sind Kleingruppen für experimentelle Tests oder systematisch rekrutierte User-Gruppen. Es können auch Echt-User modelliert werden, um mögliche relevante Faktoren für die Preisbildung besser zu untersuchen. Im vorliegenden Bericht wird dargestellt, wie PRIMMING diese zwei Erhebungsmethoden verfolgt und ineinanderfließen lässt – das Echt-User Panel sowie die Abfrage über Personas und Szenarien, die darauf aufbauen soll.

Die Nutzung von **Echt-User** Daten zur Messung von personalisierter Preisdiskriminierung erlaubt es die Auswahl von Shops nicht entlang von Warenkörben oder Branchen zu treffen, sondern auf Basis gemessener Unregelmäßigkeiten bei den Preisen. Auf diese Art und Weise werden verdächtige Shops ermittelt, die durch einen Crawler zunächst in erhöhter zeitlicher Frequenz und aufbauend darauf durch andere Faktoren genauer untersucht werden. Dieser Ansatz ist bezüglich der Menge an Datenpunkten ein Unikum.

Die **Personas und Abfrage durch Szenarien** ist ein experimentell daran anknüpfender Vorgang, um der Komplexität der datenbasierten Preisgestaltung gerecht zu werden. Ausgehend von dem Wissen zu gesammelten Daten und möglichen Kombinationen zwecks Personalisierungen im Internet (z.B. für Werbung, individualisierte Suchergebnisse, etc.), geht es darum Echt-User ähnliche Personas zu modellieren, die auch in ihrem Verhalten möglich individuell online nach Produkten suchen und diese erwerben.

1. Die Nutzung von Echt-User Daten

1.1. Einleitung

Das PRIMMING-Projekt arbeitet mit Echt-User-Daten, die von der Browser-Erweiterung „Ciuvo“ bereitgestellt werden. Ciuvo ist ein mobiler Preisvergleich – für das Projekt werden die Daten über die angezeigten Preise gespeichert und hin auf Unregelmäßigkeiten analysiert. Damit unterscheidet sich der Zugang des Projekts wesentlich von bisherigen Erhebungen in der Funktionsweise der Echt-User-Einbindung. Im Folgenden finden Sie eine Zusammenfassung von Erhebungen, die eine vergleichbare Fragestellung mit Echt-User-Daten behandelten.

1.1.1. Andere Forschungsvorhaben mit Echt-User-Ansätzen

Mikians et al. arbeiteten im Jahr **2013** mit der Browser-Extension „\$heriff“, die einen mobilen Preisvergleich ermöglicht. Der User markiert den Preis und dieser wird von 14 unterschiedlichen Standorten abgefragt und mögliche Unterschiede werden mitgeteilt. Die ForscherInnen arbeiteten mit einem Datensatz aus Jänner bis Mai 2013 von 1.500 Abfragen von insgesamt **340 Echt-Usern**.

Im Jahr **2014** ermittelte die Studie „Measuring Price Discrimination and Steering on E-commerce Web Sites“ – anhand von **300 Echt-Usern** – Daten zu Preisdiskriminierung und Steering (je 100 User für E-Commerce, Hotels und Leihwägen). Die Accounts wurden über Amazon Mechanical Turk angeworben – einer Vermittlungsplattform für digitale Arbeit. Sie wurden jeweils um das Ausführen einer Human Intelligence Task (HIT) gebeten, bei der sie eine Seite aufrufen sollten und ihre Daten wurden aufgezeichnet. Ihre Cookies im Web-Browser blieben bestehen (vgl. Hannak et al. 2014).

Iordanou et al. nutzten im Jahr **2017** wie Mikians et al. auch die Browser-Extension (\$heriff) mit u.a. Echt-User-Daten aus einem Jahr – insgesamt 5.700 Anfragen, für rund 4.800 Produkte von 1.994 Websites.

Im Jahr **2019** erhoben deutsche ForscherInnen u.a. mit Studierendengruppen (Sample 61 StudentInnen) Preise von Online-Shops. Es wurden demographische Angaben und Verhaltensweisen der teilnehmenden Studierenden erhoben, wie z.B. Alter, Hobbys, Apps, Gerätetyp (vgl. Badmaeva/Hüllmann 2019).

1.1.2. Bedeutung von Echt-User Daten

Durch die Nutzung von Echt-User Daten im Projekt wird es möglich das Forschungsvorhaben anhand von Verdachtsfällen zu konzipieren. Als weitere Stärke hinzu, dass es nicht von Anfang an notwendig wird die Komplexität von Echt-Usern zu modellieren.

Diese Form der Echt-User Einbindung hat als größte Schwäche, dass sie keine Kontrolle über die KonsumentInnen-Eigenschaften bringt. Dies gilt allgemeiner auch für bisherige Echt-User Panels, wie im Vorhergehenden exemplarisch an anderen Erhebungen gezeigt. Es konnten auch in diesen Fällen nur eingeschränkt Standort oder Browser Settings und Verhalten auf der Seite kontrolliert werden.

Echt-User Panels ermöglichen es dynamische Preisdifferenzierungen zu erheben, doch welche Faktoren für eine möglicherweise personenbasierte Preisdifferenzierung ausschlaggebend sind, bleibt schwierig zu ermitteln. Um dem zu begegnen können die Daten mit Multipler Regression analysiert werden. Der Preis wird mit variierenden Faktoren zum User in Verbindung gesetzt. Ebenso kann mit experimentellen Settings der Impact bestimmter Faktoren getestet werden. Laut Hindermann (2018) ist dies der geläufigste Umgang von Forschungsvorhaben mit dem Problem der Faktorermittlung bei Echt-User Panels.

Im Vorliegenden Projekt wird dem Problem über die Erweiterung des preis.wert-Tools um Personas begegnet.

1.1.3. Ciuvo als begleitender Preisvergleich

Ciuvo ist eine Browser-Erweiterung für den ad-hoc Preisvergleich. Sobald ein User auf eine Produktseite eines Onlineshops kommt, werden alternative, günstigere Angebote, Gutscheine oder zusätzliche Information gesucht. Der User erfährt in Millisekunden bis einigen Sekunden über die Ciuvo Browser-Erweiterung im Web-Browser, ob es bessere Angebote gibt.

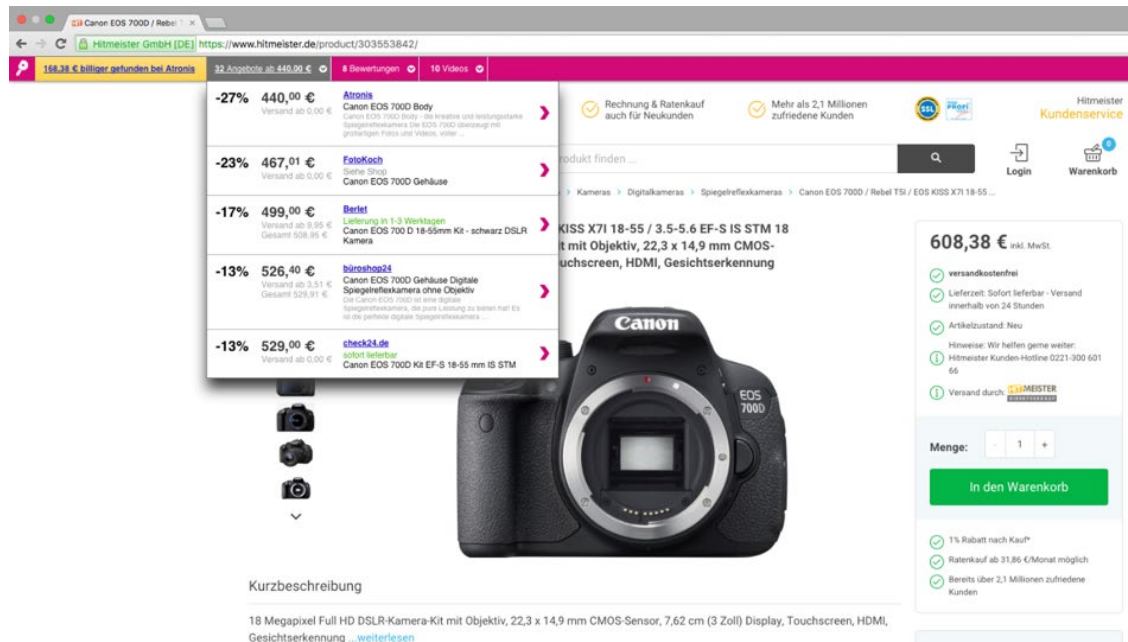


Abbildung 1: Screenshot zur Ciuvo-Browser Extension

Wie funktioniert die Ciuvo Browser Extension

Nach der Installation der Browser-Erweiterung durch den User, holt sich die Software erstmalig und später regelmäßig eine Whitelist vom Ciuvo-Server ab. Auf dieser Whitelist sind mehr als 70.000 Onlineshops weltweit vermerkt. Die Erweiterung wird nur aktiv, wenn die besuchten Domains auf der Whitelist stehen. Dies soll sicherstellen, dass die Browser-Erweiterung keine Daten von nicht relevanten Webseiten schickt.

Ist die Domain auf der Whitelist, fragt der Client beim Server nach dem aktuellen Scraper für eben diese Domain an. Die Scraper sind in einer eigens entwickelten Markup-Sprache geschrieben, da die Browser aus Sicherheitsgründen kein nachgeladenes JavaScript über Browser-Erweiterungen erlauben.

Content Scrapers Language

▶ Example Scraper

```

1 $title = trim(sizzle('meta[itemprop="name"]:first', 'content')) // FUZZY
2 $price = trim(sizzle('meta[itemprop="price"]', 'content'))
3 $availability = sizzle('meta[itemprop="availability"]', 'content')
4 $canonical_url = sizzle('link[rel="canonical"]', 'href')
5 $currency = sizzle('meta[itemprop="priceCurrency"]', 'content')
6 $image = sizzle('img[itemprop="image"]', 'src')
7 $identifier = sizzle('itemprop="productID"', 'textContent')
8 $brand = sizzle('itemprop="brand">[itemprop="name"]', 'content')
9 $condition = sizzle('meta[itemprop="itemCondition"]', 'content')
10
11 // Mobile
12 if (not $title) {
13   $title = trim(sizzle('.pdp-product-title', 'textContent'))
14   $price = sizzle('.basePrice:first', 'textContent')
15   $image = sizzle('.pdpImage:img:first', 'src')
16   $identifier = sizzle('input[name="productids"]', 'value')
17 }
18
19 require $price, $title
20 return $availability, $brand, $canonical_url, $condition, $currency, $identifier, $image, $price, $title
    
```

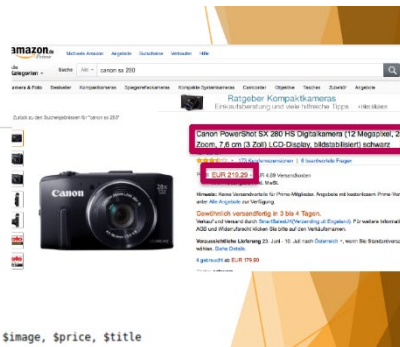


Abbildung 2: Beispiel: Scapercode in CSL (Content Scaper Language)

Die Scraper extrahieren relevante Daten wie Produkttitel, Beschreibung, Bild-URL oder IDs von der Shop-Seite und der Client übermittelt diese an den Server. Dort wird eine Vielzahl an Datenbanken und auch Partner-APIs abgefragt, um das Produkt zu identifizieren und entsprechende gleiche Angebote von anderem Händler zu finden. Diese werden an den Client gestreamt und dieser stellt sie als Preisvergleich im Browser dar.

Backend interaction overview

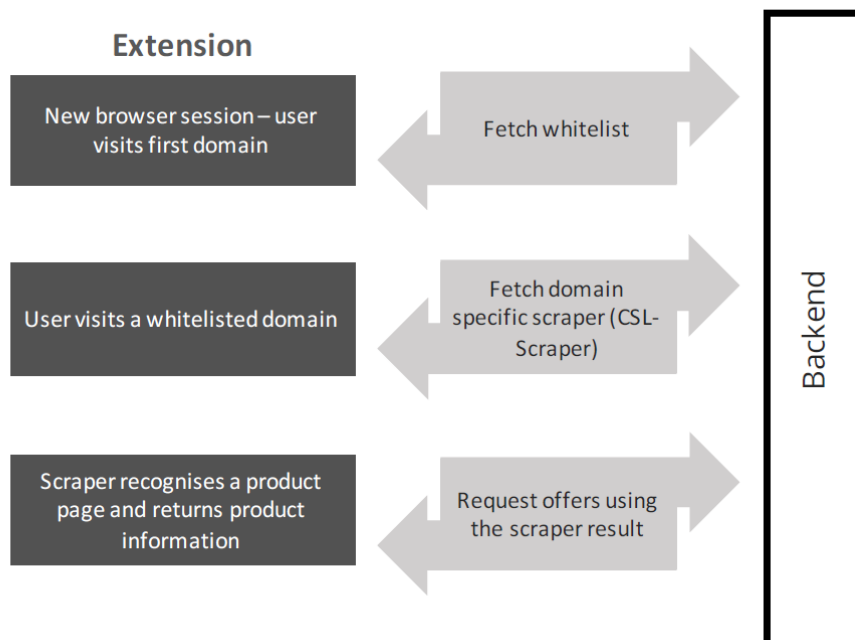


Abbildung 3: Ablauf Client-Server Interaktion bei Ciuvo

Ciuvo hat mehr als 30.4 Millionen aktive User weltweit monatlich. Diese User generieren täglich durchschnittlich 10 Millionen Touchpoints mit Shopseiten, Produkten oder anderen Angeboten. Ciuvo ist in 40 Ländern aktiv. Das bedeutet, in diesen Ländern werden Domains (70.000+) erfasst und auf diesen ist das Angebot von Ciuvo verfügbar. Die größten Märkte (wirtschaftlich betrachtet) sind hier USA, Frankreich, United Kingdom, Deutschland und Brasilien.

Aufgrund des grenzüberschreitenden Online-Traffics und damit verbundenen E-Commerce Aktivitäten hat Ciuvo aber User aus 244 Ländern registriert, wobei ca. 8% der User nicht aus den aktiven 40 Ländern stammen. Die Länder mit den meisten User sind, Brasilien, Frankreich, USA, Deutschland und Polen. Die User verwenden weltweit hauptsächlich Google Chrome als Webbrowser, was sich mit 74% Anteil bei unseren Usern niederschlägt. Weitere 23% nimmt Mozilla Firefox ein. Die restlichen 3% teilen sich auf Opera, IE/Edge und Safari. Die Nutzung der Browser ist durch unsere Verfügbarkeit eingeschränkt (neben den erwähnten werden keine anderen Browser unterstützt - inoffiziell alle Chromium Varianten).

1.2. Erhebung

Die Erhebung verläuft folgendermaßen:

1. Schritt: Echt-User Daten Sammlung durch die Ciuvo Browser-Extension
2. Schritt: Ermittlung von Unregelmäßigkeiten im Sinne häufiger Preisveränderungen
3. Schritt: Ausschluss jener Produkte/Shops, deren Preisveränderungen durch z.B. unterschiedliche Mengenangaben bedingt sein könnten
4. Schritt: Einbettung der zu untersuchenden Produkte und Shops in das preis.wert Erhebungsframework
5. Schritt: Systematische Untersuchung und limitierter Faktoren-Varianz durch das preis.wert Tool. In dieser Phase werden die verdächtigen Produkte einzelner Shops mit einer erhöhten Intensität durch den Crawler bezüglich ihrer Preisbildungsdynamiken untersucht. Geplant wird ein Abruf des Preises alle 20 Minuten – es werden regelmäßige Wechsel an untersuchten Produkten vorhersehen.

1.2.1. Der preis.wert-Crawler

Das preis.wert Tool wurde für den Einsatz in einer Linux Umgebung entwickelt und unter Ubuntu 18.04.3 LTS getestet. Alle im Einsatz befindlichen Komponenten stehen auch für einen Betrieb unter Windows mit geringem Adaptionaufwand zur Verfügung. Entwickelt wurde der Crawler im Rahmen des netidee-Projekts (preis.wert¹) im Jahr 2018/2019. Die bisherigen Funktionskomponenten umfassen:

- User Agents: Auswahl unterschiedlicher Endgeräte für Preisabfrage z.B. Windows Notebook, Apple Notebook, iPhone und Android Phone
- VPN: Um Abfragen aus unterschiedlichen Geolocations durchzuführen, können Virtual Private Networks (VPNs) angebunden werden.
- Cookies: Steuerungsmöglichkeit ob dem Webshop Informationen, die dieser in vorangegangenen Besuchen im Browser des Benutzers abgelegt hat, übermittelt werden oder nicht.

¹ <https://netidee.at/preiswert>

- Scheduling: Es kann die Häufigkeit der Abfragen reguliert werden.

1.2.2. Vorgehensweise

In dieser Phase ist das Ziel, die Datengrundlage in Form von Preismesspunkten der potenziell verdächtigen Produkte und Shops zu verdichten. Hierbei wird auf den Einsatz des Bestehenden preis.wert Crawlers unter Berücksichtigung der User-Agent, VPN und zurückgegriffen. Die Abfragen der einzelnen Produkte und ihrer Preise in den einzelnen Dimensionen erfolgt unmittelbar aufeinanderfolgend.

Tabelle 1 Dimensionen der abgefragten Preispunkte

Dimension	User-Agent	Abfrage Standort	Cookie Verhalten
WWO	Windows Notebook	Wien	Ohne Cookies
WWC	Windows Notebook	Wien	Mit Cookies
WGO	Windows Notebook	Graz	Ohne Cookies
WGC	Windows Notebook	Graz	Mit Cookies
AWO	Apple Notebook	Wien	Ohne Cookies
AWC	Apple Notebook	Wien	Mit Cookies
AGO	Apple Notebook	Graz	Ohne Cookies
AGC	Apple Notebook	Graz	Mit Cookies
CPOINT	Linux User-Agent	Wien	Ohne Cookies

Die Adaptierung der Crawler-Software zur Preis-Extraktion in Form der benötigten CSS und XPATH Selektoren aus den über 80 unterschiedlichen Shopsystemen und Produktseiten erfolgt durch Integration der CIUVO Library. Neben den acht Dimensionen, die über den Crawler direkt auf dem AIT Server erhoben werden, erfolgt Messung eines zusätzlichen Kontrollpreispunktes CPOINT am CIUVO Backend nach jedem Abfragezyklus. Dazu wird das Preisvergleichs Plugin über Selenium, ein quelloffenes Framework für automatisierte Softwaretests von Webanwendungen, in einer Google Chrome

Instanz bereitgestellt und die Produktseite aufgerufen. Die Auswertung und Visualisierung der Messergebnisse erfolgen durch das preis.wert Dashboard.

1.2.3. Ausblick auf mögliche Ausbaustufen der Echt-User Datensammlung

Als mögliche Alternativen zu dem Echt-User Datenpool, das von PRIMMING genutzt werden könnte, gelten User-Panels, die eine genauere Definition von User-Angaben zulassen würden. Zu den Schwächen des gewählten Ansatzes zählt zum Beispiel, dass personenbasierte Daten zu den Usern in dieser Form nicht erhoben werden. Es lassen sich also keine Aussagen darüber treffen, welche personenbasierten Diskriminierungen vorliegen. Eine Möglichkeit wäre eine **eigene Version auf Basis der CIUVO Extension** zu veröffentlichen, bei der User ihre Zustimmung zu einer ausgeweiteten Angabe und Erfassung von personenbasierten Daten zu Forschungszwecken geben.

2. **Personas und Abfrageszenarien**

2.1. *Methodische Vorüberlegungen zu Personas*

Das Projekt PRIMMING untersucht dynamisch generierte Preisdifferenzierungen online, mit dem Ziel durch innovative Messmethoden Wissen zu relevanten personenbezogenen Faktoren bei der Preisbildung zu generieren. Im Projekt geht es also nicht um „Gender Pricing“ im Sinne von weiblich oder männlich vermarkteter, sehr ähnlicher Produkte. Es ist weniger davon auszugehen, dass direkte Formen der Diskriminierung nach Geschlecht zum Tragen kommen, als dass es sich um indirekte Diskriminierungen handeln könnte.

Die Untersuchung von McDonald (2015) zeigt, inwiefern es bei Versicherungsverträgen zu indirekten Formen der Diskriminierung nach dem Faktor Geschlecht über berufliche Faktoren kommt. Um die Auswirkungen von algorithmischer Preisbildung zu untersuchen, wird im Folgenden ein intersektionaler Zugang gewählt, der sich die Überschneidungen von Diskriminierungsformen ansieht. Notwendig wird dies, weil dynamische Preissetzungen auf komplexe Informationen über eine Person und Preisbereitschaften zu. Die deutsche Antidiskriminierungsstelle verweist darauf, dass diese Informationen zeitnah und bezogen auf bestimmte Produkte deutlich präziser wären (2017). Es wird deshalb mit Personas gearbeitet, die nicht als KonsumentInnen-Gruppen konzipiert werden bezogen auf ihre Zahlungsbereitschaft, sondern umfassender eine Vielzahl an Faktoren einbezieht wie Angaben zu Browser, Verhalten auf der Website, oder Interessen.

Für die Erarbeitung von Personas wurde am 9. Dezember 2019 ein Workshop mit neun Expertinnen und Experten zur Bildung von Personas und User Journeys durchgeführt. Neben dem Konsortium nahmen der AK-Experte Michael Dunkl (Digitalisierung und Konsumentenschutz) und Sabine Rada (Wirtschaftskammer Österreich) teil. In einem ersten Schritt ging es um die Bildung von Personas.

Das dabei verfolgte Ziel war es Personas zu modellieren, die der Komplexität von Echt-User-Profilen gerecht werden. Gewählt wurde dafür ein Reverse Engineering-Ansatz. Angeleitet von Fragen wie „Wie drückt sich ein Einkommen in Online-Verhalten aus?“, „Woran zeigt sich eine zögerliche Persönlichkeit beim Einkaufen online?“ oder „Wie erkennt ein Unternehmen, dass jemand bald eine Wohnung neu einrichten muss?“, sollten Personen modelliert werden.

Ausgehend von den Arbeiten zu Datensammlungen online (Christl 2017); Im Folgenden sind dazu einige Beispiele für die Detailnotwendigkeiten angegeben.

- **Gender**
- **Sexuelle Orientierung**
- **Beruflicher Status**
- **Genutzte Geräte:** Smartphone, Computer, Laptop, Tablet, Smart-Watch, Smart Home
- **Betriebssysteme**
- **Browser**
- **Spracheinstellungen**
- **Internet service provider**
- **Cookies:** Ja/nein
- **Permissions:** geolocation, notifications, persistent-storage, push
- **Sichtbarkeit von „Bars“ im Browser:** Menü, Personal, Status, Tool.
- **Meist genutzte Apps:** Office, Kalender, Internet, Kamera, Video/Musik.
Ausgehend von den meist genutzten Apps, können Schlussfolgerungen zur Persönlichkeit von Menschen getroffen werden.
- **Anrufverhalten:** Anhand der Anzahl verpasster Anrufe, der Anzahl einzigartiger Anrufkontakte, der Dauer der eingehenden Anrufe, können Schlussfolgerungen zur Persönlichkeit gezogen werden.
- **Texting-Verhalten:** Länge der gesendeten Nachrichten, Unique Contacts für versandte Nachrichten, Response Rate, Intervall zwischen Anruf/Textnachrichten, Prozentanteil der selbst-initiierten Kommunikation
- **Websiteauswahl:** Anhand dessen ob jemand z.B. bei DIY die Website deviantART.com, Tumblr.com oder Etsy.com besucht, lassen sich Schlussfolgerungen zur Persönlichkeit treffen. Das betrifft z.B. die Auswahl von Streaming Dienste
- **Soziale Netzwerke**
- **Verwendete Suchmaschine**
- **Schlüsselmomente im Leben:** Marketinganalysen erlauben es wichtige Momente im Leben von Konsumentinnen und Konsumenten zu identifizieren, ab denen ihr Verhalten flexibel wird und sich ihr Verhalten für längere Zeiten langfristig verändern könnte. Beispiele dafür sind: die Geburt eines Kindes,

eine Heirat bzw. Scheidung, ein Umzug, Studienbeginn. Solche Momente können anhand des Surfverhaltens bzw. Online-Aktivitäten vorausgesagt werden.

- **Interessen:** Für personalisierte Werbung auf Google wird z.B. anhand von gesuchten Begriffen, besuchten Websites und Emails etc. auf Interessen geschlossen. Eine Auswahl möglicher Interessen, die als solche darin gelistet sind: Autovermietung und Taxidienste, Berg- und Skigebiete, Computer- und Videospiele, Familienstand, Bücher und Literaturgenre, Comics und Zeichentrick, Branche, Audiogeräte, Serien/Filmgeschmack, Fahrzeuge, Fahrzeugkauf, Fastfood/Feinkost, Fußball, Frauenspezifische Themen, Indie- und Alternative Musik, Internationale Nachrichten, Fotografie und digitale Kunst, Fitness, etc.

2.2. Modellierung von Personas

Exemplarische andere User-Modellierungen aus anderen Projekten:

In Mikians et al. (2012) wurden zwei entgegengesetzte Segmente an Personas trainiert. Unterschieden wurde zwischen wohlhabenden und preisbewussten Konsumentinnen und Konsumenten (engl. „affluent and budget customers“). Während die preisbewussten User zunächst Preisvergleichsseiten besuchten, gingen wohlhabende User auf Luxusproduktseiten. Anhand von US-Daten zu den von Usern dieser Segmente besuchten Websites wurde eine Woche lang ein System trainiert, das zwischen 9 Uhr morgens bis Mitternacht diese besuchte. Es wurde sichergestellt, dass das Setzen von Cookies möglich sei. Als zweites wurde der „Referer“-Header genutzt, der anzeigt woher eine Anfrage kommt.

In ihrer Erhebung gingen Vissers et al. (2014) den Anekdoten nach, dass Flugpreise personalisiert werden. Dafür erstellten sie insgesamt 66 Unique User Profile. Ihr Scraper wurde für sieben verschiedene Profile emuliert (Internet Explorer 9 auf Windows 7, Chrome 29 auf Windows 7, Firefox 22 auf OS X 10.7, Safari 6 auf OS X 10.7 und Safari Mobile 5 auf iOS iPad). Zusätzlich wurden mit dem Do Not Track (DNT) Header Profile zu Chrome, Safari und Internet Explorer gemacht.

Gearbeitet wurde mit drei KonsumentInnen-Profilen: wohlhabend, budget und FlugvergleicherInnen. Für jedes dieser Profile wurde ein getrennter Cookie Jar erstellt. Es wurden vier verschiedene Cookie Settings gesetzt: (1) nur KonsumentInnen-Profil Cookies (zuvor gesammelt), (2) KonsumentInnen-Profil

Cookies with First-Party Cookies², (3) nur First Party Cookies (keine Third Party Cookies) und (4) keine Cookies. Der Scraper wurde von zwei Locations aus betrieben (New York/USA und Leuven/Belgien)

Schleusener and Hosell (2015) erstellten eine Studie und Gutachten zu personalisierter Preisbildung im Auftrag des Sachverständigenrats für Verbraucherfragen. In diesem Zusammenhang erstellten sie drei Profile von Konsumentinnen und Konsumenten: (1) Luxus, (2) Normal, (3) Clean. Es wurden das genutzte Endgerät und Betriebssystem, die Suchhistorie des Users, Browserverlauf, Cookies, Verwendung von AdBlockern und unterschiedliche Standorte in Deutschland verwendet.

2.3. *Methodische Vorüberlegungen zu Abfrageszenarien*

Die bisherigen Untersuchungen zu personalisierter Preisdiskriminierung treffen Unterschiede teilweise im Sinne von Gruppen und ihrer Zahlungsbereitschaft (z.B. Luxus, Budget), eingeschränkt bezogen auf ihre Interessen (Besuch einer Website eines Automobilmagazins, danach Autohandel-Website) und die Dauer zwischen erstmaligem Besuch und Abschluss eines Kaufvertrags (z.B. eine Woche).

Darüberhinausgehend möchte das vorliegende Projekt der Komplexität der Datenanalyse, wie sie aktuell möglich wäre und insofern für Preisgestaltung online angewandt werden könnte, gerecht werden. Dafür wurden Journey Maps erstellt – die bei den ersten Handlungen starten, die ein Interesse für einen Kauf/Buchung ausdrücken und beim Abbruch bzw. der Bestellung/Buchung selbst enden.

Die Journey Maps wurden anhand von zuvor erstellten Personas in 2er bzw. 3er Teams erstellt. Jeder Schritt sollte dokumentiert werden anhand der folgenden Dimensionen:

- **Status:** Um welchen Teil des Prozesses handelt es sich z.B. Exploration von Möglichkeiten, Entscheidungsfindung, Abwägungen, Zweifel, Verwerfen, etc.
- **Handlung:** Was macht die Person in diesem Moment konkret? Bitte achten auf eine technisch übersetzbare Tätigkeit z.B. Person klickt auf den zweiten Link, klickt nach kurzem Blick auf zurück, klickt auf den dritten Link.

² In dem Fall, die von der Airline selbst gesetzten Cookies.

- **Ziel:** Notwendig für die Plausibilitätsüberprüfung im Nachhinein - welches Ziel verfolgt die Person in diesem Moment?
- **Ort:** Auf welcher Website/App befindet sich die Person, wenn sie diese Handlung setzt?
- **Zeitliche Dimension:** Wie lange braucht die Person für diese Handlung/Handlungsabfolge? Wie viel Zeit ist vergangen seit dem letzten Schritt?
- **Parallele Aktivitäten online:** Welche anderen Schritte werden parallel unternommen? Chattet die Person gerade in einem anderen Tab? Wird etwas unterbrochen durch eine andere Handlung, z.B. Check des Instagram Profils?
- **Technisches:** Welcher Browser wird genutzt? Welches Gerät wird genutzt? Welche Angaben sind in diesem Zusammenhang noch wichtig? (z.B. werden die Cookies gelöscht, sobald der Browser sich schließt)
- **Notizen:** Bitte um ergänzende Angaben zu möglicherweise wichtig zu beachtenden Details, zusätzliche Anmerkungen zu den Handlungen, etc.

2.4. Modellierung von Abfrageszenarien

Im Folgenden geben wir einen exemplarischen Ausblick auf die geplante Überführung ausgewählter, in den Personas Creation Workshop erhobenen, Verhaltensweisen im Sinne ihrer möglichen technischen Überführung.

Tabelle 2 Exemplarische Überführung Personas in technische Abläufe

Gewünschtes Verhalten	Technische Übersetzung	Technische Vorbedingungen	Technische Abbildung
Benutzer zeigt Interesse für ein Produkt eines Webshops	Produktseite (Input) wird als URL im Browser aufgerufen, das Produkt über den Button in den Warenkorb gelegt; 5 Minuten gewartet; der Warenkorb angeklickt und das Produkt entfernt.	Existierende Geräte Konfiguration User Account angelegt und Benutzer eingeloggt.	Der Ablauf wird mittels Selenium Browsersteuerung in zwei Einzelabläufen im Webshop aufgezeichnet und Test zur Bestätigung der erfolgreich durchgeführten Handlungen werden festgelegt. Die Session wird nicht beendet.

Shop wird via Handy besucht	Startseite des Shops (Input) wird über den Browser eines Android Geräts aufgerufen	Emuliertes Android Gerät Existierende Geräte Konfiguration	Appium Framework: Android Device erstellen und Google Chrome App installieren
Maria, 65 Jahre, mittleres Einkommen, technisch unterfahren, Interesse für Schmuck	Alter über Informationen in User Accounts und Font-Einstellung des Betriebssystems Technischen Skills über Wahl des Betriebssystems, Suchmaschinenwahl und langsamen Surfverhalten Interesse für Schmuck durch Besuch von Webseiten	Betriebssystem Image konfiguriert User Accounts für Bing und Amazon angelegt Suchmaschinen Abfrage aufgezeichnet	Dieser Ablauf ist ein Beispiel für ein Endprodukt zusammengesetzter Workflows, die als Vorbedingungen einer eigentlichen Handlung durchgeführt wurde

2.5. Ausblick: Stufenweise Vorgehensweise

Die in Tabelle 2 Exemplarische Überführung Personals in technische Abläufe beispielhaft gezeigte Übersetzung zeichnet sich von links nach rechts durch eine Steigerung im Grad der benötigten Detaillierung, Bereitstellung technischer Systemen und deren Interaktion zur Erfüllung und Messung von Zuständen. Zustände liegen entweder lokal vor (z.B. Konfiguration eines Browsers) oder online, wobei hier nochmals zwischen explizit vorhandener Information (z.B. hinterlegten Zahlungsdaten in einem Shop oder Kundenprofil) und implizit, nicht direkt messbarer Informationszuständen (z.B. gespeicherte Daten durch Cookies oder getrackte Interessen durch Werbenetzwerke, Blackbox Online-Shop woher bezieht dieser Informationen über einen Benutzer) zu unterscheiden ist. Aktionen sind kontrollierbare Handlungen, die zu Zustandsänderungen führen.

Das Projekt nimmt hierbei Anleihe aus dem Feld des Business Process Modelling: „Das Informationsmodell beschreibt, was Dinge (Entitäten) fachlich miteinander zu tun haben. Wie sie strukturell zusammenhängen. Es beschreibt die möglichen, konsistenten und relevanten Zustände, die statische Struktur eines Systems.

Ein Prozess überführt ein System von einem konsistenten Zustand in einen anderen. Prozesse sind nie das Ziel, sie sind immer der Weg. Prozesse beschreiben, wie eine Zustandsänderung herbeigeführt wird. Sie beschreiben nicht die Zustände als solche. Das Business Process Model beschreibt, was Akteure wie mit den Dingen (Entitäten) tun. Es beschreibt, wie durch eine

Aktivität, unter Einsatz von (technischen) Hilfsmitteln, das System von einem konsistenten Zustand in einen anderen übergeht.“ Stefan Berner (2019)

PRIMMING verfolgt eine schrittweise Modellierung des Informationsmodells, Zustandssystems und Prozesse sowie deren technische Umsetzung.

- **Abbildung möglichst simpler Aktionen** (z.B. Einloggen in einen bestehenden Account eines Webshops). Definition messbarer Zustände. Technische Implementierung (z.B. Workflow in Selenium)
- **Verkettung von technischen Systemen** (z.B. Erstellen eines Android Emulators und anschließendem Aufruf einer Produktseite in Google Chrome). Definition von Workflows, Schnittstellen, I/O & Konfigurationsmöglichkeiten.
- **Toolgestützte Modellierung** der Informationsmodelle, Zustandssysteme und Business Process Execution Language (BPEL) zur Beschreibung der einzelnen Aktivitäten.

Die schrittweise Erarbeitung des Informationsmodells, der Definition messbarer Zustände und kombinierbaren Aktionskatalogs erlaubt einerseits den Grad der Komplexität und Detaillierung im Laufe des Projektes punktuell zu steigern sowie gemessene Preispunkte und deren Effekte eindeutig und wohldefinierten expliziten Systemzuständen und Handlungen zuzuordnen. Das Ziel von PRIMMING ist es eine Toolbox zu erstellen, die es Nicht-Technikern erlaubt Personas in Form ihrer Handlungen im Netz (durch Auswahl an vorhandenen Interaktionsmechanismen) zu definieren und zur Preismessung einzusetzen.

3. Quellenverzeichnis

3.1. Literatur

- Antidiskriminierungsstelle** (2017): Preisdifferenzierung nach Geschlecht in Deutschland. Forschungsbericht. Eine Studie im Auftrag der Antidiskriminierungsstelle des Bundes, durchgeführt von Heiden, I./Wersig, M. Fachhochschule Dortmund.
- Hannak, Aniko, Gary Soeller, und David Lazer (2014)**. „Measuring Price Discrimination and Steering on E-Commerce Web Sites“, 14.
- McDonald, S.** (2015): Indirect Gender Discrimination and the „Test-Achats Ruling“: An Examination of the UK Motor Insurance Market, Newcastle University Economics.
- Mikians, Jakub & Gyarmati, László & Erramilli, Vijay & Laoutaris, Nikolaos. (2013)**. Crowd-assisted Search for Price Discrimination in E-Commerce: First results. 10.1145/2535372.2535415.
- Mikians, Jakub, László Gyarmati, Vijay Erramilli, und Nikolaos Laoutaris.** „Detecting Price and Search Discrimination on the Internet“. In Proceedings of the 11th ACM Workshop on Hot Topics in Networks - HotNets-XI, 79–84. Redmond, Washington: ACM Press, 2012. <https://doi.org/10.1145/2390231.2390245>.
- Oxera. „When algorithms set prices: winners and losers“, 2017. <https://www.oxera.com/wp-content/uploads/2018/07/When-algorithms-set-prices-winners-and-losers.pdf.pdf>.
- Schleusener, Michael, und Sarah Hosell.** „Expertise zum Thema ‚Personalisierte Preisdifferenzierung im Online-Handel‘“. Studien und Gutachten im Auftrag des Sachverständigenrats für Verbraucherfragen. Berlin, 2016. https://www.svr-verbraucherfragen.de/wp-content/uploads/eWeb-Research-Center_Preisdifferenzierung-im-Onlinehandel.pdf.
- Stefan Berner (2019)**. „BPM considered harmful: Warum Business Process Modeling nicht am Anfang einer Anforderungsdefinition stehen sollte, Mai 2019“. In: Lecture Notes in Informatics: Modellierung 2016. <https://foryouandyourcustomers.com/magazin/bpm-considered-harmful-warum-business-process-modeling-nicht-am-anfang-einer-anforderungsdefinition-stehen-sollte/>
- Thomas Vissers, Nick Nikiforakis, Nataliia Bielova, Wouter Joosen (2014)**. Crying Wolf? On the Price Discrimination of Online Airline Tickets. 7th Workshop on Hot Topics in Privacy Enhancing Technologies(HotPETs 2014), Jul 2014, Amsterdam, Netherlands.
- Tsagana, Badmaeva, und Joschka Andreas Hüllmann (2019)**. „Investigating Personalized Price Discrimination of Textile-, Electronics- and General Stores in German Online Retail“. In: 14th International Conference on Wirtschaftsinformatik, February 24-27, 2019, Siegen, Germany