

Landesweites Digitales Besuchermanagement für den Tourismus in Schleswig-Holstein (LAB-TOUR SH)

Bericht

Wir fördern Wirtschaft



Durch die Europäische Union - Europäischer Fonds für regionale Entwicklung (EFRE), REACT-EU als Teil der Reaktion der EU auf die Covid-19-Pandemie finanziert.

Schleswig-Holstein
Der echte Norden

Deutsches Institut für Tourismusforschung

Fachhochschule Westküste
Fritz-Thiedemann-Ring 20
25746 Heide
Ansprechpartner: Prof. Dr. Julian Reif
Mail: reif@fh-westkueste.de
Telefon: 0481 8555 573

FuE-Zentrum Fachhochschule Kiel GmbH

Schwentinestr. 24
24149 Kiel
Ansprechpartner: Prof. Dr. Michael Prange
Mail: michael.prange@fh-kiel.de

ADDIX GmbH

Kaistr. 101
24114 Kiel
Ansprechpartner: Björn Schwarze
Mail: bs@addix.net

Lufthansa Industry Solutions AS GmbH

Schützenwall 1
22844 Norderstedt
Ansprechpartner: Moritz Röder
Mail: moritz.roeder@lhind.dlh.de

Herausgeber:

Fachhochschule Westküste
Deutsches Institut für Tourismusforschung
Fritz-Thiedemann-Ring 20
25746 Heide/Holstein

Datum: Juni, 2023

Autor*innen:

Julian Reif, Denise Engelhardt, Tim Staubert, Vincent Krieg, Colin Mill, Nele Höftmann, Boje Sjut, Niklas von Boguszewski, Eric Horster, Michael Prange, Christoph Wemheuer, Jaroslaw Piwonski, Tobias Möller, Frank Radzio, Alexander Brinkmann, Jonas Schoenwald

Inhaltsverzeichnis

1.	Einführung	10
1.1	Projekthintergrund	10
1.2	Zielsetzung des Projektes	11
1.3	Im Projekt betrachtete Use Cases	12
1.4	Berichtsaufbau	13
2.	Grundlagen eines digitalen Besuchermanagements im Tourismus	14
2.1	Datengenerierung	15
2.2	Datenmanagement und -infrastruktur	16
2.3	Recommender	17
2.4	Ausspielung	18
3.	Datengenerierung	18
3.1	Datengrundlagen	19
3.1.1	Lokale Sensorik	19
3.1.2	Globale Sensorik	24
3.1.3	#sh_wlan	29
3.1.4	Weitere Daten	29
3.2	Deskriptive Analyse lokaler und globaler Sensorik	34
3.2.1	Lokale Sensorik	34
3.2.2	Globale Sensorik	42
3.3	Ausgewählte Vergleiche von lokaler und globaler Sensorik	62
3.4	Weitere Analysen	68
3.5	Zwischenfazit	70
4.	Datenmanagement und -infrastruktur	74
4.1	Datenfluss	74
4.2	Datenschnittstellen und Datenhubs	77
4.2.1	Motivation	77
4.2.2	Anforderungen	77
4.2.3	Umsetzung	78
4.2.4	Learnings und Potentiale	82
4.3	Datenmodelle	82
4.3.1	Anforderungen	83
4.3.2	Umsetzung	84
4.3.3	Technische Details	85
4.3.4	Learnings und Potentiale	90
4.4	Integrationsplattform	91

4.4.1	Motivation	91
4.4.2	Anforderungen	92
4.4.3	Umsetzung.....	93
4.4.4	Technische Details	95
4.4.5	Learnings und Potentiale.....	96
4.5	Zwischenfazit.....	96
5.	Recommender	97
5.1	Auslastungsprognose	98
5.1.1	Anforderungen an das Prognosemodul	99
5.1.2	Datengrundlagen für die Prognose	103
5.1.3	Data Cleaning.....	105
5.1.4	Logik und Modelle	112
5.1.5	Technische Details	117
5.1.6	Learnings und Potentiale.....	121
5.2	Alternativengenerierung	124
5.2.1	Anforderungen an einen Recommender.....	125
5.2.2	Datengrundlage	129
5.2.3	Datenbereinigung.....	132
5.2.4	Logik und Modelle	135
5.2.5	Technische Details	154
5.2.6	Learnings und Potentiale.....	156
5.3	Zwischenfazit.....	161
6.	Ausspielung	163
6.1	Konzeptionslogik	163
6.2	Relevanzabschätzung verschiedener Informationskanäle	164
6.3	Analyse der Informationskanäle.....	185
6.4	Deployment des Recommenders in einem exemplarischen Frontend	214
6.5	Pilotierung des Recommenders	218
6.5.1	Das Kampagnensystem	218
6.5.2	Funktionsweise #sh_wlan und Kampagnen	219
6.5.3	Dynamischer Inhalt und das Ausspielen von Empfehlungen	220
6.6	Testszzenarien.....	224
6.7	Zwischenfazit.....	228
7.	Fazit	229
7.1	Kernergebnisse und Forschungsdesiderate	229
7.2	Implikationen für die Forschung: Ausgewählte weitere Ansätze.....	232

7.3	Implikationen für die Praxis.....	233
7.4	Ausblick.....	234

Abbildungsverzeichnis

Abbildung 1: Digitale Besuchermessung und -management als Basis einer smarten und nachhaltigen Destinationsentwicklung.....	14
Abbildung 2: Prozess des digitalen Besuchermanagements.....	15
Abbildung 3: Messstandorte auf Fehmarn.....	20
Abbildung 4: Messstandorte an der Schlei.....	20
Abbildung 5: Messstandorte in Grömitz	21
Abbildung 6: Messstandorte in der Holsteinischen Schweiz	21
Abbildung 7: Messstandorte entlang von Radrouten an der Schlei.....	22
Abbildung 8: Mess-Systeme und deren Anwendung im Tourismus	25
Abbildung 9: Zeitreihe der Summe der täglichen Einfahrten auf Parkplätzen auf Fehmarn	35
Abbildung 10: Zeitreihe der Summe der täglichen Einfahrten auf Parkplätzen in Grömitz.....	35
Abbildung 11: Zeitreihe der Summe der täglichen Einfahrten auf Parkplätzen in Kappeln und Arnis .	36
Abbildung 12: Zeitreihe der Summe der täglichen Einfahrten auf Campingplätzen und Wohnmobilstellplätzen	36
Abbildung 13: Zeitreihe der Summe der täglichen Durchfahrten von Fahrrädern entlang der Radrouten.....	37
Abbildung 14: Summe der täglichen Einfahrten auf Parkplätzen auf Fehmarn nach Wochentagen ...	38
Abbildung 15: Summe der täglichen Einfahrten auf Parkplätzen in Grömitz	38
Abbildung 16: Summe der Einfahrten bzw. Durchfahrten nach Stunden seit Messbeginn je Use Case (Parkplätze Ostsee aufgeteilt nach Fehmarn und Grömitz).....	39
Abbildung 17: Zeitreihe der Summe der täglichen Einfahrten auf Campingplätzen und Wohnmobilstellplätzen nach Fahrzeugtyp und Wochentag.....	42
Abbildung 18: Parkplätze in Grömitz: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppe	43
Abbildung 19: Parkplätze auf Fehmarn: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppen	44
Abbildung 20: Parkplätze und Hotspots an der Schlei: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppen	45
Abbildung 21: Radwege an der Schlei: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppen	45
Abbildung 22: Wohnmobilstellplätze in der Holsteinischen Schweiz: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppe	46
Abbildung 23: Standorte an der Westküste Schleswig-Holstein: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppe.....	47
Abbildung 24: Parkplätze und Strand in Scharbeutz: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppe	47
Abbildung 25: Gini-Koeffizient der Daily visits im Jahr 2022 je Polygon.....	48
Abbildung 26: Daily visits auf den Parkplätzen in Grömitz nach Wochentagen, Typ und Saison	49
Abbildung 27: Daily visits an den Hotspots der Westküste nach Wochentagen, Typ und Saison	50
Abbildung 28: Daily visits an den Hotspots der Westküste nach Stunden, Typ und Saison	50
Abbildung 29: Daily visits auf den Wohnmobilstellplätzen und Campingplätzen nach Stunden, Typ und Saison	51
Abbildung 30: Parkplätze an der Ostsee: Daily visits nach Besuchergruppen	52

Abbildung 31: Parkplätze und Strand in Scharbeutz: Daily visits nach Besuchergruppen	53
Abbildung 32: Parkplätze und Radwege an der Schlei: Daily visits nach Besuchergruppen	54
Abbildung 33: Hotspots an der Westküste: Daily visits nach Besuchergruppen	55
Abbildung 34: Wohnmobilstellplätze und Campingplätze in der Holsteinischen Schweiz: Daily visits nach Besuchergruppen.....	55
Abbildung 35: Wohnmobilstellplätze und Campingplätze in der Holsteinischen Schweiz: Daily visits nach Besuchergruppen und Kalenderwochen	56
Abbildung 36: Quell-Zielgebietsbeziehungen von Tagestourist*innen am Großparkplatz Kappeln	57
Abbildung 37: Quell-Zielgebietsbeziehungen von Tagestourist*innen an der Perlebucht, Büsum.....	58
Abbildung 38: Quell-Zielgebietsbeziehungen von Übernachtungstourist*innen auf dem Campingplatz Spitzenort	59
Abbildung 39: Quell-Zielgebietsbeziehungen von Tagestourist*innen am Strand in Scharbeutz	60
Abbildung 40: Z-Werte von Daily Visits und Incoming-Messwerten Parkplatz Meeschendorf, Fehmarn	63
Abbildung 41: Z-Werte von Daily visits und Incoming-Messwerten am Strand in Scharbeutz.....	64
Abbildung 42: Z-Werte von Daily visits und Incoming-Messwerten in der Perlebucht Büsum	65
Abbildung 43: Z-Werte von Daily visits und Incoming-Messwerten an der Seebrücke in St. Peter-Ording.....	65
Abbildung 44: Volumen von daily visits (Insel Amrum) und Wifi-Signale (Fähranleger Insel Amrum) .	66
Abbildung 45: Korrelation von daily visits (Insel Amrum) und Wifi-Signale (Fähranleger Insel Amrum)	66
Abbildung 46: Z-Werte von Daily Visits und Incoming-Messwerten in der Einkaufsstraße Holm, Flensburg	67
Abbildung 47: Einfahrten Parkplatz Burgtiefe nach Datum	68
Abbildung 48: Einfahrten Parkplatz Burgtiefe nach Temperatur	69
Abbildung 49: Anzahl Device Ids, #sh_wlan clients und visitors in Kiel vor, während und nach der Kieler Woche	70
Abbildung 50: Anzahl unterschiedlicher Clients pro Access Point für einige ausgewählte APs an der Kiellinie vor, während und nach der Kieler Woche 2022	70
Abbildung 51: Datenfluss im Projekt LAB-TOUR SH	75
Abbildung 52: Gesamtübersicht über die Architektur der Integrationsplattform	93
Abbildung 53: Dynamik von Ein- und Ausfahrten als Proxy für die Auslastung eines Parkplatzes.....	107
Abbildung 54: Drifts in den lokalen Sensordaten am Beispiel von Kappeln und Fehmarn.....	108
Abbildung 55: Beispielkurve einer korrigierten Parkplatzbelegung.....	110
Abbildung 56: Korrektur von Auslastungsdaten im Zeitverlauf: Original und Bearbeitung.....	111
Abbildung 57: Vergleich zwischen Realdaten und Modellierung am Beispiel des Scharbeutzer Strandes.....	116
Abbildung 58: Architekturbild des Prognosemoduls.....	119
Abbildung 59: Prognose und Messwerte im Zeitverlauf.....	123
Abbildung 60: Prognose und Messwerte im Zeitverlauf, Juni bis Oktober 2022	123
Abbildung 61: POI in Schleswig-Holstein.....	130
Abbildung 62: Anzahl von Erlebnis-POI nach Typ.....	131
Abbildung 63: Anzahl von funktionalen POI nach Typ	132
Abbildung 64: Beispiel für Konzept- und Entitätsknoten	137
Abbildung 65: Meta-Knowledge-Graph in LAB-TOUR SH.....	138
Abbildung 66: Cluster von Kernbegriffen für Aktivitäten oder POI.....	158
Abbildung 67: Geräte zur Internetnutzung 2022	165
Abbildung 68: Besitz von Endgeräten	165

Abbildung 69: Marktanteil von Endgeräten 2022 in Deutschland	166
Abbildung 70: Gerätenutzung: Tagesreichweiten 2021 nach Alter	167
Abbildung 71: Nutzung von Geräten auf Urlaubsreisen	167
Abbildung 72: Tägliche Nutzungsdauer von Medien	169
Abbildung 73: Nutzung mediales Internet	170
Abbildung 74: Nutzung mediales Internet nach Alter	171
Abbildung 75: Nutzungsdauer mediales Internet	171
Abbildung 76: Nutzungsdauer mediales Internet nach Alter	172
Abbildung 77: Informationsquellen für die Reiseplanung	173
Abbildung 78: Nutzung von Apps im Urlaub 2022	173
Abbildung 79: Smartphone Nutzung im Urlaub	174
Abbildung 80: Nutzung von Informationsquellen zur Information über touristische Angebote/Ausflugsmöglichkeiten am Speicherkoog, in St- Peter-Ording und in Büsum	176
Abbildung 81: Häufigkeit der Nutzung von webbasierten Informationsquellen	177
Abbildung 82: Informationsthemen	178
Abbildung 83: Informationszeitpunkt	180
Abbildung 84: Inspiration und Information bei der Campingplatzwahl durch Offline- Kommunikationswege	181
Abbildung 85: Inspiration und Information bei der Campingplatzwahl durch Offline- Kommunikationswege nach Planung Campingplatzbesuche	182
Abbildung 86: Inspiration und Information bei der Campingplatzwahl durch Online- Kommunikationswege durch Planung Campingplatzbesuche	182
Abbildung 87: Interesse an digitalen Services entlang der Customer Journey	183
Abbildung 88: Inspiration und Information bei der Campingplatzwahl durch Camping-Portale bzw. - apps	184
Abbildung 89: Marktanteil der beliebtesten Suchmaschinen in Deutschland	186
Abbildung 90: Tagesreichweiten von Suchmaschinen in Deutschland	187
Abbildung 91: Relevanz von Suchmaschinen als Kontakthersteller zu Online-Infos	187
Abbildung 92: Nutzung von Angeboten/Apps zur Information über Verbindungen	188
Abbildung 93: Informationsquellen im Internet für eine unbekannte Bahnverbindung in SH	189
Abbildung 94: Informationsquellen von Radreisenden während der Reise.	190
Abbildung 95: Nutzung von Apps beim Wandern nach Altersgruppe	191
Abbildung 96: Nutzung von Apps beim Wandern nach Anbieter	191
Abbildung 97: Sichtbarkeitsindex von Destinationswebsites in Schleswig-Holstein	193
Abbildung 98: Sichtbarkeitsindex von relevanten Informationskanälen/Websites für Parkplätze an der Ostsee	194
Abbildung 99: Sichtbarkeitsindex von relevanten Informationskanälen/Websites für Parkplätze an der Ostsee	195
Abbildung 100: Sichtbarkeitsindex von relevanten Informationskanälen/Websites für Wohnmobilstellplätze an der Holsteinischen Schweiz	196
Abbildung 101: Sichtbarkeitsindex von relevanten Informationskanälen/Websites für Radrouten und Hotspots an der Schlei	197
Abbildung 102: Tourensuche bei Outdooractive	198
Abbildung 103: Tourenplaner bei Outdooractive	198
Abbildung 104: Lübecker Bucht Guide	199
Abbildung 105: Strandticker des Lübecker Bucht Guides	199
Abbildung 106: Ausflugsticker Bayern	200
Abbildung 107: Ausflugsticker Bayern	200

Abbildung 108: Digitaler Reiseführer Wangerland	201
Abbildung 109: Google Popular Times für das Seenotrettungsmuseum auf Fehmarn	202
Abbildung 110: Tourensuche bei Naturtrip für die Region Wilder Kaiser.....	202
Abbildung 111: Tourensuche bei Komoot.....	203
Abbildung 112: Routenplaner bei Komoot.....	204
Abbildung 113: UrbnUps	204
Abbildung 114: CityKey	205
Abbildung 115: ADAC Trips App	206
Abbildung 116: DB Ausflug App	207
Abbildung 117: ByTheWays App	208
Abbildung 118: RegioApp.....	208
Abbildung 119: UmweltNAVI Niedersachsen App	209
Abbildung 120: Peuka Parkplatz-Sharing App	210
Abbildung 121: Frontend: Karte mit POI in Schleswig-Holstein.....	215
Abbildung 122: Frontend: Detailansicht	215
Abbildung 123: Frontend: Detailansicht zum Besuchsaufkommen und generierten Alternativen am Beispiel Fehmarn Parkplatz Burgtiefe Südstrand	216
Abbildung 124: Frontend: Detailansicht zu Attraktionen in der Umgebung	217
Abbildung 125: Website (Kampagne #sh_wlan) mit den Platzhaltern und per Skript generierten dynamischen Inhalten der aktuellen Parkplatzauslastung.....	219
Abbildung 126: Anmeldeprozess an das #sh_wlan inklusive der eventuellen Ausspielung der Anmelde- oder Kampagneseite.....	220
Abbildung 127: Sensorik am Parkplatz Mönkeberg	222
Abbildung 128: Ausspielung einer Kampagne am #sh_wlan Access Point Mönkeberg auf mobilem Endgerät	223
Abbildung 129: Ausspielung einer Kampagne im Rahmen von DorfFunk.....	224
Abbildung 130: TestszENARIO bei Nutzung bzw. Nicht-Nutzung des Recommenders am Beispiel des Use Cases "Parkplätze Ostsee".....	225
Abbildung 131: TestszENARIO bei Nutzung bzw. Nicht-Nutzung des Recommenders am Beispiel des Use Cases "Wohnmobilstellplätze Holsteinische Schweiz"	226
Abbildung 132: TestszENARIO bei Nutzung bzw. Nicht-Nutzung des Recommenders am Beispiel des Use Cases "Hotspots Schlei"	227
Abbildung 133: TestszENARIO bei Nutzung bzw. Nicht-Nutzung des Recommenders am Beispiel des Use Cases "Radrouten Schlei"	228

Tabellenverzeichnis

Tabelle 1: Parameter FLEX Sensor.....	23
Tabelle 2: Parameter ZELT Evo Sensor	23
Tabelle 3: Untersuchungsgebiete in Schleswig-Holstein.....	27
Tabelle 4: Home- und Worklocations der Polygone	28
Tabelle 5: Beispiel visitors und daily visits im Polygon.....	28
Tabelle 6: Im Projekt eingebundene Datenquellen von besucherstrom.info	31
Tabelle 7: Touristische Frequenzdaten in Schleswig-Holstein. Eigene Recherche. Keine Gewähr	34
Tabelle 8: Übersicht der Messstandorte seit Messbeginn	41
Tabelle 9: Top 3-Quellmärkte der als Tages- und Übernachtungstourist*innen klassifizierten Daily visits je Untersuchungsregion im Jahr 2022.....	62
Tabelle 10: Pearson Korrelation und Bestimmtheitsmaß nach Betrachtungsgebiet	63
Tabelle 11: Im Projekt definierte Datenmodelle	86
Tabelle 12: Gemeinsame Attribute in allen definierten Datenmodellen.....	88
Tabelle 13: Attribute im Occupancy Datenmodell.....	89
Tabelle 14: Attribute im OccupancySeries Datenmodell	89
Tabelle 15: Attribute im TriggerRecommendation Datenmodell	90
Tabelle 16: Attribute im TouristRecommendation Datenmodell	90
Tabelle 17: Klassifikation von Ein- und Ausfahrten	106
Tabelle 18: Features des PoC	112
Tabelle 19: Vergleich von Regressionsmodellen	114
Tabelle 20: Beispiel von POI-Merkmalen anhand der Badestelle Elpersbüttel.....	135
Tabelle 21: Kombination von Aktivitäten bei Urlaubs- und Freizeitreisen ab 1 Übernachtung in Schleswig-Holstein im Jahr 2019.	141

1. Einführung

1.1 Projekthintergrund

Der Tourismus ist ein bedeutender Wirtschaftsfaktor in Schleswig-Holstein. Er trägt mit Bruttowertschöpfungseffekten in Höhe von rund 7 Mrd. Euro zu knapp 8% der gesamten Wirtschaftsleistung Schleswig-Holsteins bei und sorgt für rund 12% der Gesamtbeschäftigung (Krüger et al., 2021). Trotz dieser positiven wirtschaftlichen Wirkungen kann der Tourismus auch negative Folgen mit sich bringen und für Umwelt und Einwohnende ein Belastungsfaktor sein. Auf der einen Seite wird für die Schleswig-Holsteiner*innen Einkommen generiert und Infrastruktur gesichert, auf der anderen werden punktuell zu bestimmten Zeiträumen touristische Tragfähigkeitsgrenzen erreicht und bisweilen überschritten. Dies passiert in sensiblen Naturlandschaften, beispielsweise im Nationalpark Wattenmeer, aber auch in nachfragestarken Tourismusorten. Auch wenn nicht von einem „Overtourism-Problem“ in Schleswig-Holstein gesprochen werden kann, zeigen Studien zur Tourismusakzeptanz, dass insbesondere der Verkehr und Verstopfung (*Congestion*) von Einfahrtsstraßen etc. ein Problem darstellen (Seeler, 2023). Aber nicht nur die Einheimischen sind von negativen Begleiterscheinungen des Tourismus betroffen. Auch die Gäste selbst können durch eine verminderten Erlebnisqualität beeinträchtigt werden. So zeigen beispielsweise Studien an Nord- und Ostseeküste, dass aus Sicht der Gäste in manchen Orten Überfüllungen (*Crowding*) wahrgenommen werden (Schmücker, Reif & Bergler, 2023).

Der Umgang mit den hier grob skizzierten Überlastungen erfordert ein strategisches Management im Tourismus. Dabei lassen sich zwei grundsätzliche Ansätze unterscheiden: die Reduzierung und Begrenzung des Tourismus in Schleswig-Holstein im Sinne einer bewussten Abkehr von einem Wirtschaftswachstum (*Degrowth*) oder ein verbessertes Management des Tourismus im Land. Eine praktikable Strategie besteht darin, Gästen und Einheimischen umfassende Orientierungshilfen an die Hand zu geben, Informationsasymmetrien abzubauen und Transaktionskosten zu reduzieren mit dem Ziel, Besucherströme im Land effektiv zu lenken. Besucherdruck an touristischen Hotspots kann beispielsweise durch die bewusste Empfehlung von Coldspots, also touristischen Attraktionspunkten, die weniger häufig besucht werden, reduziert werden. Derartige Lenkungsmaßnahmen können wesentlich dazu beitragen, die Akzeptanz des Tourismus zu erhalten, negative Auswirkungen auf die natürliche Umwelt zu minimieren und gleichzeitig wirtschaftliche Interessen zu berücksichtigen. Ein strategisch ausgerichtetes Besuchermanagement ist demnach ein Teil einer nachhaltigen Destinationsentwicklung und maßgeblicher Baustein, um den Tourismus in Schleswig-Holstein zukunftsfähig und smart aufzustellen.

Gleichwohl sieht sich eine erfolgreiche Umsetzung von Managementmaßnahmen auf der Ebene von Flächendestinationen wie Schleswig-Holstein mit verschiedenen Herausforderungen konfrontiert. Dabei sind wiederum für das Management von Besuchern zwei Punkte von hoher Relevanz: (1) die Verfügbarkeit und der Zugang zu raum-zeitlich fein aufgeschlüsselten Daten über Besucherfrequenzen und -ströme sowie (2) die Bereitstellung von inspirierenden und informativen Angeboten bzw. Lenkungsmöglichkeiten für (potenzielle) Besucher*innen über geeignete Kanäle. Für die erste Herausforderung gab es bislang keine einheitliche Datenquelle, die aussagekräftige Daten für bestimmte touristische Hotspots oder gar flächendeckend für Schleswig-Holstein liefern konnte. Derzeit sind also aktuelle und zukünftige Auslastungen bzw. Frequenzen von Besucher*innen in Schleswig-Holstein unbekannt (Wann kommen wie viele Gäste von wo an den Strand von St. Peter-Ording, wie lange bleiben sie und wo fahren sie danach hin?) oder basieren auf tradiertem Wissen der Destinationsmanager*innen. Bei der Ausspielung erschweren die vielfältigen Möglichkeiten der Informationsweitergabe (beispielsweise über App, Website, PWA, Stelen) ein operatives Management. Wenn also Daten über Besucherfrequenzen und -ströme vorliegen, wie werden potenzielle Gäste, aktuelle Gäste und Einheimische mit diesen Informationen versorgt?

Digitale Empfehlungssysteme können im Bereich des digitalen Besuchermanagements eine Schlüsselrolle spielen, indem sie sowohl Vorhersagen über zukünftige Belegungen bestimmter Orte liefern als auch Alternativen zu stark frequentierten Zielen vorschlagen. Trotz der Entwicklung von theoretischen Modellen für solche Systeme fehlen derzeit allerdings praktikable Lösungen (Schmücker, Reif, Horster et al., 2023). So erfüllen beispielsweise bestehende Systeme wie Smartphone-Apps oder Routenplaner nicht die Anforderungen an ein digitales Besuchermanagement. Gleichwohl finden sich in der Literatur empirische Belege dafür, dass Tourist*innen erfolgreich zu weniger besuchten, von einer DMO empfohlenen Orten gelenkt werden können, ohne dass dies die Qualität ihrer Erfahrung mindert (Mitas et al., 2023).

Digitale Besuchermanagementsysteme, wie hier skizziert, sind derzeit noch in der Entwicklungsphase. Sowohl in der Praxis als auch in der akademischen Diskussion sind keine integrierten Gesamtsysteme bekannt, die es erlauben, Besucher*innen effektiv zu managen. Obwohl die Notwendigkeit der raumzeitlichen Entzerrung von Gästen seit Jahren diskutiert wird, fehlte bisher der Anstoß zur Umsetzung. Die Auswirkungen der Corona-Pandemie haben diesen Impuls jedoch gegeben und führten zu einer „zweifachen digitalen Transformation“ (Schmücker & Reif, 2021). Einerseits hat die transformative Wirkung der Pandemie die Produktion von Daten zu Besucherfrequenzen verstärkt, andererseits bietet die Nutzung dieser Daten die Möglichkeit einer nachhaltigen Transformation des Destinationsmanagements, die über die Zeit der Pandemie hinaus Bestand hat.

Die derzeit in Schleswig-Holstein aktiven Systeme zur Frequenzmessung und -auspielung (bspw. an der Lübecker Bucht, in St. Peter-Ording, in Büsum) sind daher eine gute Grundlage, aber es bedarf noch umfangreicher Forschungs- und Entwicklungsarbeiten, um die vorhandenen Ansätze auszubauen, damit diese Systeme nicht nur lokal wirksam sind. Integrierte Lösungen sind noch nicht vorhanden, haben aber ein großes Potenzial sowohl in wirtschaftlicher Hinsicht als auch in Bezug auf die Nachhaltigkeit (Vermeidung ökologischer, sozialer und perzeptueller Überlastungen).

An dieser Stelle setzt das Verbundprojekt aus Wissenschaft und Praxis an, ein derartiges System zu entwickeln und zu konzeptionieren. Die Verknüpfung der wissenschaftlichen Partner*innen aus der Domäne Tourismus (Deutsches Institut für Tourismusforschung, Fachhochschule Westküste) und der Domäne Data Science (Fachhochschule Kiel), der Integration der Partner*innen aus der IT-Wirtschaft in Schleswig-Holstein (ADDIX GmbH und Lufthansa Industry Solutions AS GmbH) und unter Einbezug der Destinationspraxispartner*innen bei der konkreten Umsetzung an Fallbeispielen ist eine einzigartige Chance, derartige Systeme zu erproben.

1.2 Zielsetzung des Projektes

Ziel des Projektes ist ein Konzept für ein landesweites effektives digitales Besuchermanagementsystem zu konzeptionieren und soweit möglich zu implementieren. Damit wird ein Beitrag für eine smarte Tourismusedwicklung in Schleswig-Holstein geliefert, bei der über unterschiedliche Quellen Daten zu Besucherströmen gesammelt, aufbereitet und mit dem Ziel der Steigerung von wirtschaftlicher Effizienz, ökologischer Nachhaltigkeit und Erlebnisqualität in Form eines Recommenders ausgespielt werden. Die Folge kann eine nachhaltigere und ausgewogenere Tourismusedwicklung in Schleswig-Holstein sein, da temporäre und regionale Überlastungen vermieden werden können.

Das Projekt ist bewusst konzeptionell angelegt, da eine Implementierung eines Recommender, der in der Lage ist Menschen effektiv zu lenken, in der kurzen Projektlaufzeit nicht umsetzbar ist. Daher werden aus den verschiedenen Arbeitspaketen des Projektes insbesondere Empfehlungen abgeleitet, soweit möglich aber im Rahmen eines „Proof- of Concept“ (PoC) bereits implementiert. Eine detaillierte

Urlaubsplanung mit speziell auf die Bedürfnisse der Nutzenden abgestimmten Empfehlungen ist ebenfalls nicht Ziel des hier vorgestellten Recommenders. Eine solche Plattform könnte zwar auf den Projektergebnissen aufbauen, müsste diese aber um einige Funktionalitäten erweitern, deren Ausarbeitung und Umsetzung für die gesamte schleswig-holsteinische Tourismuslandschaft sehr umfassend entwickelt und getestet werden müsste. Fokus und Anspruch dieses Projektes ist es, die Grundlagen für einen Recommender zu legen, der dann weiterentwickelt werden kann.

Eine enge Abstimmung erfolgte mit dem aus Mitteln des Programms „KI-Leuchttürme“ des Bundesministeriums für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (BMUV) teilfinanzierten Projekt AI-Recommender für nachhaltigen Tourismus (AIR), bei dem das DI Tourismusforschung und die Fachhochschule Kiel ebenfalls Verbundpartner sind. Grundgedanke ist, dass die hier im Rahmen des Projektes entwickelten Ideen, die installierte Sensorik, die bezogenen Daten, die Datenarchitektur als auch der Recommender in AIR überführt sowie weitergenutzt und -entwickelt werden.

1.3 Im Projekt betrachtete Use Cases

Um einen möglichst unmittelbaren Anwendungsnutzen in Schleswig-Holstein zu erzielen, wird das Projekt an konkreten Fallbeispielen (Use Cases) erprobt. Dafür werden in den Beispielregionen Sensorik installiert, um Frequenzen zu messen, und der Recommender auf die spezifischen Belange dieser Use Cases ausgerichtet. Hierfür wurden die Destinationsmanagementorganisationen (DMO) in Schleswig-Holstein dazu aufgerufen, Use Case-Vorschläge für das Projekt einzureichen. Konkret sollten Problemstellungen erläutert werden, für die ein digitales Besuchermanagement auf Basis von Sensordaten als potenzieller Lösungsansatz in Frage kommt. Anschließend wurden die Use Cases anhand eines Relevanzbewertungsverfahrens gerankt. Aufgrund der im Projekt zur Verfügung stehenden finanziellen Mittel für Sensorik konnten vier Use Cases umgesetzt werden.

Beim ersten Use Case handelt es sich um **Parkplätze an der Ostsee**, explizit auf Fehmarn und in Grömitz. Die deutsche Ostseeinsel Fehmarn ist mit ihren vielen Stränden und der langen Küstenlinie ein beliebter touristischer Anlaufpunkt und vor allem bei Surfern und Kitzern sehr beliebt. Auch die in der Lübecker Bucht gelegene Gemeinde Grömitz ist ein beliebtes Ziel bei Tourist*innen. Die Parkplätze auf Fehmarn und in Grömitz überschreiten vor allem saisonbedingt immer wieder ihre Kapazitätsgrenzen. Besonders in den Ferienzeiten und bei guten Witterungsverhältnissen kommt es regelmäßig zu verstopften Zufahrten und einem erhöhten Parkplatzsuchverkehr. Das hat wiederum zur Folge, dass auch die Ortschaften durch den Verkehr stark belastet werden, da viele Gäste lange nach einem Parkplatz suchen, der am besten noch strandnah und zentral ist. Als Lösungsansatz seitens der Destinationen wurde vorgeschlagen, die Parkplätze zu digitalisieren, um die Besucher*innen zu Stoßzeiten zu lenken. Ein Besuchermanagementsystem soll frühzeitig auf überfüllte Parkplätze hinweisen.

Bei einem weiteren Use Case handelt es sich um **Wohnmobilstellplätze** in der Holsteinischen Schweiz. Beliebt bei Urlauber*innen sind die zahlreichen Seen, aber auch die vielfältigen Rad- und Wanderwege in der Destination. Auf der Suche nach geeigneten Stellplätzen fahren Wohnmobilstellplätze oft verschiedene Wohnmobilstellplätze vor Ort an. Dies führt in der Regel ebenfalls zu einem stark erhöhten Verkehrsaufkommen. Die Auslastungen von freien Wohnmobilstellplätzen sind temporär, besonders zu Saisonzeiten, nur sehr schwer kalkulierbar. Wunsch der DMO war es, durch eine digitale Besucherlenkung die Aufmerksamkeit auf ländlich geprägte Stellplätze zu lenken und Alternativen über freie Verfügbarkeiten aufzuzeigen. Die freien Kontingente sollten beispielsweise für Tourist Informationen, digitale Infostellen, oder auch direkt über das Smartphone der Besucher*innen einsehbar sein.

Den dritten Use Case bilden **Radrouten an der Schlei**. Die Schlei ist ein Meeresarm der Ostsee, der die Landschaften Angeln und Schwansen trennt. Entlang der Strecke dieses Ostseefjords gibt es verschiedene beliebte Themenwege für Radfahrer*innen und die Region ist vor allem als Naherholungsgebiet beliebt. Die Themenwege an der Schlei erfreuen sich unterschiedlicher Beliebtheit bei den Besucher*innen. Die Strecken entlang des Wassers sind besonders beliebt, sodass es oft zu Überfüllungen kommt. An vielen Stellen ist es aufgrund der schmalen Radwege nicht möglich, mit dem Fahrrad nebeneinander zu fahren und entgegenkommender Fahrradverkehr kann nicht passieren. Das führt unter anderem zu Gästebeschwerden bezüglich der Strecke. Die ländlicheren Radrouten sind nicht so stark frequentiert, sodass hier noch freie Kapazitäten für Radfahrer*innen bestehen würden. Mit Hilfe eines digitalen Besuchermanagement soll der Radverkehr besser in der Region verteilt werden. Die nachfragestarken Routen sollen entlastet und die schwachen Routen gefördert werden. Auf Basis einer Messung können tagesaktuelle Vorschläge für weniger frequentierte Wege gemacht werden.

Der vierte Use Case „**Hotspots Schlei**“ befindet sich ebenfalls in der Region Schlei. Einige Orte in der Region werden von den Gästen besonders stark frequentiert, jedoch bestand hier aufgrund des ländlich geprägten Raums in den letzten Jahren keine Problematik. Ein deutlicher Anstieg der Übernachtungskapazitäten an der Schlei führt nun jedoch zu einer deutlich gestiegenen Nachfrage, die mittlerweile in einigen Orten als Belastung wahrgenommen wird. Dies betrifft insbesondere Orte wie Kappeln und Arnis. Der potenzielle Lösungsansatz der DMO sieht vor, für beide Orte Echtzeitdaten zu erheben, um einen Überblick über das aktuelle Ausmaß an touristischer Nachfrage zu erhalten. Wenn die Belastung durch den Gästeverkehr tagesaktuell zu groß wird, soll dies den Gästen kommuniziert und automatisch alternative Ausflugsziele vorgeschlagen werden. Dadurch wird sich eine Lenkungswirkung und somit gleichmäßigere Verteilung der Besucherströme in der Region erhofft. Als Indikator sollen die zentralen Parkplätze in Kappeln und Arnis dienen.

1.4 Berichtsaufbau

Dem hier in Abschnitt 1 dargelegten Rahmen des Projektes folgt in Abschnitt 2, "Digitales Besuchermanagement im Tourismus", ein theoretischer und konzeptioneller Überblick. Abschnitt 2 fungiert damit als Einführung in den Ansatz eines digitalen Besuchermanagements. Dabei werden auch die zentralen Aspekte jedes Teilbereichs des Projektes kurz umrissen, um die Kontextualisierung der nachfolgenden Arbeitsergebnisse zu erleichtern. Im darauffolgenden Abschnitt 3, "Datengenerierung", liegt der Schwerpunkt auf den innerhalb des Projektes genutzten Datenquellen. Sowohl die Daten aus lokal installierter Sensorik als auch Daten aus globaler Sensorik werden deskriptiv analysiert und interpretiert. Abschnitt 4 befasst sich mit den Anforderungen an die Dateninfrastruktur eines digitalen Besuchermanagementsystems und legt besonderen Wert auf Schnittstellen und Data Hubs, Datenmodelle und die Integrationsplattform. Der folgende Abschnitt 5, "Recommender", konzentriert sich auf die Prognose und Generierung von Alternativen. In diesem Zusammenhang werden sowohl theoretische Voraussetzungen als auch die praktische Umsetzung im Projekt ausführlich erörtert. Abschnitt 6 behandelt das Deployment, also die unterschiedlichen Möglichkeiten der Datenauspielung sowie deren Umsetzung im Projekt selbst. Besondere Aufmerksamkeit liegt hier auf der Bewertung der Relevanz verschiedener Auspielkanäle in Schleswig-Holstein sowie der exemplarischen Umsetzung des Recommenders in einer Frontend-Anwendung. In Abschnitt 7 werden die Ergebnisse zusammengefasst, die Implikationen für zukünftige Forschungen und die Praxis in Reisezielen aufgezeigt und ein Ausblick geboten.

2. Grundlagen eines digitalen Besuchermanagements im Tourismus

Digitales Besuchermanagement im Tourismus beschreibt einen Ansatz, durch ein besseres Management mit Überlastungserscheinungen umzugehen. Konkret bezieht es sich auf den Einsatz digitaler Methoden zur Steuerung von Reise- und Ausflugsentscheidungen, um Überfüllung an Reisezielen zu verhindern. Dies schließt die Datengenerierung und -speicherung, die Entwicklung von Alternativen und die Präsentation der Resultate über digitale Kontaktpunkte mit ein (Schmücker, Keller et al. (2023)). Ein wichtiger Aspekt dabei ist es, die Besucher*innen durch (digitale) Tools so zu beeinflussen, dass Einwohnende sowie Umwelt und Klima möglichst wenig gestört und belastet werden. Dies wird zunehmend zu einer wichtigen Aufgabe von Destinationsmanagementorganisationen (DMO). Digitales Besuchermanagement kann so einen Beitrag zu einer nachhaltigen Tourismusentwicklung leisten (vgl. Abb. 1).

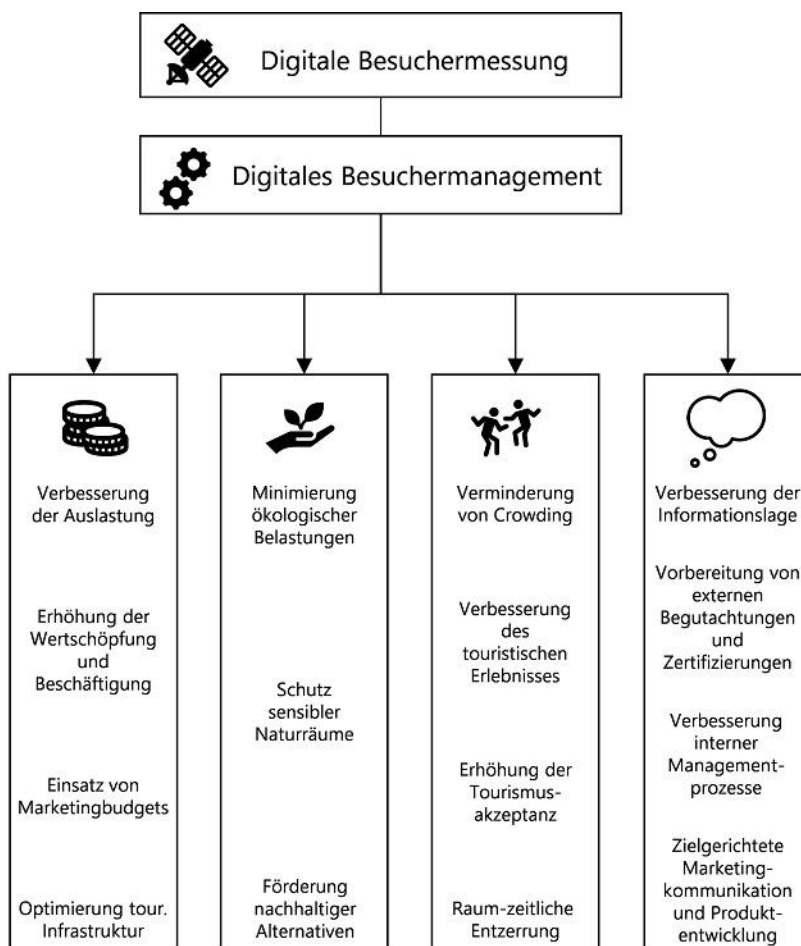


Abbildung 1: Digitale Besuchermessung und -management als Basis einer smarten und nachhaltigen Destinationsentwicklung

Quelle: Schmücker & Reif, 2022, S. 11

Nicht alle Besucher*innen lassen sich aktiv lenken, aber dennoch kann ein Besuchermanagement für Destinationen viele Chancen bieten, wie Entzerrung der Besucherströme, bessere Verteilung der Auslastung von Points of Interest (POI) und Attraktivierung von weniger bekannten POI. Als Grundlage sind

unterschiedliche Daten zur Besuchermessung notwendig. Durch die stetige Messung von Besucherfrequenzen, -strömen und Aktionsräumen in Echtzeit und die Rückschau auf historische Daten lassen sich neben der Live-Auslastung ebenso Modellierungen und Vorhersagen für die Zukunft entwickeln. Die Überlastung von POI soll vermieden und geeignete Alternativen können aufgezeigt werden – nicht nur während des Aufenthaltes, sondern bereits bei der Reiseplanung. Ausgangspunkt eines digitalen Besuchermanagements ist stets eine Analyse der Überlastungsproblematik mit einer damit verbundenen Bedarfsermittlung sowie der Festlegung der Zielsetzung. Es ist aber auch möglich, das digitale Besuchermanagement als Präventivmaßnahme zu implementieren, damit Überlastungen gar nicht erst auftreten.

Ein zeitgemäßes effektives Besuchermanagementsystem sollte für den Tourismusstandort Schleswig-Holstein auf folgenden Grundlagen basieren:

- Generierung von Frequenzen und Besucherströmen auf Grundlage verschiedenster Sensorik
- Digital-dynamisches System, das unterwegs und zu Hause genutzt werden kann
- Nutzung von öffentlichen und freien Datahubs zur Integration verschiedener Sensoren und Datenquellen
- Verwendung von Datenmodellen und KI-Algorithmen für verschiedene Anwendungsfälle zur Analyse der erfassten Daten und zur Erstellung von Prognosen als Basis für die Besucherlenkung
- Diskriminierungsfreie (möglichst lizenzoffene) Bereitstellung der Daten und Prognoseergebnisse für eine möglichst große Weiternutzung auf Anbieterseite und damit Erzielung einer hohen Reichweite bei Einwohnende sowie Besucher*innen
- Nutzung von Nudging-Mechanismen mit Bereitstellung von Alternativen statt reiner Information oder Restriktionen durch Verbote und Sperrungen
- Aktive Integration von Ausspielungskanälen zur Erzielung möglichst hoher Reichweite (Deployment)

Daraus ergibt sich der für das Projekt maßgebliche Prozess des digitalen Besuchermanagements, der auch im Projekt LAB-TOUR SH verfolgt wurde (vgl. Abb. 2).

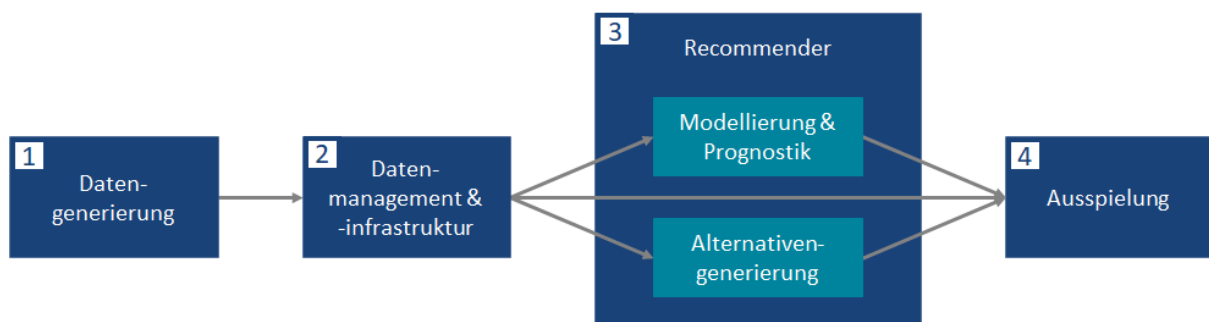


Abbildung 2: Prozess des digitalen Besuchermanagements

Quelle: Eigene Darstellung auf Basis von Schmücker, Keller et al. (2023)

2.1 Datengenerierung

Bei der Datengenerierung geht es um die Messung von Besuchern und damit um die notwendige Datengrundlage für ein digitales Besuchermanagementsystem. Es werden Daten von verschiedenen Quellen erhoben bzw. beschafft. Dabei kann beispielsweise zwischen lokal verbauter (z. B. Bodensensoren) und globaler Sensorik (z. B. Mobilfunkdaten), die flächendeckende Messwerte bereitstellen

kann, unterschieden werden. Zu Beginn gilt es, die Standorte zu bestimmen, für die Daten über Besucher*innen gesammelt werden sollen. Für diese Standorte muss recherchiert werden, ob bereits bestehende Datenquellen verfügbar sind. Das können zum Beispiel Daten aus Schrankensystemen an Parkplätzen oder auch Verkehrsdaten von Dauerkontrollstellen sein. Die Identifizierung der bereits bestehenden Datenquellen bildet die Basis für alle folgenden Schritte. Anschließend muss überprüft werden, ob die bestehenden Datenquellen eine ausreichende Datengrundlage bieten oder weitere Datenquellen, wie beispielsweise lokale und globale Sensorik, genutzt werden müssen. Die Daten aus sowohl bestehenden als auch zusätzlichen Quellen sollen in Kombination einen möglichst umfassenden Blick auf die Besucherfrequenzen, -ströme und Aktionsräume geben.

Nach der Identifizierung der Standorte und der bereits bestehenden Datenquellen werden die Standorte genauer betrachtet. Zum einen muss untersucht werden, welche (technischen) Standortbesonderheiten gegeben sind oder je Standort geschaffen werden muss. Hierunter fallen beispielsweise Lichtverhältnisse, Stromversorgung, Mobilfunk- oder Breitbandversorgung, aber auch Aspekte des Datenschutzes. Zum anderen muss festgelegt werden, was an jedem Standort gemessen werden soll (Personen, Verkehr, Auslastung, Richtungserkennung etc.). Diese Bewertung führt zu der finalen Festlegung von Standorten und benötigten (ggf. neuen) Datenquellen. Sind beispielsweise für die lokale Sensorik keine Sensoren vorhanden, muss evaluiert werden, welche Art der Sensorik sich am jeweiligen Standort anbietet, um diese anschließend zu installieren. Nach erfolgreicher Installation kann mit der Datengenerierung begonnen werden.

Die generierten Daten werden so gesammelt, gespeichert und aufbereitet, dass deren Verfügbarkeit dauerhaft sichergestellt ist. Auch bildet die Datenaufbereitung eine wichtige Grundlage für die weitere Analysen. Die Daten müssen so aufbereitet werden, dass sie sowohl dem jeweiligen Projekt(-ziel) als auch den Datenanforderungen entsprechen. Bei letzterem geht es beispielsweise um die Anforderungen an die Aggregation und den Umfang der Daten sowie die eventuelle Nutzung der Rohdaten. Zudem ist zu beachten, dass die mittels Sensorik erhobenen Daten überprüft werden. Dies kann beispielsweise hinsichtlich der Zuverlässigkeit der Sensoren im Rahmen eines Realitätschecks durch manuelle Zählungen oder durch eine Prüfung der Gültigkeit anhand vorliegender Daten, zum Beispiel aus Schrankensystemen etc., erfolgen. Falls zu einem späteren Zeitpunkt weitere Daten für wichtig empfunden oder neu generiert werden, können diese zusätzlich eingebunden werden – sofern die Schnittstellen und Datenstandards (siehe hierzu Abschnitt 4 zum Datenmanagement) dies zulassen.

2.2 Datenmanagement und -infrastruktur

Für den nachhaltigen Erfolg eines digitalen Besuchermanagementsystems und die langfristige Nutzbarkeit der Daten ist es wichtig, eine offene und von der anbietenden Instanz unabhängige digitale Dateninfrastruktur zu schaffen und die Daten zielgerichtet zu managen. Die Anbindung verschiedener Data Hubs sollte dabei möglich sein, damit weitere Kontextdaten (Standorte weiterer wichtiger POI, Öffnungszeiten von Ausflugszielen usw.) integriert und miteinander in Beziehung gesetzt werden können. Um die Integration der Daten aus den verschiedenen Quellen sicherzustellen, sind die Konzeption einer oder mehrerer standardisierter Schnittstellen erforderlich. Diese stellen eine kontinuierliche Kommunikation zwischen den verschiedenen Data Hubs sicher, damit die Daten dauerhaft ausgetauscht und aktuell gehalten werden.

Ebenfalls spielen etablierte Datenmodelle eine wichtige Rolle. Diese sind die Voraussetzung für die Interoperabilität zwischen verschiedenen, auch externen, Data Hubs und garantieren darüber hinaus im Idealfall, dass die Daten maschinenlesbar und damit „KI-ready“ sind. Es sollte auf international verbreitete Standards (beispielsweise schema.org oder smartdatamodels.org) zurückgegriffen werden.

Mit Hilfe der Datenstandards werden unter anderem Objekte (z. B. Hotels), deren Attribute (z. B. Anzahl der Zimmer) und Wertebereiche (z. B. belegt oder verfügbar) sowie verschiedene weitere Datentypen beschrieben und entsprechend weitergegeben. Mitunter bestehen solche Datenmodelle noch nicht in einer solchen Detailtiefe, dass damit jeder spezifische Anwendungsfall beschrieben werden kann. In einem solchen Fall sollten aus den bestehenden Ontologien heraus neue Beschreibungen für spezifische Fälle (z. B. Beschreibung einer spezifischen Auslastungsgrenze eines Strandes, die variabel und abhängig vom Wetter ist und durch ein Modell jeweils neu berechnet wird) entstehen. So sind sie für den jeweiligen Anwendungsfall sehr individuell einsetzbar und dennoch kompatibel mit etablierten Standards und nachvollziehbar für andere. Ebenfalls wichtig für den nachhaltigen Erfolg ist die dauerhafte Datenhaltung bzw. -aufbewahrung der relevanten Daten, welche frühzeitig sichergestellt werden sollte.

Sind alle Voraussetzungen erfüllt, kann die gesamte Dateninfrastruktur aufgesetzt werden. Mit Hilfe dieser Dateninfrastruktur können in der Folge alle Daten aus verschiedenen Quellen und Hubs zusammengeführt und nutzbar gemacht werden. Die Grundlage für die Anbindung bilden die vorab beschriebenen Schritte zur Konzeption der Schnittstelle(n) und der Datenstandards. Auch ist die Festlegung der notwendigen Datenströme ein wichtiger Faktor. Durch die Schnittstelle(n) werden sowohl der Dateneingang in als auch der Datenausgang aus der Dateninfrastruktur ermöglicht. So können externe Daten (z. B. aus lokaler Sensorik) einzeln eingespeist und im Anschluss gebündelt wieder weitergegeben werden. Gleichzeitig können Analyseergebnisse in die Dateninfrastruktur zurückgespielt werden.

2.3 Recommender

Das Empfehlungsmodul des Besuchermanagements basiert auf zwei Säulen. Zum einen müssen Modelle und Prognosen zur Auslastung von verschiedenen POI entwickelt und erprobt werden und zum anderen sollen Alternativen für zu besuchende Orte ausgespielt werden. Beides wird später auf den Ausspielkanälen dargestellt. Eine wesentliche Rolle spielen hierfür KI-gestützte Verfahren, die Auslastungen in der Zukunft prognostizieren und Alternativen vorschlagen, wenn bestimmte Orte voraussichtlich überlastet sind. Die Datengrundlage für ein solches System wird in den vorab beschriebenen Schritten gelegt. Nun müssen die notwendige Datenbasis und weitere Metadaten für die Modellierung definiert werden. Letztere werden zur Beschreibung verschiedener touristischer Objekte unter Verwendung von (internationalen) Standards genutzt. Dies bezieht sich sowohl auf die erhobenen Frequenzdaten als auch auf alle externen Daten, wie zum Beispiel das Wetter, Veranstaltungen oder auch Ferienzeiten, die Einfluss auf die Prognosen haben können.

Wichtige Schritte auf dem Weg zur Recommender-Entwicklung sind die Datenbereinigung sowie die Sicherstellung der Datenqualität. Bei der Datenbereinigung werden unter anderem fehlerhafte, ungenaue oder doppelte Daten identifiziert und korrigiert. Liegen die relevanten Daten anschließend in ausreichender Qualität, das heißt nicht nur korrekt, vollständig und auf aktuellem Stand, sondern auch in standardisierter Form vor, wird mit der Entwicklung und dem Testen von Modellen zur Auslastungsprognostik auf Basis von KI-Algorithmen begonnen. Sobald diese fertiggestellt sind, werden Auslastungsprognosen ermittelt und weitergegeben.

Anhand der Auslastungsprognosen für verschiedene POI wird ein Empfehlungsmodul konzipiert, welches verschiedene Alternativen, beispielsweise für einen Parkplatz generiert. Dabei gilt es, die Alternativen auf Basis verschiedener Parameter, beispielsweise die Ähnlichkeiten der Alternativen zum ursprünglichen Wunsch oder die Entfernung, zu berechnen. Anhand der tatsächlichen Auslastung wird anschließend weiter getestet und bei Bedarf angepasst.

2.4 Auspielung

Die Prognosen und die Alternativempfehlungen müssen so aufbereitet werden, dass sie in einer ansprechenden und verständlichen Art und Weise dort abrufbar sind, wo Besucherinnen sich ohnehin digital über ihre Reise informieren. Diese Daten können zudem vor Ort über digitale Schilder, Infostellen usw. verteilt werden. Darüber hinaus ist es sinnvoll, die gebündelten Daten sowie deren Analyseergebnisse für die Destinationsvertreter*innen in Form eines Destination-Dashboards verfügbar zu machen, um auch von dieser Seite bei Bedarf Maßnahmen zur Beeinflussung der Besucherströme ergreifen zu können. Dabei ist es wichtig, Destinationsspezifika und die Anforderungen der Zielgruppen zu berücksichtigen. Beispielsweise sollte geschaut werden, welche Apps oder PWA die priorisierten Zielgruppen bereits nutzen. Für eine auf Aktivtourismus spezialisierte Destination spielt die Auspielung in digitalen Tourenplanern eine größere Rolle als für klassische Städteziele.

Mit Hilfe einer Bedarfsanalyse können die relevanten Auspielkanäle identifiziert werden. Die Auswahl und Priorisierung der möglichen Auspielkanäle wird unter anderem anhand ihrer Relevanz für die Destination und die Besucher*innen vorgenommen. Mögliche Touchpoints können sowohl im Vorfeld als auch vor Ort digital abgerufen werden. Als digitale Touchpoints kommen Touren- oder Routenplaner, Websites und Apps von Destinationen oder auch von Hotels, Ausflugszielen oder Mobilitätsanbietern in Frage. Ziel sollte es sein, eine möglichst hohe Verbreitung der Informationen zu erzielen, damit Gäste diese wichtigen Informationen genau dort und dann erhalten, wenn sie diese benötigen – ohne dabei zwingend eine Anwendung nutzen zu müssen, die ihnen bisher unbekannt ist.

Durch die Entwicklung und Nutzung standardisierter Datenmodelle und Schnittstellen können die Daten und die Prognostik sowie die Generierung von Alternativen in verschiedene Auspielkanäle integriert werden. Nach der Auswahl der Auspielkanäle und der Vorbereitung der Daten wird die Auspielung der Prognosen und die Alternativempfehlungen konzipiert, technisch umgesetzt und pilotiert. Anhand von Tests und Analysen können eventuelle Fehler oder Ungenauigkeiten festgestellt und im weiteren Verlauf die notwendigen Anpassungen vorgenommen werden. Die Prognosen und Alternativvorschläge sollten nun den Besucher*innen angezeigt und eine Lenkung dieser zur Vermeidung von Überlastungen bei POI ermöglichen.

3. Datengenerierung

Die Datengenerierung ist die Grundlage für das digitale Besuchermanagement in Schleswig-Holstein. Um die Datengrundlage zu generieren, wurde im Rahmen von zwei komplexen Vergaben Daten aus lokaler und globaler Sensorik beschafft, die Aussagen über Frequenzen und Besucherströme geben sollen. Das Arbeitspaket der "Messung und Sensorik" hat folgende Ziele:

- Schaffung der Datengrundlage mit Hilfe von lokaler Sensorik
- Anreicherung der Daten mit Hilfe von globaler Sensorik
- Deskriptive Datenanalyse zum besseren Verständnis der Besucherfrequenzen und -ströme
- Datenplausibilisierung mit der touristischen Realität

3.1 Datengrundlagen

3.1.1 Lokale Sensorik

Im Rahmen des Projektes LAB-TOUR SH werden am Beispiel der vier Use Cases mit Hilfe von lokal verbauter Sensorik (ZELT und FLEX-Sensoren der Firma Eco Counter) an 54 Messpunkten bzw. 29 Messstandorten in Schleswig-Holstein Frequenzen von Fahrrädern (ZELT, 10 Sensoren) und Fahrzeugen (FLEX-Sensoren, 44 Sensoren) gemessen.

- 1) Use Case Parkplätze Ostsee: Parkplätze auf Fehmarn und in Grömitz: 24 Messpunkte an 11 Standorten
- 2) Use Case Wohnmobilstellplätze Holsteinische Schweiz: 13 Messpunkte an 5 Standorten
- 3) Use Case Hotspots Schlei: 7 Messpunkte an 3 Standorten
- 4) Use Case Radrouten Schlei: 10 Messpunkte/Messstandorte entlang von Radrouten

Die folgenden Abbildungen zeigen die Standorte, an denen im Rahmen des Projektes Sensoren installiert wurden bzw. mit Stand zum 20.06.2023 installiert sind. Dabei kann es vorkommen, dass an einem Standort mehrere Sensoren installiert wurden, beispielsweise weil ein Parkplatz mehrere Ein- und Ausfahrten hat oder dass eine Einfahrt (oder Ausfahrt) zu breit ist, so dass zwei Sensoren angebracht werden mussten. Eine detaillierte Übersicht der verbauten Sensorik findet sich im Anhang. Die ersten Sensoren wurden im August 2022 auf Fehmarn installiert und dann sukzessive während der Projektlaufzeit an den weiteren Standorten. Im Verlauf der Projektzeit kam es jedoch zu starken Verzögerungen in der Umsetzung der Sensorikinstallation, da das Einholen von Genehmigungen mit einem hohen Abstimmungsbedarf verbunden war und Zuständigkeiten, insbesondere bei den Radrouten, die divergierende Grundstückseigner haben, unklar waren. Da nicht alle ursprünglich anvisierten Standorte umgesetzt werden konnten, wurde in Abstimmung mit den DMOs Alternativen gesucht. Kurz vor Projektende waren jedoch noch nicht alle Sensoren installiert. Die fehlende Sensorik bezieht sich auf Standorte im Use Case Radrouten Schlei (5 zu installierende Sensoren) aber auch auf einen als Alternativstandort vorgeschlagenen Parkplatz in Grömitz (2 zu installierende Sensoren) und eine Nachrüstung in Grömitz (1 zu installierender Sensor).

Die Daten sind auf der Plattform Eco Visio für alle Projekt- und Verbundpartner*innen freigeschaltet und können dort abgerufen und analysiert werden.¹ Gleichzeitig besteht die Möglichkeit, die Daten via API abzurufen (siehe Abschnitt 4). Darüber hinaus werden alle Daten nach Projektende auf der Open Data Plattform des Landes Schleswig-Holstein als öffentlicher Download zur Verfügung gestellt. Die Datenlieferung der lokalen Sensorik erfolgt bis zum 30.06.2024, also über das Projektende hinaus. Die Daten können weiterhin im Rahmen der Forschung und Destinationspraxis genutzt werden – es sei denn, es müssen Batterien bei den Sensoren getauscht werden.

¹ Über den Zugang www.eco-visio.net können auf Anfrage auf der [Projektwebsite](#) die Daten mit Leserechten eingesehen und analysiert werden.



Abbildung 3: Messstandorte auf Fehmarn



Abbildung 4: Messstandorte an der Schlei



Abbildung 5: Messstandorte in Grömitz



Abbildung 6: Messstandorte in der Holsteinischen Schweiz



Abbildung 7: Messstandorte entlang von Radrouten an der Schlei

Mit Hilfe des Bodensensors FLEX Sensor werden Ein- und Ausfahrten von Fahrzeugen auf Parkplätzen gezählt. Der kabellose Fahrzeugsensor wird im Asphalt in der Mitte der Fahrspur eingelassen. Der Sensor hat eine integrierte Batterie und ein integriertes Modem. Demnach ist ein autonomer Betrieb möglich. Mit einer Batterie sind, je nach Übertragungsrate, kontinuierliche Messungen bis zu 10 Jahre möglich. Bei dem Sensor handelt es sich um einen Magnetometer, der die Veränderungen des Erdmagnetfeldes misst, sobald er von einem (eisenhaltigen) Fahrzeug überfahren wird. Dabei wird ein magnetisches Profil des jeweiligen Fahrzeugs erstellt, wodurch der Sensor in der Lage ist, bis zu fünf verschiedene Klassen von Fahrzeugen zu unterscheiden (Pkw, Motorrad, Kleinbusse, Lkw, Sonstiges). Der Sensor gibt in 15-minütiger zeitlicher Auflösung Informationen über die Anzahl der Durchfahrten (Ein- und Ausfahrten) sowie die Art des Fahrzeuges. Laut Hersteller liegt die gesamte Messgenauigkeit bei 99%, bei der Klassenermittlung bei 95%. Weitere Parameter des FLEX Sensors fasst Tabelle 1 zusammen:

Physische Beschaffenheit	<ul style="list-style-type: none"> - Höhe: 115 mm - Durchmesser: 155 mm - Betriebstemperaturen: -35°C bis 70°C
Standort	<ul style="list-style-type: none"> - Möglichst fließender Verkehr - Möglichst auf gerader Strecke - Gesamter Verkehr sollte die Schleifen überfahren - Standort ohne erhebliche elektromagnetische Interferenzen - Asphalt, Pflaster (bis 10cm Dicke), Wassergebundene Decke
Installation	<ul style="list-style-type: none"> - In der Mitte der zu messenden Fahrspur

	<ul style="list-style-type: none"> - Durchmesser Bohrloch: > 170 mm - Tiefe Bohrloch: 180 mm - Dauer: ca. 30-60 min - Bohrloch wird inklusive Zähltechnik versiegelt - Genehmigung notwendig - Im öffentlichen Verkehr wird zusätzlich eine verkehrsrechtliche Anordnung benötigt
Daten	<ul style="list-style-type: none"> - Bis zu 5 Klassifizierungen: PKW, Motorräder, Kleinbusse, LKW, Unbestimmte Nutzungsart - Erfassung von Überfahrten inkl. Richtung - Datenspeicherung in 15-Minuten-Intervallen - Datenübertragung (GMS) in 15-Minuten-Intervallen
Genauigkeit	<ul style="list-style-type: none"> - Klassifizierungsfehler bei einem 5-Klassen-Schema: 5% - Durchschnittlicher Geschwindigkeitsfehler: 1% - Durchschnittlicher Längenfehler: 1%

Tabelle 1: Parameter FLEX Sensor

Bei dem ZELT Evo Sensor wird ebenfalls eine unterirdische Zähltechnik bzw. eine Zähltechnik im Schacht angewandt. Es handelt sich um eine induktionsschleifenbasierte Messung, bei der hauptsächlich die metallischen Teile in der Nähe der Schleifen, also den Rädern, identifiziert werden. Es wird sozusagen der elektro-magnetische „Fingerabdruck“ des Fahrrads erkannt. Insgesamt werden 13 spezifische Parameter analysiert. Der Sensor kalibriert sich fortlaufend selbstständig und verfügt über eine autonome Energieversorgung. Demnach ist ein autonomer Betrieb möglich. Die Batterielaufzeit beträgt, je nach Übertragungsrate, bis zu 24 Monate. Der Sensor gibt in 15-minütiger zeitlicher Auflösung Informationen über die Anzahl der Überfahrten. Die Messgenauigkeit liegt laut Hersteller auf Radwegen bei 95%. Einen Überblick über die Parameter des ZELT Evo Sensors gibt Tabelle 2:

Physische Beschaffenheit	<ul style="list-style-type: none"> - Induktionsschleifen inklusive Schacht
Standort	<ul style="list-style-type: none"> - Möglichst fließender Verkehr - Möglichst auf gerader Strecke - Gesamter Verkehr sollte die Schleifen überfahren - Standort ohne erhebliche elektromagnetische Interferenzen - Asphalt, Pflaster (bis 10cm Dicke), Wassergebundene Decke
Installation	<ul style="list-style-type: none"> - Ca. 4-5cm tiefer und 1cm breiter Nut - Zuleitung zu Schacht im Grün/Pflaster/Wegesrand - Dauer: ca. 4-5 Stunden - i.d.R. keine Sperrung der gesamten Fahrbahn notwendig - Genehmigung notwendig - Im öffentlichen Verkehr wird zusätzlich eine verkehrsrechtliche Anordnung benötigt
Daten	<ul style="list-style-type: none"> - Erfassung von u.a. Radfahrenden & PKW/Bus inkl. Richtung - Datenspeicherung in 15-Minuten-Intervallen - Datenübertragung (GMS) in 15-Minuten-Intervallen
Genauigkeit	<ul style="list-style-type: none"> - 90% auf der Straße - 95% auf Radwegen

Tabelle 2: Parameter ZELT Evo Sensor

3.1.2 Globale Sensorik

Die im Rahmen der lokal verbauten Sensorik gemessenen Daten dienen als Basis für ein Prognosemodul des Recommenders im Projekt. Sie geben also Auskunft über die Frequentierung der von den Praxispartner*innen von hohem Besucheraufkommen identifizierte Standorte. Sie haben jedoch den Nachteil, dass sie keine Auskunft über die Zielgruppe der Nutzenden (Handelt es sich um Einwohnende oder Tourist*innen, die die Orte aufsuchen?), über die Quellmärkte der Zielgruppe oder über den Aufenthaltsort vor und nach dem Besuch des Ortes – so genannte Quell-Zielgebietsbeziehungen – geben (Wo waren die Menschen am Morgen bevor sie den Ort aufsuchten und wo waren sie am Abend?). Diese Auskunft können jedoch Big Data Quellen (globale Sensorik) liefern, die im Rahmen des Projektes in Form von App-basierten Location-Events des GPS-Empfängers im Smartphone beschafft wurden. Ziel des Datenbezuges ist die Erforschung der Anreicherung und Kalibrierung der lokal generierten Sensordaten mit Hilfe von globaler Sensorik (vgl. Abb. 8). Dabei ist die im Projekt verbaute lokale Sensorik als Teil der Kategorie B-1 (Standortmessungen) zu verstehen, die in der Lage sind, Frequenzen an einzelnen Standorten zu zählen (Ziel 1) während die App-Location Events in diesem Falle Multispot-Systeme sind, die durch die Aneinanderreihung von Signalen (Tracking) Aussagen über alle vier Zielbereiche digitaler Besuchermessung geben können. Der Vorteil dieser Tracking-Systeme liegt darin, dass sie in der Lage sind, Nutzende wieder zu identifizieren und so Rückschlüsse über Besucherströme geben zu können. Konkret bedeutet dies, dass für den Fall der Anreicherung für die Orte der lokalen Sensorik Aussagen erhalten werden sollen:

- Touristische Klassifizierung (Einheimische, Tagestourist*innen, Übernachtungstourist*innen und Sonstige),
- Wohnorte bzw. gewöhnliche Aufenthaltsorte (Home-Zone) der Besucher*innen*innen am Mess-Standort
- Aufenthaltsort am Morgen des Tages
- Aufenthaltsort am Abend des Tages
- Alter und
- Geschlecht

Im Falle der Kalibrierung sollen für die Orte ebenfalls Frequenzdaten von Besuchenden gegeben werden. Gleichzeitig können diese Daten zu Volumen im Falle der Kalibrierung auch für die Prognose eines Empfehlungsmoduls genutzt werden. Trotz in der Theorie elaborierter Konzepte und Diskussionen von verschiedenen Datenquellen (Hardy, 2020; Schmücker & Reif, 2022a), gibt es bislang keine in der Literatur und Praxis bekannten Anwendungsfälle, in der Daten aus lokaler und globaler Sensorik gegenübergestellt wurden. Zwar gibt es Arbeiten, die globale Daten, beispielsweise aus Smartphone-Applikationen und Daten aus dem Mobilfunknetz, vergleichen sowie mit Referenzdaten aus Destinationen validieren (Schmücker & Reif, 2022b) und auch anonymisierte App-Locations werden in der Destinationspraxis, beispielsweise auf Hawaii², genutzt. Das Projekt LAB-TOUR SH bringt jedoch erstmals diese beiden Datenquellen (Kombination aus globaler und lokaler Sensorik) zusammen und liefert Ansätze zur Vorgehensweise und erste Ergebnisse.

² <https://www.hawaiitourismauthority.org/research/symphony-dashboards/>

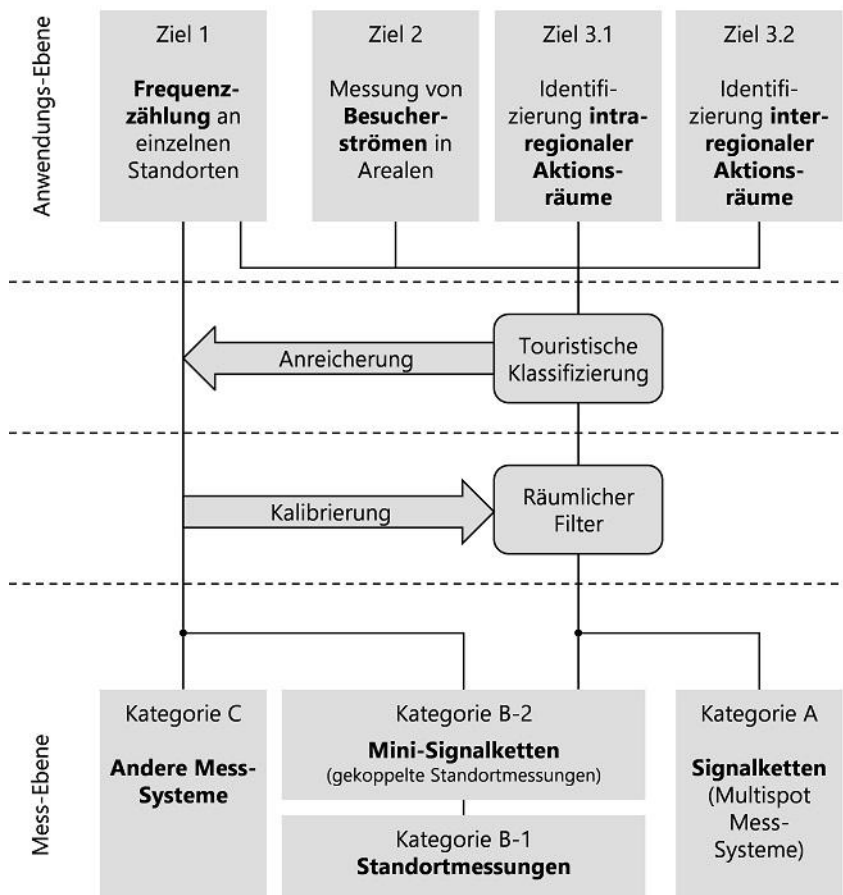


Abbildung 8: Mess-Systeme und deren Anwendung im Tourismus

Quelle: Schmücker & Reif, 2022a, S. 27

Für das Projekt stehen Daten globaler Sensorik vom Unternehmen thaltegos GmbH zur Verfügung, die wiederum auf Daten des Unternehmens Placesense³ zurückgreifen und diese im deutschen Markt vertreiben. Placesense, ein sogenannter Geo Location Broker, nutzt unter Einhaltung des Datenschutzes die Lokalisierungsdaten, die bei der Nutzung verschiedener Smartphone-Applikationen durch den GPS-Empfänger des Smartphones anfallen. Grundlage hierfür ist die Werbe-ID des Smartphones, die es erlaubt Nutzungsdatensätze zu erstellen, die dann beispielsweise im Einzelhandel oder der Immobilienbranche genutzt werden. Aufgrund ihrer granularen räumlichen und zeitlichen Auflösung, lassen sich diese Daten auch für die Tourismusforschung und -praxis nutzen.

Für LAB-TOUR SH stehen für 38 Polygone in Schleswig-Holstein anonymisierte Location Events zur Verfügung (siehe Tab. 3). Die Auswahl und Definition der Polygone erfolgte nach den Use Cases und der zum Zeitpunkt der Datenbeauftragung installierten lokalen Sensorik. Da zum Zeitpunkt der Datenbeauftragung noch nicht alle lokalen Sensoren installiert waren, konnten nicht für alle in LAB-TOUR SH vorgesehenen Standorte globale Daten bezogen werden. Die Auswahl der Polygone erfolgte in diesem Falle auf Basis von vorliegenden Referenzdaten aus anderen Sensoren (Insel Amrum, Perlebucht, Holm in Flensburg, Polygon des Veranstaltungsortes der Kieler Woche, Parkplätze und gesamter Strand der Lübecker Bucht, Strandparkplatz und Seebrücke St. Peter-Ording). Die weiteren Polygone (Büsum Hauptdeich, Büsum Einkaufsstraße und Museumsdeich, Strand Sylt und Promenade St. Peter-Ording)

³ <https://placesense.ai/>

wurden als bekannte Hotspots bewusst an der Westküste ausgewählt. Hier liegen keine Referenzdaten vor.

Use Case	Name des Zielgebiets (Polygons)	ID
Amrum	Insel Amrum	20
Büsum	Perlebucht	21
Büsum	Hauptdeich	22
Büsum	Einkaufsstraße und Museumshafen	23
Flensburg	Einkaufsstraße Holm	24
Hotspots Schlei	Großparkplatz Kappeln	12
Hotspots Schlei	Arnis Parkplatz Neuer Damm	15
Hotspots Schlei	Kappeln Ellenberger Parkplatz	16
Kiel	Kieler Woche	25
Lübecker Bucht	Parkplatz Scharbeutz P3	26
Lübecker Bucht	Parkplatz Scharbeutz P2	27
Lübecker Bucht	Lübecker Bucht Strand gesamt	28
Lübecker Bucht	Parkplatz Scharbeutz Badweg	29
Lübecker Bucht	Parkplatz Scharbeutz P4	30
Parkplätze Ostsee	Parkplatz Burgtiefe	2
Parkplätze Ostsee	Parkplatz Südstrand Mitte	3
Parkplätze Ostsee	Parkplatz Meeschendorf	4
Parkplätze Ostsee	Parkplatz Gold	5
Parkplätze Ostsee	Parkplatz Orth	6
Parkplätze Ostsee	Parkplatz Grüner Brink	7
Parkplätze Ostsee	Parkplatz Königsredder Straße	8
Parkplätze Ostsee	Parkplatz Seebrückenvorplatz Seestr. 31	9
Parkplätze Ostsee	Parkplatz Deichweg	10
Parkplätze Ostsee	Großraumparkplatz Grömitz	11
Radwege Schlei	Standort 23 Reitanlage Dorfstr Radweg	31
Radwege Schlei	Standort 22 Gulde	32
Radwege Schlei	Standort 21 Drecht	33
Radwege Schlei	Standort 17 Radweg 1 Zuwegung Museum Haithabu	34
St. Peter-Ording	Promenade Bad	35
St. Peter-Ording	Seebrücke Bad	36
St. Peter-Ording	Strandparkplatz Ordinger Strand	37

Sylt	Strand Sylt	38
Wohnmobilstellplätze Holsteinische Schweiz	Reisemobilpark Eutin	1
Wohnmobilstellplätze Holsteinische Schweiz	Naturcamping Prinzenholz	13
Wohnmobilstellplätze Holsteinische Schweiz	Malente Camping an der Schwentine	14
Wohnmobilstellplätze Holsteinische Schweiz	Womo Stop Malente	17
Wohnmobilstellplätze Holsteinische Schweiz	Camping Gut Ascheberg Lange Wisch	18
Wohnmobilstellplätze Holsteinische Schweiz	Camping Spitzenort	19

Tabelle 3: Untersuchungsgebiete in Schleswig-Holstein

Durch die Möglichkeit der zeitlichen und räumlichen Auflösung der GPS-Signale von Endgeräten, lassen sich auf Basis von Regelmäßigkeiten des Mobilitätsverhaltens dieser Endgeräte Aussagen über übliche Aufenthaltsorte des Endgerätes am Tag und am Abend treffen. Mit Hilfe dieser Erkenntnis lassen sich mit einfachen Heuristiken Aussagen über "touristische" und "nicht-touristische" Signale treffen. Für diese Art der touristischen Klassifizierung werden folgende Definitionen verwendet:

- **Einheimische:** Personen, die ihren Wohn-Standort *innerhalb* der „Homelocation des Zielgebietes“ haben **und** im Zielgebiet auftauchen.
- **Tagestourist*innen:** Personen, die ihren Wohn-Standort *außerhalb* **und** ihren Arbeits-Standort *außerhalb* der „Homelocation des Zielgebietes“ haben **und** im Zielgebiet auftauchen **und nicht** an zwei aufeinander folgenden Tagen getrackt wurden.
- **Übernachtungstourist*innen:** Personen, die ihren Wohn-Standort *außerhalb* **und** ihren Arbeits-Standort *außerhalb* der „Homelocation des Zielgebietes“ haben **und** im Zielgebiet auftauchen **und** an zwei aufeinander folgenden Tagen getrackt wurden.
- **Sonstige (Einpender):** Personen, die ihren Wohn-Standort *außerhalb* und ihren Arbeits-Standort *innerhalb* der „Homelocation des Zielgebietes“ haben **und** im Zielgebiet auftauchen.

Die Home- und Worklocations der Zielgebiete werden gemäß der Definition in Tabelle 4 festgelegt, wobei eine Use Case spezifische Definition vorgenommen wird. Zum Verständnis ist es wichtig, dass sich die Homezone eines Smartphones anhand des häufigsten Übernachtungsortes innerhalb der letzten 30 Tage berechnet. Demnach kann die Homezone auch wechseln.

Use Case	Home- und Worklocations
Amrum	Gemeinden: Nebel ODER Norddorf ODER Wittdün
Büsum	Gemeinde: Büsum
Flensburg	Gemeinde: Flensburg
Hotspots Schlei	Gemeinden: Kappeln ODER Arnis
Kiel	Gemeinde: Kiel
Lübecker Bucht	Gemeinde: Scharbeutz
Parkplätze Ostsee	Gemeinden: Fehmarn (für Polygone auf Fehmarn) und Grömitz (für Polygone in Grömitz)
Radwege Schlei	Kreis: Schleswig-Flensburg

St. Peter-Ording

Gemeinde: St. Peter-Ording

Sylt

Gemeinden: List auf Sylt ODER Kampen (Sylt), ODER Wenningstedt-Braderup ODER Sylt ODER Hörnum (Sylt)

Wohnmobilstellplätze Holsteinische Schweiz

Gemeinden: Eutin ODER Malente ODER Plön ODER Ascheberg

Tabelle 4: Home- und Worklocations der Polygone

Für das Projekt stehen Daten auf Basis von Kalenderwochen für einen Zweijahreszeitraum, beginnend vom 30.06.2021 – 30.06.2023 zur Verfügung. Darunter der Projektzeitraum LAB-TOUR SH (01.01.2022 – 30.06.2023). Die deskriptive Analyse der Daten in Abschnitt 3.2 bezieht sich jedoch mit wenigen Ausnahmen auf das Kalenderjahr 2022.

Für die Kalibrierung werden die Anzahl der *visitors* und *daily visits* je Kalenderwoche bereitgestellt. Unter *visitors* wird die Anzahl an eindeutigen Besucher*innen (unique devices) innerhalb einer Kalenderwoche verstanden. Falls ein/e Besucher*in täglich im Polygon erfasst wird, wird er dennoch als ein/e Besucher*in erfasst. Diese Besuchenden erzeugen *daily visits*, also eine Anzahl an täglichen Besuchen innerhalb der Kalenderwoche (Anzahl von Besuchstagen in einer Woche). Falls ein/e Besucher*in täglich am Zielgebiet erfasst wird, sammelt die Person 7 *daily visits*. Das fiktive Beispiel in Tabelle 5 zeigt für ein Polygon in Summe 3 *visitors*, die 12 *daily visits* erzeugen.

Bsp.	Mo	Di	Mi	Do	Fr	Sa	So	Summe
Visitor A	x	x	x	x	x	x	x	1
Visitor A daily visits	x	x	x	x	x	x	x	7
Visitor B	x	x	x	x				1
Visitor B daily visits	x	x	x	x				4
Visitor C	x							1
Visitor C daily visits	x							1

Tabelle 5: Beispiel visitors und daily visits im Polygon

Darüber hinaus stehen typische Tages- und Wochenverläufe auf Stundenbasis für die *visitors* und *daily visits* zur Verfügung. Hierfür wurden folgende Saisonzeiten definiert:

- 1 Frühling:** März, April, Mai
- 2 Sommer:** Juni, Juli, August
- 3 Herbst:** September, Oktober, November
- 4 Winter:** Dezember, Januar, Februar

In Abschnitt 3.2 werden die im Rahmen von LAB-TOUR SH erworbenen Daten analysiert. Gleichwohl kann eine umfassende Auswertung aller hier vorliegenden Daten nicht im Rahmen des Berichtes erfolgen. Für die weitere Verwendung und öffentliche Nutzung der Daten für Forschung, Lehre und Praxis

werden daher die Daten online im Excel-Rohformat zur Verfügung gestellt und können so für weitere Projekte genutzt werden.⁴

3.1.3 #sh_wlan

Mit dem #sh_wlan stehen durch die Projektbeteiligung der ADDIX GmbH dem Projekt auch diese Daten zur Verfügung. Beim #sh_wlan⁵ handelt es sich um ein freies WLAN-Angebot in Schleswig-Holstein. Diese Art von Datenquelle kann mit Blick auf Abbildung 8 grundsätzlich der Kategorie B-1, Standortmessungen zugeordnet werden und ist wie die oben beschriebenen lokalen Sensoren Teil von stationärer Sensorik. Allerdings ist es im #sh_wlan in Teilen auch möglich mit Hilfe anonymisierter Daten Mini-Signalketten zu produzieren um Besucherströme nachzuvollziehen. Demnach fällt das #sh_wlan im vorliegenden Fall in die Kategorie B-2 (Mini-Signalketten), die zum Tracking fähig sind. Die ADDIX GmbH betreibt technisch die Access Points (Zugangspunkte), an denen #sh_wlan ausgestrahlt und damit WLAN-fähigen Geräten (Clients) Zugang zum Internet ermöglicht wird. Die Sparkassen in Schleswig-Holstein fördern das #sh_wlan als Digitalisierungsinitiative. Insgesamt steht das #sh_wlan an etwa 800 Standorten und 3.000 Access Points zur Verfügung. Im Jahr 2022 wurde es von etwa 1,3 Millionen Menschen genutzt.

Beim ersten Login stimmen die Benutzer*innen explizit den allgemeinen Geschäftsbedingungen zu, danach findet ein sogenannter automatischer Re-Login statt. Dieser Service sorgt dafür, dass sich die Benutzer*innen nicht fortwährend zwischen verschiedenen Access Points, an denen #sh_wlan ausgestrahlt wird, neu anmelden müssen. Damit das technisch möglich ist, wird von jedem Client beim ersten Login eine eindeutige, ihm zugeordnete Kennung gespeichert. Diese wird von dem jeweiligen Access Points an den WLAN-Controller (WLC) weitergeleitet. Der WLC legt zusätzlich eine eindeutige Kennung des übermittelnden Access Points unter einem Zeitstempel in der Datenbank ab. Jedem Access Point ist intern eine Geokoordinate zugeordnet. Je nach verbautem Access Point ist eine Verbindung ab einem Umkreis von etwa 30-50 Metern im Innenbereich (indoor) und etwa 150 Metern im Außenbereich (outdoor) möglich.

Die Access Points können somit als Sensoren für #sh_wlan-Nutzer*innen betrachtet werden. Durch die eindeutige Kennung sind die Clients unterscheidbar, d. h. auch zählbar. Im Rahmen des LAB-TOUR SH-Projekts wird pro Access Point, an dem #sh_wlan ausgestrahlt wird, alle 15 Minuten die Anzahl der Clients aggregiert und an die Datenplattform übertragen. Die eindeutigen Kennungen der Nutzer*innen werden nicht weitergegeben. Es muss auf jeden Fall ein Access Point vor Ort vorhanden sein, damit Daten gesammelt werden können. Dieser Access Point fungiert als Sensor. Die Nutzung des #sh_wlans als Ausspielmöglichkeit wird in Abschnitt 6.5 erläutert.

3.1.4 Weitere Daten

Im Rahmen des Projektes wurden weitere Datenquellen recherchiert, die Aussagen über Bewegungen von Personen erlauben. Eine wichtige Quelle hierbei waren Daten von Besucherstrom.info (SH Beach) des Verbundpartners LHIND. Besucherstrom.info lässt Datenquellen zu, die über ein REST-Interface Daten übertragen werden können. Datenquellen sind in erster Linie die Datenplattformen der Sensorhersteller (Datenplattform), nicht die Sensoren selbst, die die Daten erheben. Die Daten werden üblicherweise im JSON-Format entgegengenommen und müssen im Wesentlichen Sensor-ID, Zeitstempel

⁴ Die Daten stehen als halböffentlicher Download auf der Projektwebsite des DI Tourismusforschung unter diesem [Link](#) zur Verfügung.

⁵ <https://sh-wlan.de/>

und Zählinformation. Enthalten. Je nach Art der Datenplattform ist es möglich, Zeiträume abzufragen, die entweder alle Zählungen für den abgefragten Zeitraum enthalten oder voraggregierte Zählungen für diesen Zeitraum. Die Datenquellen senden entweder eventbasiert für jede Zählung, zu festgelegten Intervallen, oder werden zu festgelegten Intervallen abgefragt.

Besucherstrom.info ordnet die Zählungen einzelner Sensoren sog. Bereichen (Areas) zu. Ein oder mehrere Sensoren können einen Bereich bilden. Dadurch lassen sich für den Bereich alle ausgestatteten Ein- und Ausgänge synchron erfassen. Werden diese vollständig erfasst, lässt sich für diesen Bereich der aktuelle Füllstand im Zeitverlauf ermitteln. Mehrere Bereiche lassen sich zu einer Bereichsgruppe zusammenfassen. Diese Bereichsgruppe summiert die Zählungen aller ihr zugeordneten Bereiche. Damit lässt sich also eine Hierarchie von Bereichen abbilden.

Je nach Art des Sensors sind unterschiedliche Zählarten möglich:

- **Zähllinien-Messung:** Messung von Übertritten in Ein- und Austrittsrichtung über eine virtuelle Zähllinie, die die Bewegung von Objekten misst
- **Flächenzählung:** Zählung von Objekten zu einem Zeitpunkt auf einer virtuellen Fläche
- **Belegungserkennung für Parkplätze:** Erkennung vom Belegungszustand einzelner Parkbuchten eines Parkplatzes zur Bestimmung des Füllstandes des Parkplatzes
- **Zählung per Schrankenöffnung:** Öffnen und Schließen von Zugängen wie Schranken oder Drehkreuzen werden erfasst. Damit lassen sich sowohl die Bewegungen als auch der Füllstand bestimmen.

Es werden folgende Sensoren verwendet:

- Überkopf-Zählsensoren (Evocount, Xovis) Zähllinien-Messung
- Netzwerkkameras (AXIS): Flächenzählung
- Parkplatzsensoren (Cleverciti): Belegungserkennung
- Parkschraken (Siemens): Zählung per Schrankenöffnung

Folgende Datenquellen je Standort wurden im Rahmen des Projektes genutzt (vgl. Tab. 6):

	Personenzählung	Fahrzeugzählung
St. Peter-Ording	Seebrücke, Köhlbrand (Evocount): Abfrage von Sensordaten jede Minute, Sensor teilt aggregierte Zu- und Abgänge der letzten Minute mit. Aufgrund der Charakteristik einer Passage werden hier nur die Bewegungsdaten betrachtet, eine Aufsummierung zu einem Füllstand ist nicht sinnvoll.	Strandparkplatz Ording, Parkplatz Marktplatz (AXIS): Es werden jede Minute Bilder des Bereichs gemacht, und die befindlichen Fahrzeuge anonym gezählt. Die Zählinformation wird mit Zeitstempel übertragen. Die Bilder werden nach Auswertung sofort gelöscht

	<p>Marktplatz, Kulinarische Meile, Seebrücke Vorplatz (AXIS): Es wird jede Minute Bilder des Bereichs gemacht und die befindlichen Personen anonym gezählt. Die Zählinformation wird mit Zeitstempel übertragen. Die Bilder werden nach Auswertung sofort gelöscht</p>	
Scharbeutz	<p>Strandabschnitte 9-29 (Evocount): Messung erfolgt für Zu- und Eintritt. Minütliche-Datenabfrage, aggregierte Zu- und Abgänge der letzten Minute</p>	<p>Parkplatz Strandallee II, Parkplatz Badeweg (CleverCiti): Es wird die Belegung der Parkbuchten des Parkplatzes zurückgegeben. In Besucherstrom wird die Belegung des Parkplatzes angezeigt, wird alle zwei Minuten abgefragt</p> <p>Parkplatz P2, P3 und P4 (Siemens Parkanlage): Abfrage alle zwei Minuten, Füllständebasierend auf Öffnen und Schließen der Parkschranksen, um den Füllstand zu ermitteln.</p>

Tabelle 6: Im Projekt eingebundene Datenquellen von besucherstrom.info

Die nachfolgende Tabelle 7 zeigt weitere touristische Frequenzdaten, die in Schleswig-Holstein zur Verfügung stehen:

Kurzbezeichnung	Beschreibung	Anzahl Messpunkte	Zeitliche Auflösung	Sensortyp/ -hersteller	Dateneigentümer
KVG Kiel	Positionsdaten der KVG Busse	ca. 160	sekündlich (wenn Bus im Einsatz)	ADDIX WLAN Router	ADDIX GmbH
Laboe Fußgänger	Personenzählung Promenade	1	k.A.	Eco-Counter PIR	Tourismusbetrieb Ostseebad Laboe
Amrum Wifi	Wificounter, beispielsweise am Hafen	6	sekündlich	hubermedia destination.box	Amrum Touristik

Kurzbezeichnung	Beschreibung	Anzahl Messpunkte	Zeitliche Auflösung	Sensortyp/-hersteller	Dateneigentümer
Büsum Wattinsel/ Perlebucht	Lidarsensor	3	k.A.	Evocount Lidar	TMS Büsum
Kiel Parkleitsystem	Livedaten Parkplätze Innenstadt	ca. 14	Live	k.A.	Stadt Kiel
Lübeck Parkleitsystem	Livedaten Parkplätze Innenstadt	18 (von insgesamt 45)	Live	k.A.	KWL Lübeck
Travemünde Parkleitsystem	Livedaten Parkplätze Travemünde	5 (von insgesamt 18)	Live	k.A.	KWL Lübeck
TA.SH Badestellenfinder	Auslastung Badestellen in SH	6	Live	k.A.	TASH Kiel
SPO Seebrücke & Strandzugang Köhlbrand	Personenzählung Bewegungen Strandzugänge	2	Live (minütliche Granularität)	Evocount optisch/Laser	TZ-SPO
SPO Strandparkplatz Ording	Fahrzeugzählung	1	Live (alle 20 Minuten Update kompletter Parkplatz; jede Minute Update ein Frame)	Axis PTZ-Kamera/LHIND Fahrzeugzählung KI	TZ-SPO
SPO Seebrücke Vorplatz, Kulinarische Meile, Marktplatz (Personen)	Personenzählung Füllstand	3	Live (alle 5 Minuten)	Axis PTZ-Kamera/LHIND Fahrzeugzählung KI	TZ-SPO

Kurzbezeichnung	Beschreibung	Anzahl Messpunkte	Zeitliche Auflösung	Sensortyp/-hersteller	Dateneigentümer
SPO Marktplatz (Fahrzeuge)	Personenzählung Füllstand	4	Live (alle 5 Minuten)	Axis PTZ-Kamera/LHIND Fahrzeugzählung KI	TZ-SPO
SPO Wifi-Tracker	Personenzählung (abgeleitet aus Anzahl (Wifi-/Bluetooth-Geräten))	4	Live (alle 15 Minuten)	hubermedia destination.box Wifi-Tracker	TZ-SPO
TALB Strandzugänge	Personenzählung Strandzugänge	20	Live (minütliche Granularität)	Evocount optisch/Laser	TALB
TALB Siemens Großparkanlage	Personenzählung Fahrzeuge	6	Live	Siemens Parkschrankenanlage (tlw. Schranke, Bodenkontakt-schleife, Kamera)	Gemeinde Scharbeutz
TALB Parkplatz Badeweg/Parkbuchten Abschnitt Strandallee II	Personenzählung Fahrzeuge	2	Live (minütliche Granularität)	cleverciti optisch basierter Sensor	TALB
Haffkrug WiFi-/Bluetooth Tracker	Personenzählung (abgeleitet aus Anzahl (Wifi-/Bluetooth-Geräten))	11	Live (minütliche Granularität)	SonoBeacon Wifi-/Bluetooth Tracker	TALB/Gemeinde Scharbeutz
Hystreet Zählstellen in Flensburg, Schleswig, Kiel und Lübeck	Personenzählung	6	Stündlich	LasePeco	Hystreet

Kurzbezeichnung	Beschreibung	Anzahl Messpunkte	Zeitliche Auflösung	Sensortyp/-hersteller	Dateneigentümer
Ostseecard	Personenzählung	22	k.A.	k.A.	Ostsee-Tourismus-Service GmbH
Lübecker Bucht Strandticker	Personenzählung	7	Live (alle 10 Minuten)	k.A.	TALB
Lübecker Bucht Strandkorbvermietung	Personenzählung	k.A.	viermal täglich	manuell	TALB

Tabelle 7: Touristische Frequenzdaten in Schleswig-Holstein. Eigene Recherche. Keine Gewähr

3.2 Deskriptive Analyse lokaler und globaler Sensorik

3.2.1 Lokale Sensorik

Nachfolgend werden die Daten aus der lokalen Sensorik seit Messbeginn beschrieben. Auf der Insel Fehmarn wurden die Sensoren am 06.08.2022 installiert und liefern seit diesem Tag fortlaufend Daten zu den Ein- und Ausfahrten auf den Parkplätzen. Die zeitliche Auflösung der Daten liegt für alle Standorte in LAB-TOUR SH im 15 Minuten Takt vor. Um die Daten für den Bericht nun nachfolgend zu visualisieren, wurde sich für eine tageweise Aggregation der Einfahrten entschieden. Nachfolgend wird demnach für jeden Tag seit Installationsbeginn die Summe der Einfahrten dargestellt. Dabei ist zu betonen, dass es sich hierbei nicht um die Berechnung eines Füllstandes für den entsprechenden Parkplatz handelt, sondern lediglich die einfahrenden Fahrzeuge als Indikator für Besuchshäufigkeit herangezogen wurden. Es kann also nicht gesagt werden, ob es an den Tagen, an denen viel Durchfluss an diesem Standort war, auch voll war. Gleichwohl kann dies als plausibel angenommen werden. So ist in Abbildung 9 beispielsweise gut zu erkennen, dass alle Standorte auf Fehmarn – mit einigen Ausnahmen – dasselbe Muster aufweisen: Viel Frequenz in den Spätsommermonaten, gefolgt von wenig Frequenz in den Wintermonaten bis hin zu einer wieder ansteigenden Frequenz etwa ab April 2023. Deutlich zu erkennen ist aber bei allen Parkplätzen der „Sylvester-Effekt“ um den Jahreswechsel. Das wiederkehrende Muster mit Spitzen und Tiefen lässt sich auch mit unterschiedlich starken Frequenzen an Wochentagen vergleichen (siehe beispielsweise Abbildung 10). Ebenfalls wird deutlich, dass vor allem im Süden der Insel, am Südstrand, mit den beiden Parkplätzen Burgtiefe und Südstrand Mitte, die meisten Einfahrten verzeichnet werden. An beiden Parkplätzen und so auch am Parkplatz Orth war das Tagesmaxima am 20.05. erreicht (Wochenende nach Christi Himmelfahrt). Während beim östlich gelegenen Parkplatz Meeschendorf der auf Basis der Einfahrten bemessene stärkste Besucherandrang im August 2022 war, war dies beim kleinen Parkplatz Grüner Brink im Norden der 3. September 2022.

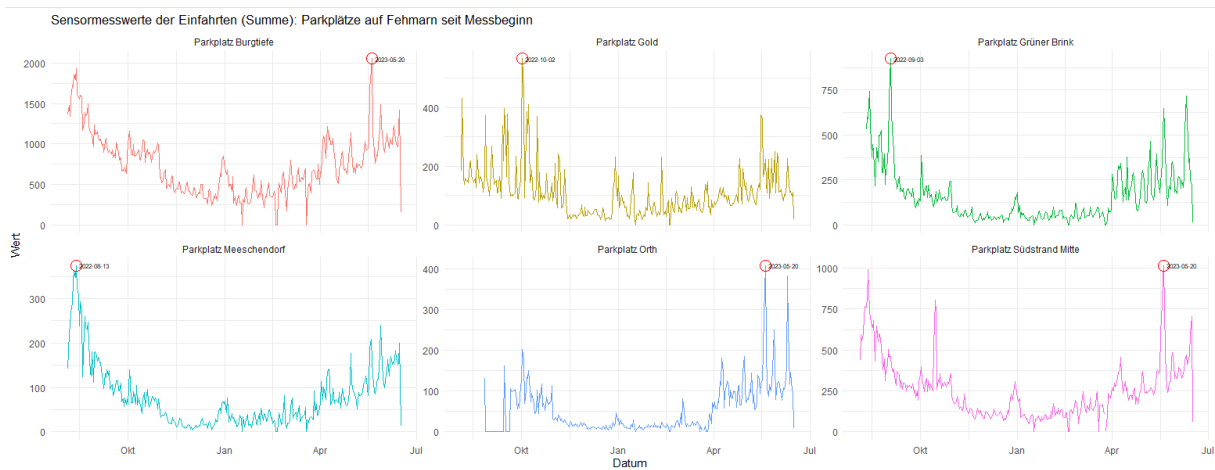


Abbildung 9: Zeitreihe der Summe der täglichen Einfahrten auf Parkplätzen auf Fehmarn

In Grömitz zeigt sich beim Parkplatz Königsredder Straße ein ähnliches Muster wie bei den Parkplätzen auf Fehmarn. Auch hier lassen sich mit Ausnahme beim Parkplatz Deichweg die Auswirkungen vom Jahreswechsel 2022/2023 in den Daten erkennen. In Grömitz hingegen ist der Großraumparkplatz mit Abstand der am meisten frequentierte Parkplatz von denen, die mit Sensorik ausgestattet sind. Auch in Grömitz wurden bereits Anfang August Sensoren in den Boden gelassen und so gab es kurz nach Installation beim Großraumparkplatz und beim Parkplatz Königsredder Straße am 26.08. bzw. am 13.08. die höchsten Messwerte bis Mitte Juni 2023. Bei den beiden Parkplätzen am Seebrückenvorplatz (Seestr. 31) und am Deichweg haben hingegen die Feiertage im Jahr 2023 für einen hohen Andrang von Besucher*innen gesorgt. Bei beiden Parkplätzen wurde die Sensorik jedoch erst im Oktober 2022 installiert. Der Ostersonntag war der meistbesuchte Tag beim Parkplatz Deichweg und der Pfingstsonntag beim Seebrückenvorplatz.

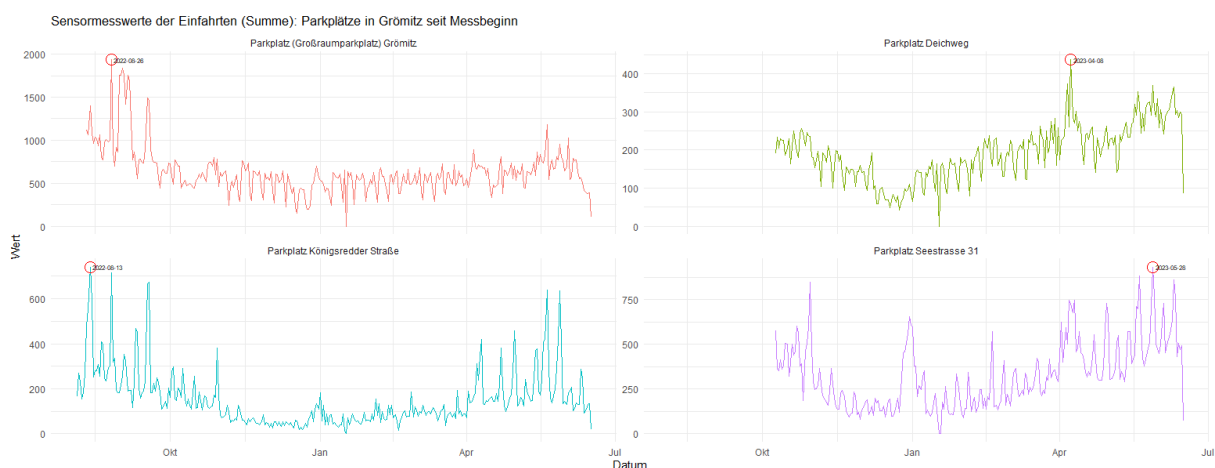


Abbildung 10: Zeitreihe der Summe der täglichen Einfahrten auf Parkplätzen in Grömitz

In Kappeln und Arnis liegt eine deutlich kleinere Zeitreihe zum aktuellen Stand vor (vgl. Abb. 11). Diese beginnt mit der Installation der Sensoren am 06.11.2022. Auch hier ist deutlich zu erkennen, dass ab April die Frequenzen der Einfahrten zunehmen, wenn auch hier der Großraumparkplatz die meisten

Einfahrten verzeichnet. Während in Kappeln das Wochenende nach Christi Himmelfahrt den meisten Andrang verzeichnete, war dies in Arnis der Pfingstsonntag.

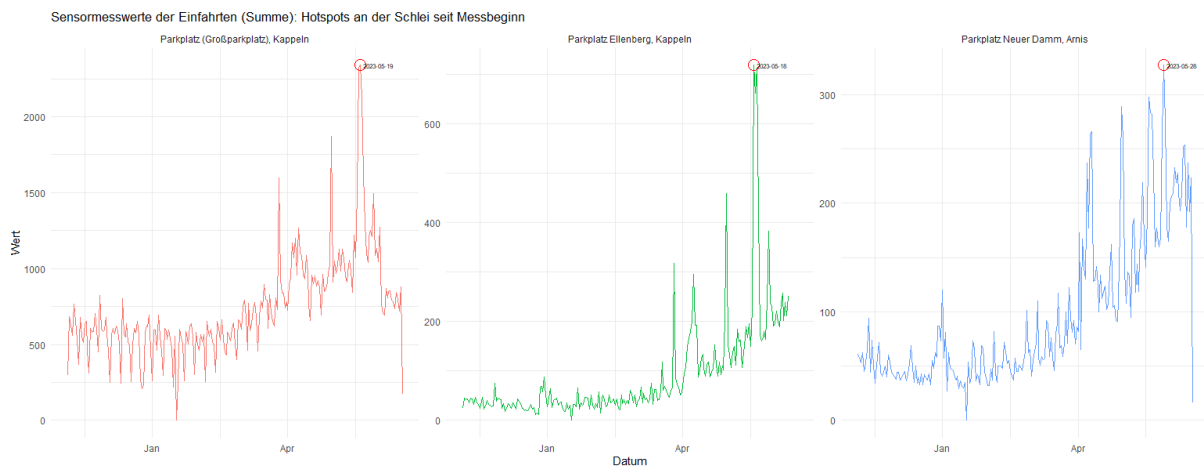


Abbildung 11: Zeitreihe der Summe der täglichen Einfahrten auf Parkplätzen in Kappeln und Arnis

Auch bei den Wohnmobilstellplätzen in der Holsteinischen Schweiz beginnt die Messreihe am 06.11.2022. Allerdings ist hier deutlich zu erkennen, dass bei den drei Campingplätzen Lange Wisch, Prinzenholz und Naturcamping kaum Frequenzen zu verzeichnen sind, da diese im Winter geschlossen sind (vgl. Abb. 12). Die wenigen Einfahrten weisen auf Versorgungs- und Instandhaltungsverkehr hin. Mit Öffnung der Campingplätze Mitte März bzw. Anfang April ist dann auch eine deutliche Steigerung der Frequenzen zu erkennen mit dann Spitzenwerten am Pfingstwochenende bzw. am 01. Mai beim Naturcamping Prinzenholz. Anders sieht dies beim Reisemobilpark in Eutin und beim „WoMo Stop“ in Malente aus. Hier wird ein in etwa gleichverteilter Durchfluss seit Messbeginn verzeichnet mit den Tagesmaxima am 18.3.2023 in Malente bzw. am 25.05. in Eutin.

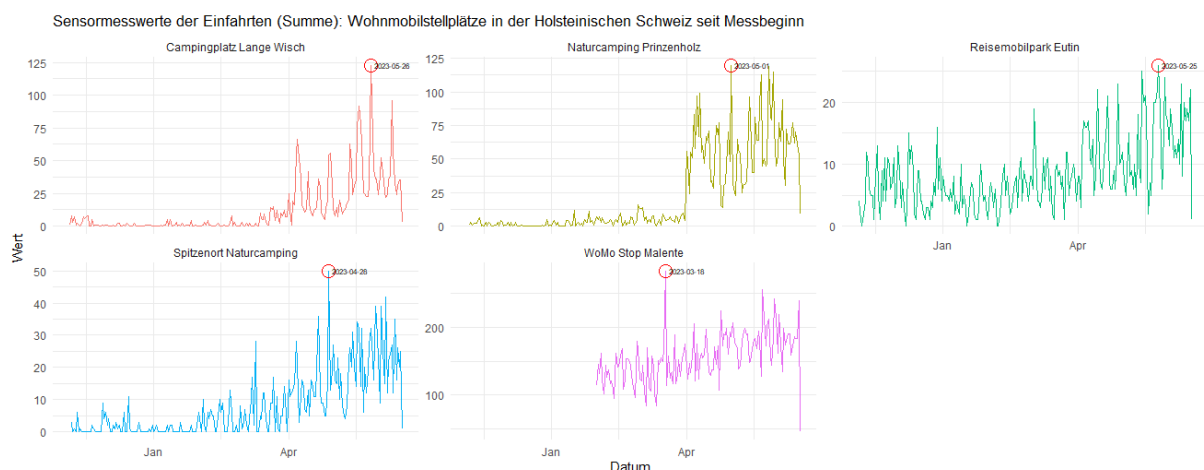


Abbildung 12: Zeitreihe der Summe der täglichen Einfahrten auf Campingplätzen und Wohnmobilstellplätzen

Einen Sonderfall stellen die Radrouten an der Schlei dar. Neben einer anderen Sensorart (siehe oben) gab es bei der Installation dieser einige Herausforderungen. Zum einen sind für die einzelnen Abschnitte der Radrouten diverse Straßenmeistereien zuständig. Die Vielzahl an Stakeholdern erschwerte die Abstimmungen bezüglich der Installationen. Darüber hinaus waren die Zuständigkeiten für einige Straßenabschnitte unklar, weshalb teilweise aufwendige Recherchen erfolgen mussten, um die zuständigen Personen auszumachen. Einige der ursprünglich geplanten Standorte lagen zudem auf Landesstraßen, bei denen sich der Erhalt einer Genehmigung als besonders herausfordernd darstellte, weshalb schließlich Alternativstandorte abseits der Landesstraßen ausgewählt wurden.

Auch wenn in Arnis der Sensor bereits am 21.02.2023 und am Campingplatz Hasselberg am 27.04.2023 installiert wurde, liegen hier erst seit dem 16.05.2023 für alle fünf Fahrradstandorte vergleichbare Daten vor. Immerhin konnten so die Feiertage Christi Himmelfahrt und Pfingsten mit in die Analyse einbezogen werden. Anders als bei den Parkplätzen handelt es sich hierbei nicht um die Summe von Einfahrten, sondern um Durchfahrten von Fahrrädern an diesem Standort (vgl. Abb. 13). Zur besseren Interpretation wurde die Fahrtrichtung der Räder mit angegeben. Christi Himmelfahrt und Pfingsten waren dann auch die beiden Feiertage mit dem höchsten Radverkehr. Kurz vor der Fähre Arnis, bei Sundsacker, war am Ostersonntag der meiste Verkehr in Richtung Fähre fahrend. Bemerkenswert hier sind die geringen Messwerte bis April und dann wieder ab Mai, die sich laut Aussagen des Sensorbetreibers auf technische Probleme mit dem Zählgerät zurückführen lassen, die nun behoben wurden.

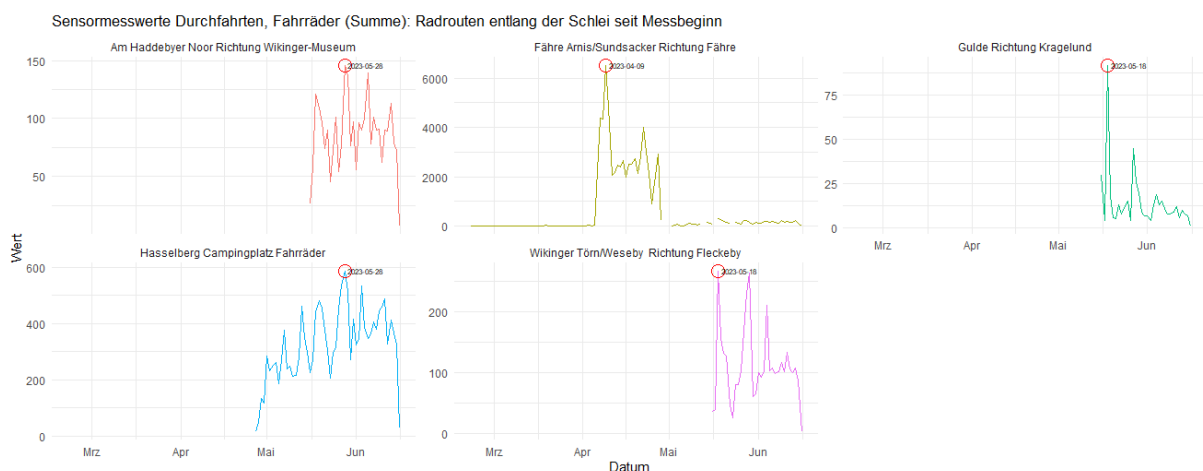


Abbildung 13: Zeitreihe der Summe der täglichen Durchfahrten von Fahrrädern entlang der Radrouten

Als weitergehende Analysen auf Basis der lokalen Sensorik wird an dieser Stelle beispielhaft für die Parkplätze auf Fehmarn und in Grömitz die Aggregation der Einfahrten nach Wochentagen visualisiert. Der Samstag ist bei allen Parkplätzen am meisten besucht. Während der Parkplatz Orth sonst unter der Woche kaum angefahren wird, weisen die Parkplätze Burgtiefe und Meeschendorf beispielsweise einen vergleichsweise ausgeglichenen Wochengang auf (vgl. Abb. 14). In Grömitz hingegen zeigen sich auch andere Spitzentage unter der Woche. So zum Beispiel der Freitag beim Großraumparkplatz oder der Sonntag beim Parkplatz an der Seebrücke (vgl. Abb. 15).

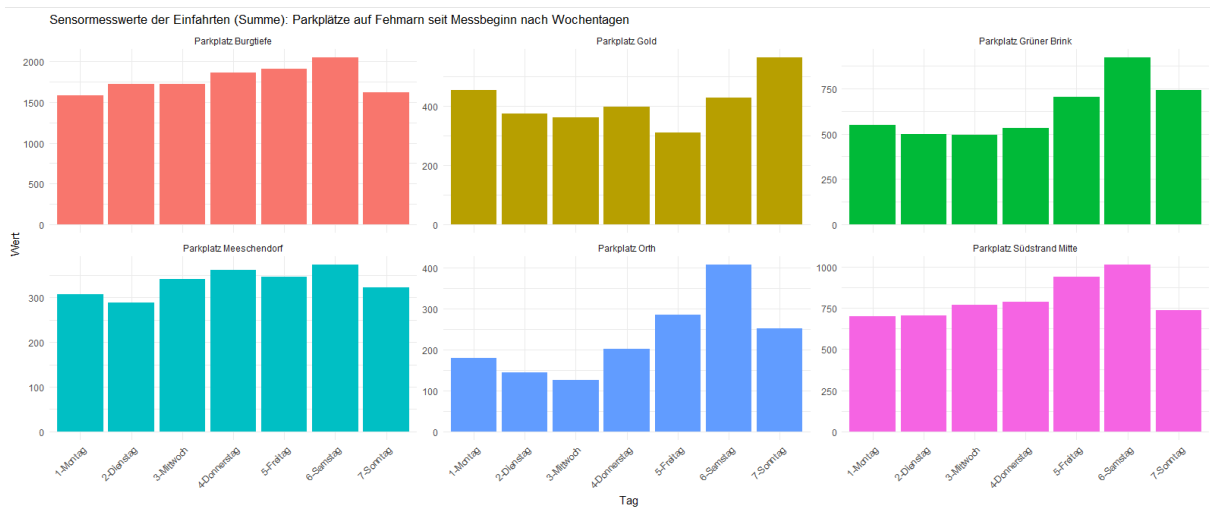


Abbildung 14: Summe der täglichen Einfahrten auf Parkplätzen auf Fehmarn nach Wochentagen

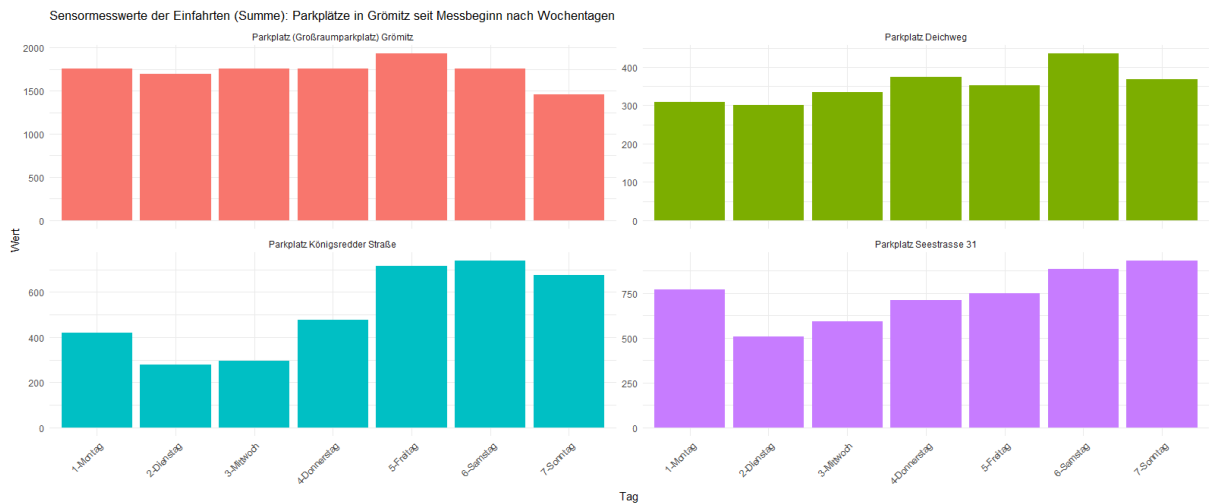


Abbildung 15: Summe der täglichen Einfahrten auf Parkplätzen in Grömitz

Wichtige Erkenntnisse liefert auch die Visualisierung der Einfahrten bzw. Durchfahrten bei Radwegen nach Stunden seit Messbeginn (vgl. Abb. 16). Mit Blick auf die Spitzenzeiten zeigt sich, dass der Großteil der Standorte seine höchsten Besucherzahlen zwischen 10:00 und 14:00 Uhr verzeichnet. Die jedoch mit einigen Ausnahmen wie der Parkplatz Südstrand Burgtiefe sowie die beiden Großparkplätze in Grömitz und in Kappeln, die ihre Spitzen später am Tag zwischen 14:00 und 17:00 Uhr haben. Auffällig ist ebenfalls, dass einige Standorte wie beispielsweise der WoMo Stop Malente, der Reisemobilpark Eutin aber auch die beiden auf Fehmarn gelegenen Parkplätze Südstrand Burgtiefe und Meeschendorf haben auch nachts eine beträchtliche Anzahl von Einfahrten verzeichnen, was darauf hindeutet, dass diese Orte wahrscheinlich als Übernachtungsorte genutzt werden. Die Messstandorte Spitzenort Naturcamping Plön und Naturcamping Prinzenholz in Eutin weisen im Vergleich zu anderen eine niedrigere Frequenz auf, während die beiden Großparkplatz Grömitz und Großparkplatz Kappeln die höchsten Werte verzeichnen. Hier steigt die Anzahl der Besucher*innen während des Tages stark an und in der Nacht wiederum stark ab. Standorte wie Südstrand Burgtiefe, Großparkplatz Kappeln und Parkplatz Ellenberg in Kappeln zeigen schon früh am Morgen einen starken Anstieg der Einfahrten.

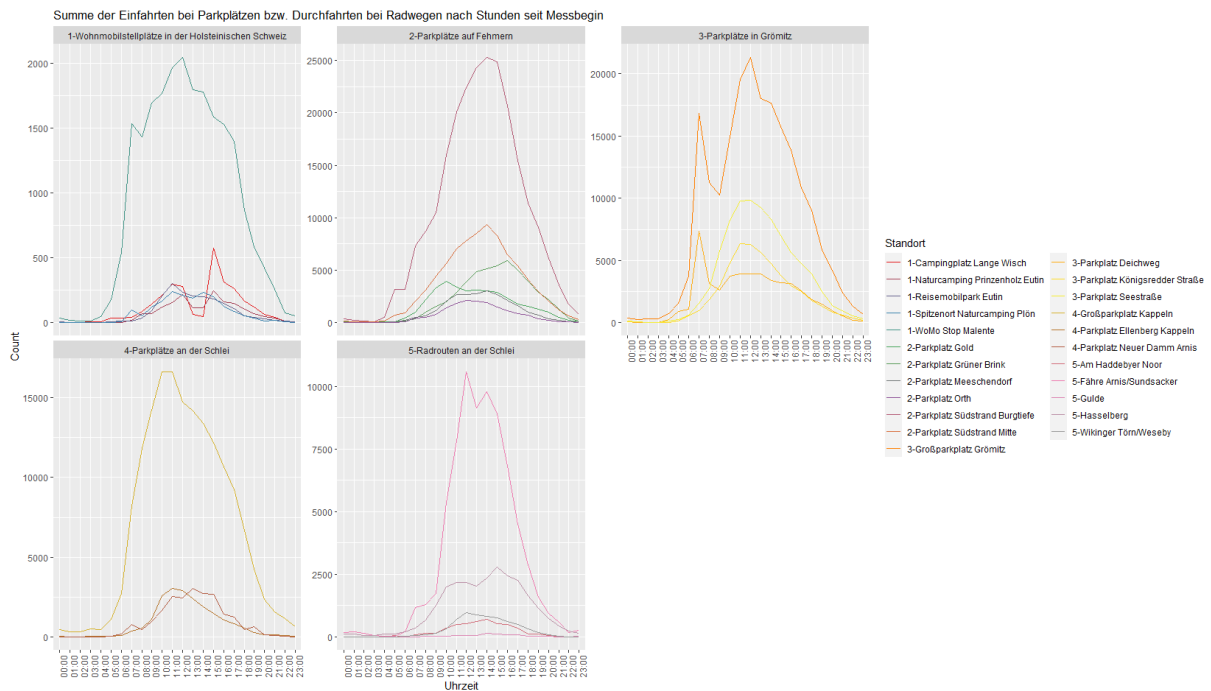


Abbildung 16: Summe der Einfahrten bzw. Durchfahrten nach Stunden seit Messbeginn je Use Case (Parkplätze Ostsee aufgeteilt nach Fehmarn und Grömitz)

Weitere Analysemöglichkeiten sind im Rahmen der lokalen Sensorik möglich. Hierfür sei auf das Dashboard Eco Visio verwiesen. Tabelle 8 gibt eine Übersicht über die von der lokalen Sensorik gemessenen Einfahrten an den Messstandorten. Dabei ist zu beachten, dass im Falle von mehreren Messpunkten an einem Standort diese in einer Zählstelle zusammengefasst wurden. Im Falle der Radrouten wurde die Summe beider Fahrtrichtungen betrachtet.

Standort	Messzeitraum	Summe der Einfahrten	Durchschnitt täglich	Werktags-Durchschnitt	Wochenend-Tagesdurchschnitt	Täglicher Median	Spitzentag	Spitzenwert	Wochentags-Höchstwert
Use Case Wohnmobilstellplätze Holsteinische Schweiz									
WoMo Stop Malente	31.01.23-16.06.23	21.598	158	166	137	157	Samstag, 18.03.2023	284	Montag
Reise mobilpark Eutin	06.11.22-16.06.23	1.870	8	9	8	7	Donnerstag, 25.05.2023	26	Freitag
Spitzenort Naturcamping Plön	07.11.22-16.06.23	1.829	8	9	6	4	Freitag, 28.04.2023	50	Freitag

Natur-camping Prinzenholz Eutin	07.11.22-16.06.23	1.678	8	7	9	4	Samstag, 29.04.2023	41	Samstag
Campingplatz Lange Wisch	06.11.22-16.06.23	2.777	12	13	12	2	Freitag, 26.05.2023	123	Freitag
Use Case Fehmarn									
Parkplatz Südstrand Burgtiefe	05.08.22-16.06.23	235.721	746	732	781	682	Samstag, 20.05.2023	2052	Samstag
Parkplatz Meeschen-dorf	05.08.22-16.06.23	24.371	77	73	86	63	Samstag, 13.08.2022	349	Sonntag
Parkplatz Südstrand Mitte	05.08.22-16.06.23	81.186	258	247	286	231	Samstag, 20.05.2023	1015	Samstag
Parkplatz Orth	27.08.22-16.06.23	15.723	54	49	66	24	Samstag, 20.05.2023	407	Samstag
Parkplatz Grüner Brink	11.08.22-16.06.23	49.168	159	140	205	113	Samstag, 03.09.2022	832	Sonntag
Parkplatz Gold	05.08.22-16.06.23	35.011	111	101	136	92	Sonntag, 02.10.2022	566	Sonntag
Use Case Grömitz									
Parkplatz Seestraße	09.10.22-16.06.23	83.482	333	293	433	308	Sonntag, 28.05.2023	932	Sonntag
Parkplatz Königsredder Straße	05.08.22-16.06.23	48.759	154	132	211	124	Samstag, 13.08.2022	740	Sonntag
Großparkplatz Grömitz	11.08.22-16.06.23	200.923	648	682	562	608	Freitag, 26.08.2022	1933	Freitag
Parkplatz Deichweg	09.10.22-16.06.23	48.408	193	204	166	192	Samstag, 08.04.2023	436	Freitag
Use Case Hotspots Schlei									
Großparkplatz Kappeln	06.11.22-16.06.23	164.436	737	763	671	655	Freitag, 19.05.2023	2339	Donnerstag
Parkplatz Ellenberg Kappeln	06.11.22-16.06.23	19.794	91	84	109	46	Donnerstag, 18.05.2023	718	Samstag

Parkplatz Neuer Damm Arnis	06.11.22-16.06.23	21.516	96	90	113	65	Sonntag, 28.05.2023	327	Sonntag
Use Case Radrouten Schlei									
Wikinger Törn/We-seby	16.05.23-16.06.23	6.551	205	179	283	176	Donners-tag, 18.05.2023	470	Sonntag
Fähre Arnis/Sunds acker	21.02.23-16.06.23	74.409	670	599	863	166	Sonntag, 09.04.2023	6526	Samstag
Am Had-debyer Noor	08.04.23-16.06.23	4.958	92	96	83	121	Sonntag, 28.05.2023	265	Freitag
Hassel-berg	27.04.23-16.06.23	25.678	503	468	597	516	Sonntag, 28.05.2023	877	Samstag
Gulde	01.05.23-16.06.23	843	26	23	36	19	Samstag, 27.05.2023	125	Samstag

Tabelle 8: Übersicht der Messstandorte seit Messbeginn

Neben der Zählung liegen die Vorteile der lokal verbauten Sensorik in der Richtungserkennung und in der Identifizierung von verschiedenen Fahrzeugklassen. Letzteres ist insbesondere für die Wohnmobilstellplätze in der Holsteinischen Schweiz von zentraler Bedeutung. Die unterschiedlichen Fahrzeugklassen werden aufgrund der Veränderungen des Magnetfeldes des Sensors bestimmt, wenn ein eisenhaltiges Fahrzeug über den Sensor fährt (siehe oben). So ist in Abbildung 17 nicht nur zu erkennen, dass der Womo-Stop in Malente der Stellplatz ist, der die höchste Frequenz verzeichnet, sondern es ist auch der Parkplatz, bei dem der Pkw dominiert. Diese Abbildung zeigt auch die Summe der Incoming-Messwerte unterteilt nach den verschiedenen Fahrzeugklassen und Wochentagen auf. Es wird nicht nur deutlich, dass an allen Stellplätzen freitags die meisten Fahrzeuge verzeichnet werden, sondern auch dass neben den Pkws Kleinbusse dominieren. Dabei handelt es sich im Wesentlichen um Wohnmobile (aber auch Campervans). Größere Wohnmobile können auch in die Kategorie der LKWs fallen. Dies ist beispielsweise am vergleichsweise hohen Anteil an LKW auf dem Campingplatz Spitzenort gut zu erkennen.

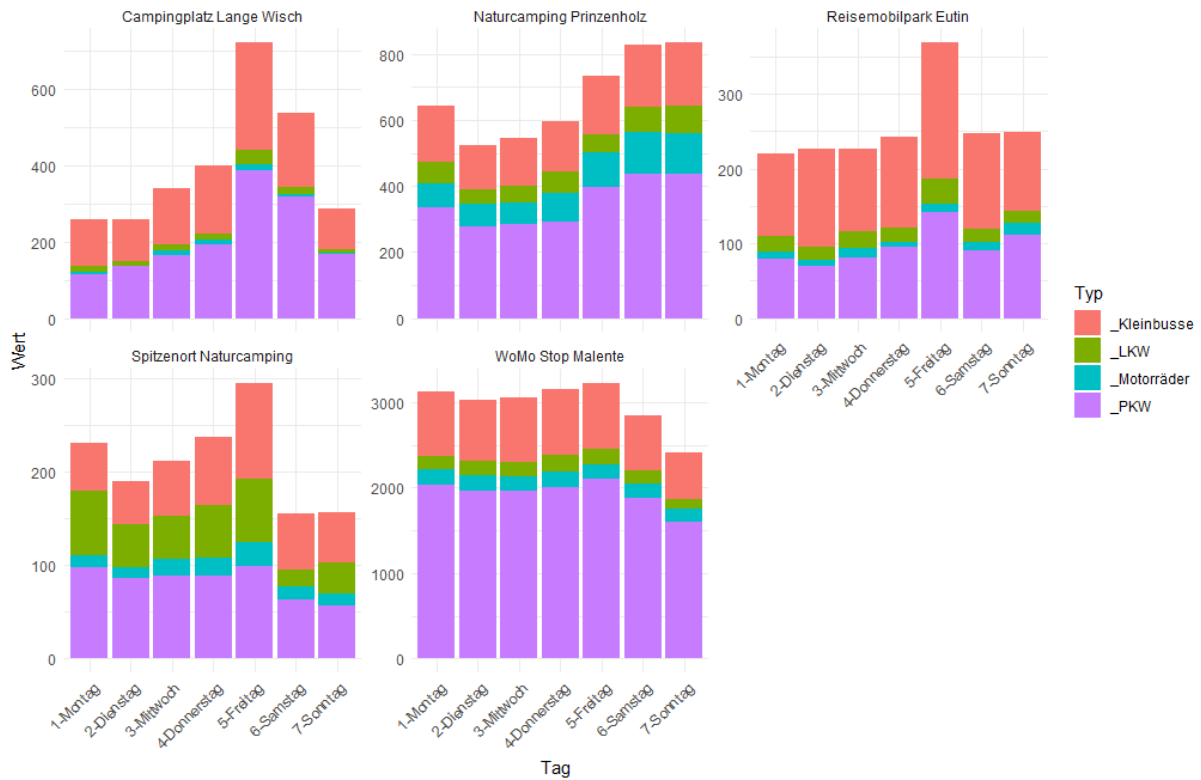


Abbildung 17: Zeitreihe der Summe der täglichen Einfahrten auf Campingplätzen und Wohnmobilstellplätzen nach Fahrzeugtyp und Wochentag

3.2.2 Globale Sensorik

Nutzung globaler Daten zur Frequenzmessung

Mit Daten aus globaler Sensorik lassen sich ebenfalls Aussagen zu Frequenzen an einzelnen POI tätigen. Ein Abgleich mit den im Rahmen von lokal installierter Sensorik generierten Frequenzdaten erfolgt für ausgewählte Standorte in Abschnitt 3.3.

Am Beispiel der Parkplätze in Grömitz lassen sich mit Hilfe von Big Data bereits erste Aussagen zur Saisonalität im Jahresverlauf zeigen. Nachfragespitzen wie beispielsweise in der Kalenderwoche 15 (Ostern) werden so sichtbar (vgl. Abb. 18). Weiterhin wird deutlich, dass die Gesamtanzahl der daily visits von Inländern im Jahr 2022 sehr gering ist. Dies wird besonders deutlich beim Parkplatz Deichweg (n=935).



Abbildung 18: Parkplätze in Grömitz: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppe

Info: Blauer Hintergrund sind die Sommerferien in Schleswig-Holstein in den Kalenderwochen 27-32.

Ähnlich verhält es sich mit den Parkplätzen auf Fehmarn (vgl. Abb. 19). Hier ist bemerkenswert, dass die kleineren Parkplätze Orth, Gold, Meeschendorf und Grüner Brink über das Jahr hinweg kaum Signale vorweisen. Dies ist u. a. auf die Polygongröße zurückzuführen, die es nicht erlaubt, dass dort ausreichend Location Events generiert werden. Auffallend beim Parkplatz Burgtiefe sind die Nachfragespitzen in den Herbstferien, welche insbesondere aus einem hohen Aufkommen von Übernachtungs-Besucher*innen resultiert.

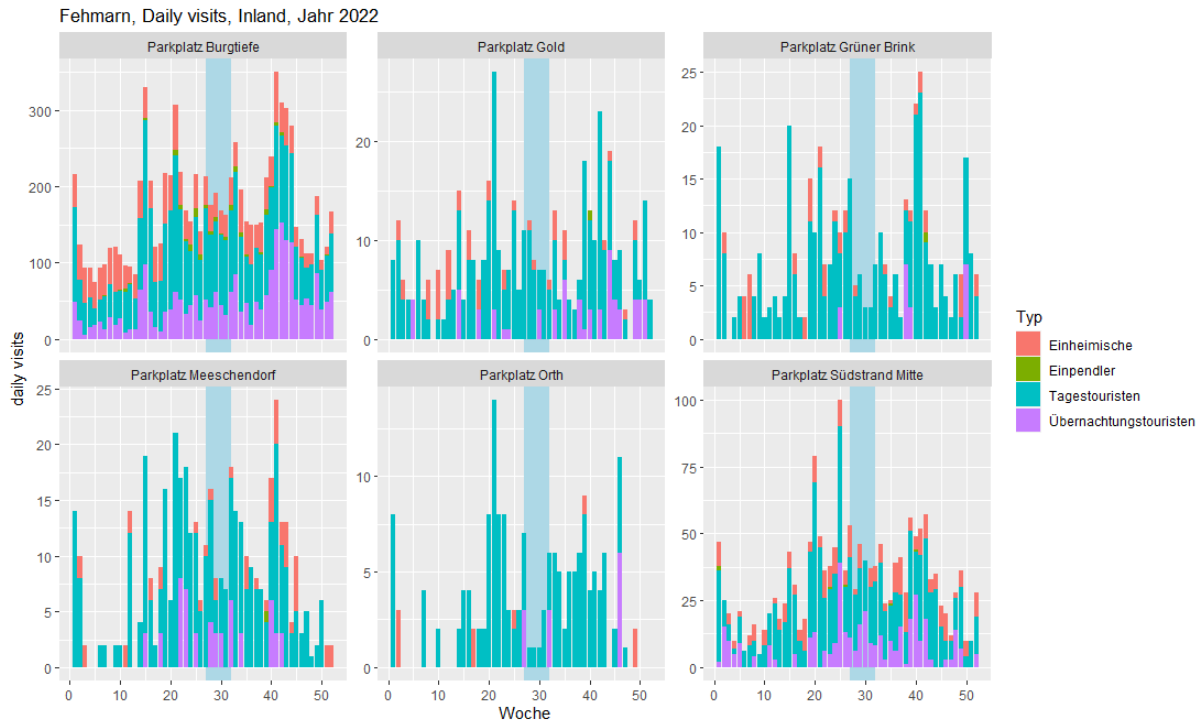


Abbildung 19: Parkplätze auf Fehmarn: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppen

Info: Blauer Hintergrund sind die Sommerferien in Schleswig-Holstein in den Kalenderwochen 27-32.

Auch bei den drei Polygonen in Kappeln und Arnis in Abbildung 20 zeigt sich die Bedeutung der für die Analyse genutzten Polygongröße. Beispielsweise werden in Kappeln ($n=3.316$) rund fünfmal so viele daily visits von Endgeräten generiert wie in Arnis ($n=661$). Auch in Kappeln zeigen sich mehrere Nachfragespitzen, so beispielsweise an Ostern (Kalenderwoche 15/16) aber auch in Kalenderwoche 20 sowie 42 des Jahres 2022.

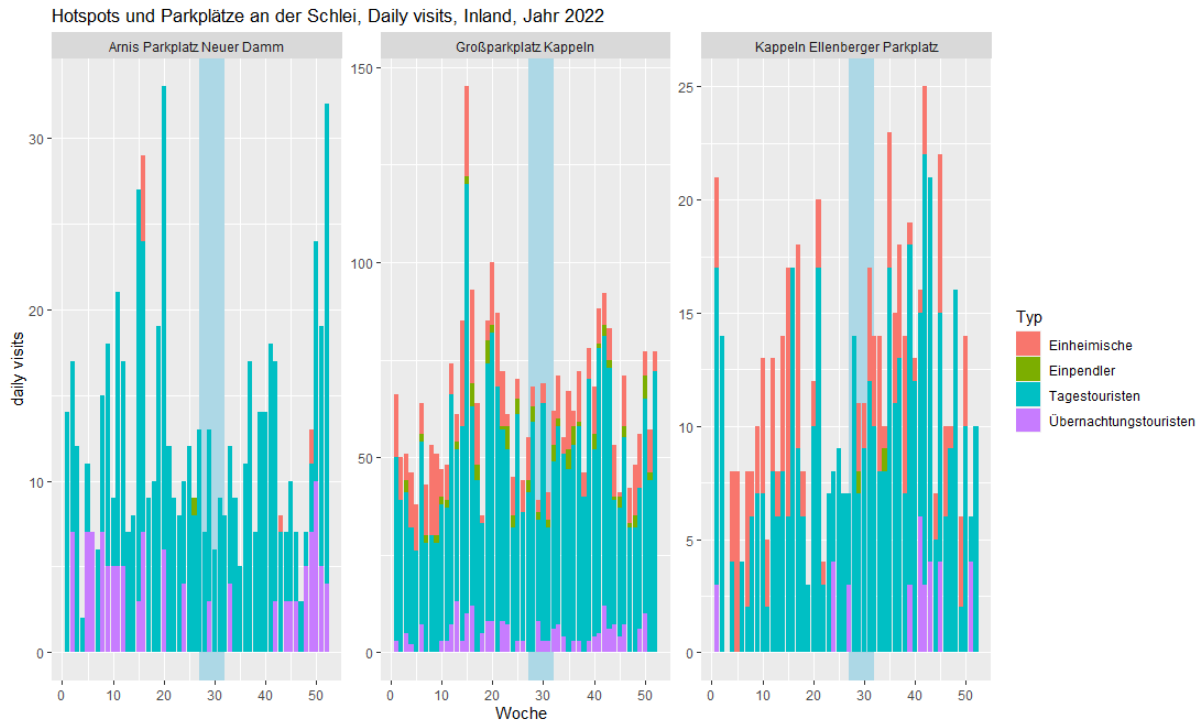


Abbildung 20: Parkplätze und Hotspots an der Schlei: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppen

Info: Blauer Hintergrund sind die Sommerferien in Schleswig-Holstein in den Kalenderwochen 27-32.

Die ausgewählten Polygone für die Radwege zeigen eine sehr geringe Reliabilität auf. Insbesondere bei den Standorten Gulde (n=11 (!)) und Drecht (n= 167) zeigen sich bei der Analyse auf Basis von Kalenderwochen deutliche Datenlücken (vgl. Abb. 21). Die Zuwegung zum Wikingermuseum Haithabu ist hingegen ab April 2022 auswertbar, wenn auch hier mit geringen Fallzahlen der daily visits (n=357).



Abbildung 21: Radwege an der Schlei: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppen

Bemerkenswert ist hingegen die Analyse der Campingplätze und Wohnmobilstellplätze. In Abbildung 22 zeigt sich vor allem bei den drei Campingplätzen Camping an der Schwentine, Spitzenort und Prinzenholz der Beginn und das Ende der Campingsaison, die in etwa von der Kalenderwoche 15-40 im Jahr 2022 reicht. Die beiden Stellplätze „WomoStop“ und der Reisemobilpark zeigen hingegen einen über das Jahr ausgeglichenen Verlauf.



Abbildung 22: Wohnmobilstellplätze in der Holsteinischen Schweiz: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppe

Auch bei den ausgewählten Hotspots an der schleswig-holsteinischen Westküste zeigt sich die Relevanz der Polygongröße bei der Anzahl der daily visits (vgl. Abb. 23). Während die gesamte Insel Amrum als Polygon im Jahr 80.318 daily visits von Inländern aufweist, sind es bei der Seebrücke in St. Peter-Ording lediglich 5.540, was im Vergleich zu den Parkplätzen jedoch deutlich mehr ist. Auch wenn das Polygon nicht besonders groß ist, handelt es sich bei der Seebrücke jedoch um einen touristischen Attraktionspunkt bei dem davon auszugehen ist, dass Personen hier verweilen und die Aussicht genießen (beispielsweise auf den Sitzbänken) und so auch eher die Möglichkeit haben das Smartphone zu nutzen (beispielsweise um zu fotografieren und die Fotos online zu teilen) und demnach auch mehr Signale generieren. An allen Standorten der Westküste ist der Peak an Ostern zu erkennen.



Abbildung 23: Standorte an der Westküste Schleswig-Holstein: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppe

Beim Strand in Scharbeutz zeigt Abbildung 24 deutlich die Nachfragespitze in den Sommerferien um die Kalenderwochen 27-32. Während die Parkplätze auch hier deutlich weniger Signale aufweisen.

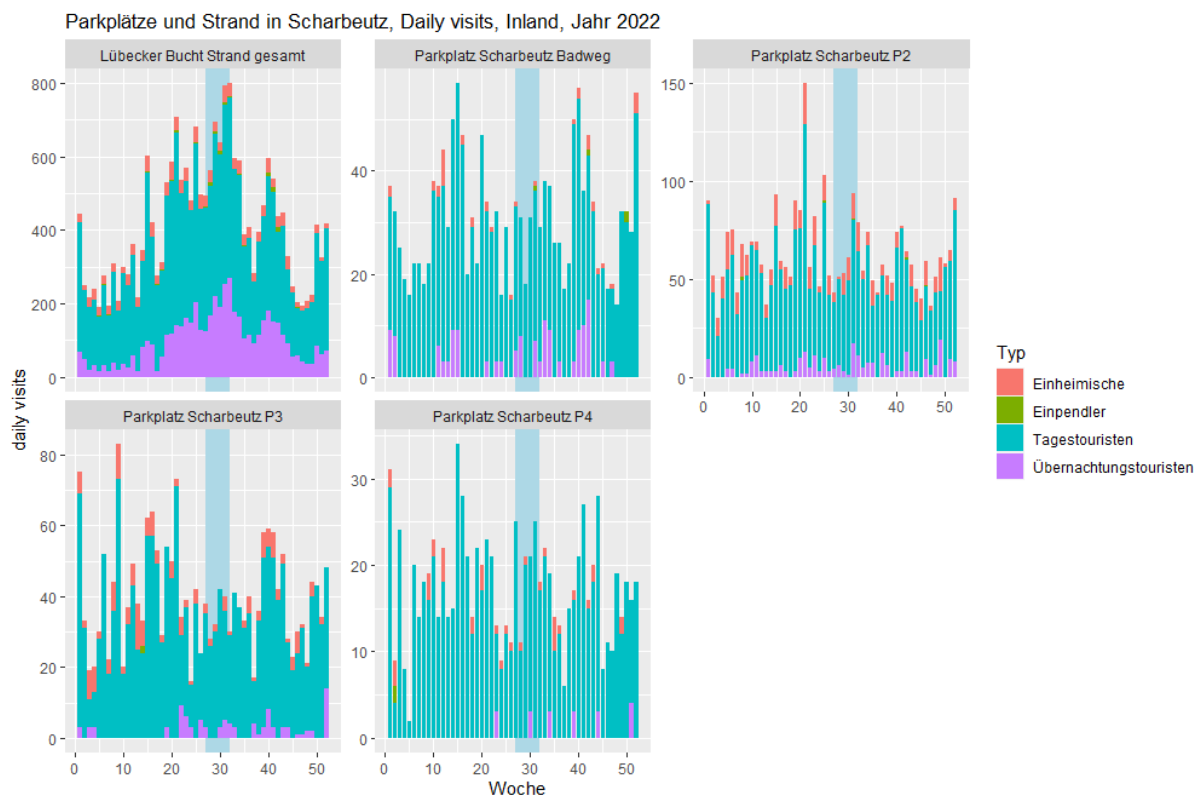


Abbildung 24: Parkplätze und Strand in Scharbeutz: Daily visits aus dem Inland im Jahr 2022 nach Besuchergruppe

Um zu überprüfen, ob die oben je Use Case aufgezeigten daily visits über das Jahr 2022 gleich verteilt sind oder ob es zu einer Konzentration der Nachfrage kommt, bieten sich Konzentrationsmaße an. Dabei gibt es verschiedene Möglichkeiten, die saisonale Variation eines POI zu verdeutlichen. Ein nützliches Maß dafür ist der Gini-Saisonalitätskoeffizient, der auf der Lorenzkurve basiert und insbesondere bei der Analyse regionaler Unterschiede von großer Bedeutung ist (Eisenstein & Reif, 2017). Der hier berechnete Koeffizient bezieht sich auf die daily visits im Jahr 2022 und gibt Aufschluss über die Verteilung dieser auf die Kalenderwochen. Die Ausprägung des Koeffizienten variiert zwischen 0 und 1. Ein Wert von 0 impliziert eine gleichmäßige Verteilung der Ereignisse über sämtliche Kalenderwochen, während ein Wert von 1 darauf hindeutet, dass sämtliche daily visits auf eine einzige Kalenderwoche fokussiert sind. Wenn der Koeffizient größer als 0,4 ist, wird von einer sehr hohen Saisonalität gesprochen (Zimmermann-Janschitz, 2014). Auffallend sind die hohen Konzentrationswerte bei den Campingplätzen in der Holsteinischen Schweiz, die sich aufgrund der Öffnungszeiten und den Nachfragespitzen in den Sommerzeiten erklären lassen oder auch die erhöhte Saisonalität bei den Radwegen an der Schlei, die witterungsbedingt eine stärkere Konzentration über das Jahr hinweg aufweisen, aber gleichzeitig auch viele Kalenderwochen ohne Daten vorliegen haben. Mit einem Gini-Koeffizienten von 0,54 weist der Parkplatz Orth auf Fehmarn die höchste und die Kieler Innenstadt (hier das Polygon wo maßgeblich die Kieler Woche stattfindet) mit einem Gini-Koeffizienten von 0.05 die niedrigste Saisonalität aller hier untersuchten Polygone auf (vgl. Abb. 25). Die Saisonalität sagt jedoch nichts über die Frequenzen aus: POI mit einer hohen Saisonalität können also auch über das Jahr hinweg eine niedrige Besucher*innenfrequenz aufweisen und vice versa.

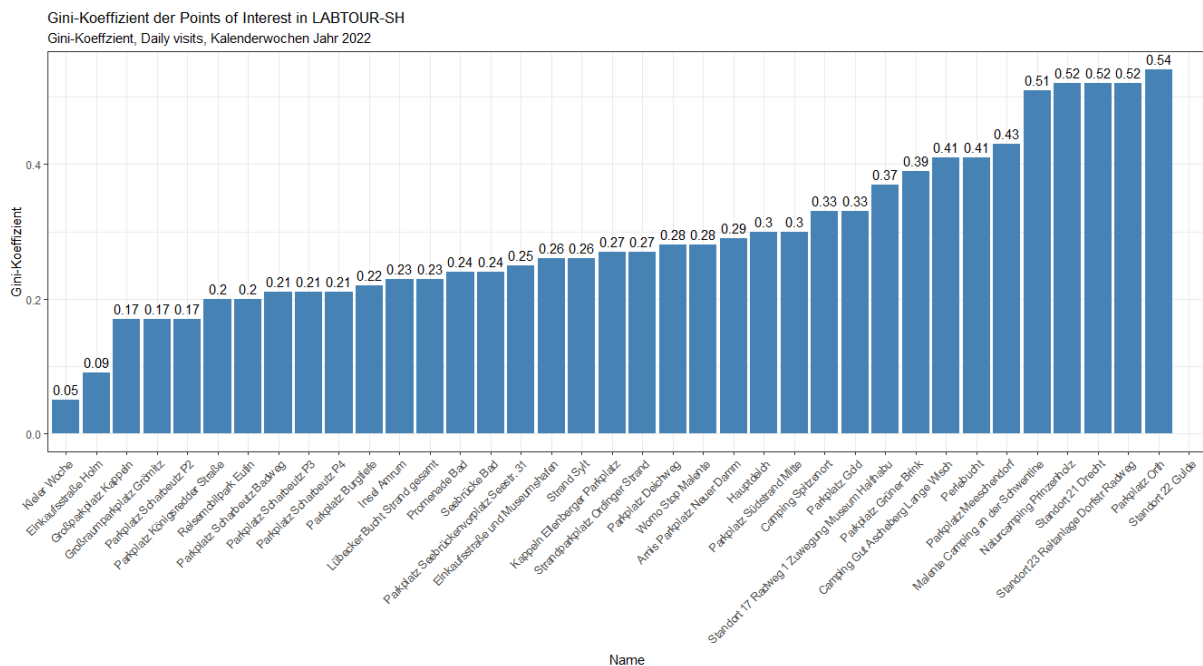


Abbildung 25: Gini-Koeffizient der Daily visits im Jahr 2022 je Polygon

Quelle: Eigene Darstellung und Berechnung; Standort Gulde: zu wenig Daten.

Neben der Analyse der daily visits nach Kalenderwochen ist es auch möglich die Daten nach Saison bzw. nach Uhrzeit und Wochentag zu betrachten. Dies erfolgt hier nicht für alle Polygone, sondern beispielhaft für ausgewählte Standorte um Besonderheiten aufzuzeigen. Die Darstellung erfolgt nach Wochentagen und nach Saisons. Dabei wird nicht das Kalenderjahr 2022 herangezogen, sondern ein

leicht versetzter 12-Monatszeitraum vom Dezember 2021 bis November 2022 wobei der Winter den Dezember 2021 bis inklusive Februar 2022 beinhaltet, der Frühling den März 2022 bis Mai 2022 usw. (siehe oben).

Abbildung 26 zeigt, dass bei allen in Grömitz betrachteten Parkplätzen Tagestourist*innen einen Großteil der daily visits ausmachen. Während Übernachtungs*touristinnen in allen Saisons vermehrt den Großraumparkplatz nutzen, generieren Einheimische hier vor allem im Frühling und Winter daily visits. Auffällig ist zudem, dass die Parkplätze Deichweg und Königsredder Straße vor allem freitags geringere Besucher*innenfrequenzen aufweisen, bevor diese zum Wochenende wieder ansteigen. Im Gegensatz dazu kann vor allem am Parkplatz Seebrückenvorplatz Seestr. 31 eine fortwährend steigende Nachfrage ab Mitte der Woche erkannt werden, die ihren Höhepunkt am Samstag bzw. Sonntag erreicht.

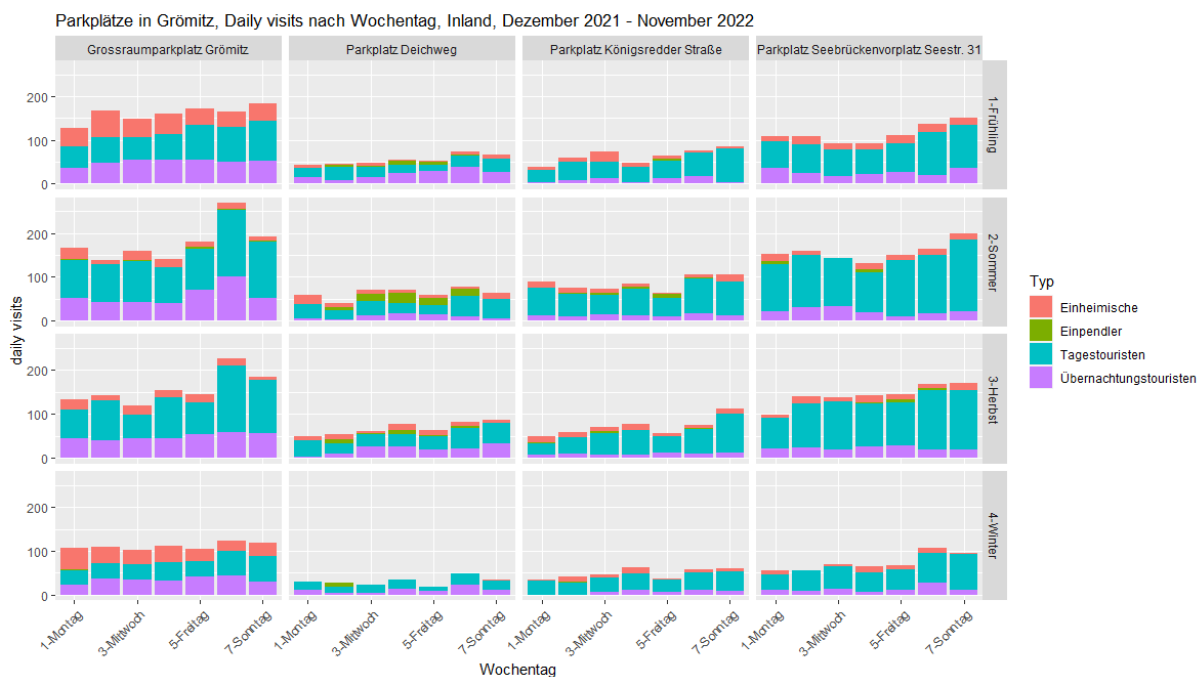


Abbildung 26: Daily visits auf den Parkplätzen in Grömitz nach Wochentagen, Typ und Saison

Die Hotspots an der Westküste zeigen stark variierende Besucher*innenfrequenzen mit der höchsten Anzahl an daily visits im Polygon Einkaufsstraße und Museumshafen und der deutlich geringsten Summe an daily visits an der Seebrücke Bad (vgl. Abb. 27). Die Orte Hauptdeich, Perlebucht, Promenade Bad und Strandparkplatz Ordinger Strand zeichnen sich zudem durch einen ähnlichen Saisonverlauf der daily visits aus, da hier die Nachfragespitzen im Sommer und Herbst mit einem deutlichen Rückgang im Winter ersichtlich werden. Obwohl gleiches für die Einkaufsstraße und den Museumshafen gilt, sind hier auch hohe Besucher*innenfrequenzen im Frühling und, wenn auch etwas geringer, im Winter auszumachen. Zusätzlich ist zu erkennen, dass lediglich in diesem Polygon Einheimische gemessen wurden, während alle weiteren Hotspots an der Westküste fast ausschließlich von Tages- bzw. Übernachtungstourist*innen frequentiert wurden. Die Analyse nach Wochentagen ergibt zudem, dass kein eindeutiger Trend zu einem hohen Besuchsaufkommen am Wochenende ersichtlich ist und die Frequentierung im Wochenverlauf variiert bzw. ähnlich hoch ist.

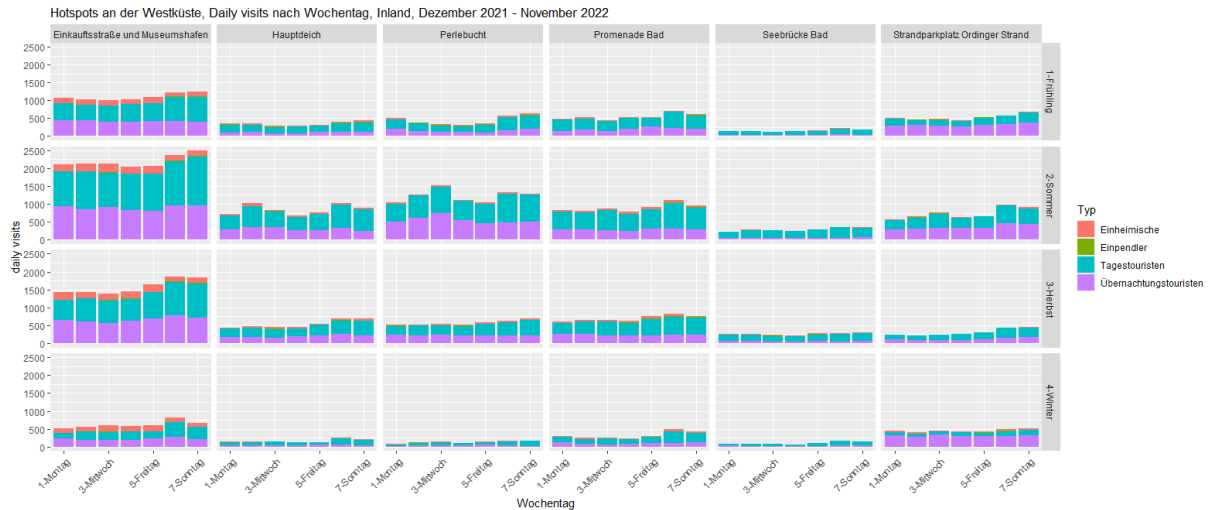


Abbildung 27: Daily visits an den Hotspots der Westküste nach Wochentagen, Typ und Saison

Abbildung 28 verdeutlicht den typischen tageszeitlichen Verlauf für Standorte an der Westküste, der im Allgemeinen die Form einer Glockenkurve annimmt. Anders sieht die zeitliche Verteilung der daily visits auf den Camping- und Wohnmobilstellplätzen in der Holsteinischen Schweiz (vgl. Abb. 29) aus. Hier kann eine bimodale Kurve identifiziert werden, bei der zwei Tagesmaxima auftreten, beispielsweise in den Morgen- und Abendstunden, an denen sich Personen in der Regel auch auf den Plätzen aufhalten. Weiterhin ist der Schwerpunkt der daily visits in der Sommersaison zu erkennen.

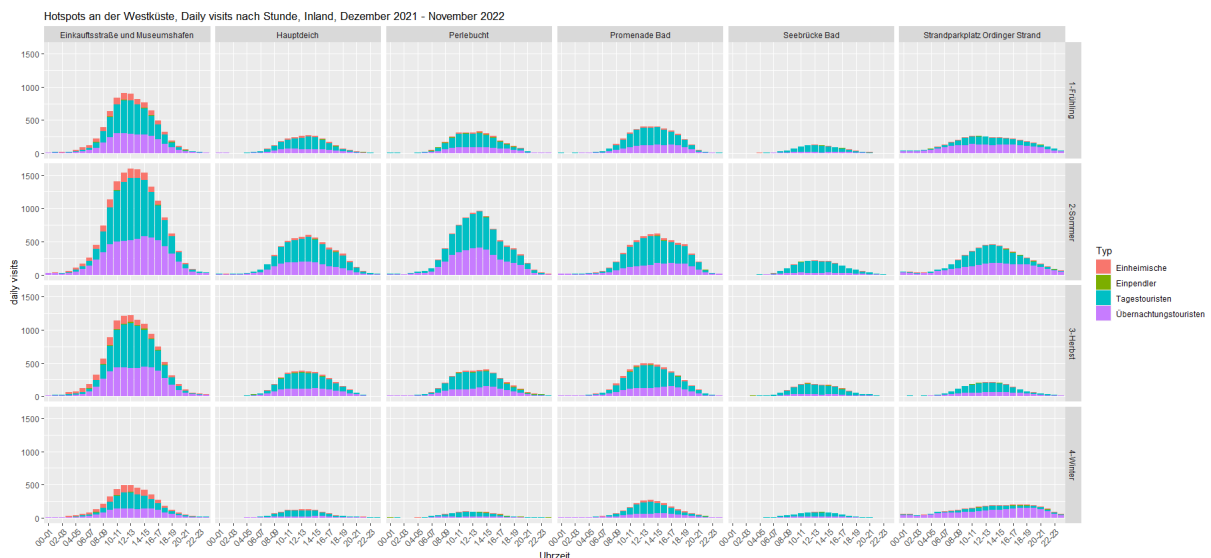


Abbildung 28: Daily visits an den Hotspots der Westküste nach Stunden, Typ und Saison

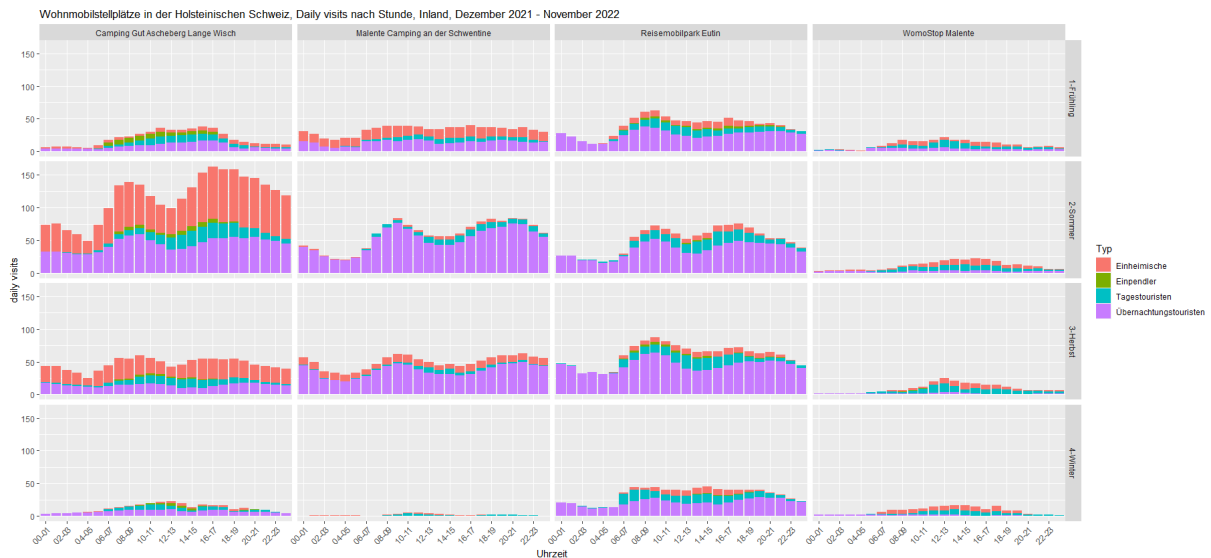


Abbildung 29: Daily visits auf den Wohnmobilstellplätzen und Campingplätzen nach Stunden, Typ und Saison

Nutzung globaler Daten zur Anreicherung

Die über lokale Sensorik generierten Frequenzdaten haben ihre Stärke in der Beantwortung der Fragestellung, wie viele Personen, Autos etc. zu einem bestimmten Zeitpunkt an einem Ort sind. Sie lassen jedoch keine Rückschlüsse darüber zu, um welche Art von Frequenzen es sich handelt. Die lokale Sensorik ist also nicht in der Lage, Signale über einen gewissen Zeitraum zu tracken. Anders ausgedrückt: Die lokale Sensorik kann zwar Aussagen darüber liefern, wann es am Parkplatz in Grömitz voll ist, nicht aber wer für die Überfüllung zuständig ist (Einwohnende, Tagestourist*innen etc.) und auch nicht, wenn jemand auf dem Parkplatz war, wo er oder so sie vorher war und wo am Abend. Diese Aussagen können aber die globale Sensorik geben.

Für die Parkplätze an der Ostsee wurden daher für das Jahr 2022 die daily visits nach den Segmenten Einwohnende, Tagestourist*innen, Übernachtungstourist*innen und Sonstige analysiert (vgl. Abb. 30). Die dazugehörigen Definitionen ergeben sich aus der unterschiedlichen Annahme von Home- und Workzone je Betrachtungsgebiet (vgl. Abschnitt 3.1). Übergreifend zeigt sich sowohl für Fehmarn als auch für Grömitz die hohe Bedeutung des Tagestourismus, der über alle untersuchten Polygone über die Hälfte der verzeichneten daily visits ausmacht. Gleichwohl zeigen sich mit Blick auf den Übernachtungstourismus Unterschiede. Während dieser beispielsweise bei den Parkplätzen in Grömitz (Deichweg und Großraumparkplatz) oder auch südlich auf Fehmarn (Südstrand und Burgtiefe) über ein Fünftel aller daily visits ausmacht, werden die Parkplätze Grüner Brink und Orth (beide Fehmarn) sowie der Parkplatz in der Königsredderstraße kaum von Übernachtungstourist*innen aufgesucht. Der Anteil der Einwohnenden ist bei den beiden Parkplätzen am Südstrand in Fehmarn mit Anteilen von 18% (Südstrand) und 22% (Burgtiefe) am höchsten.

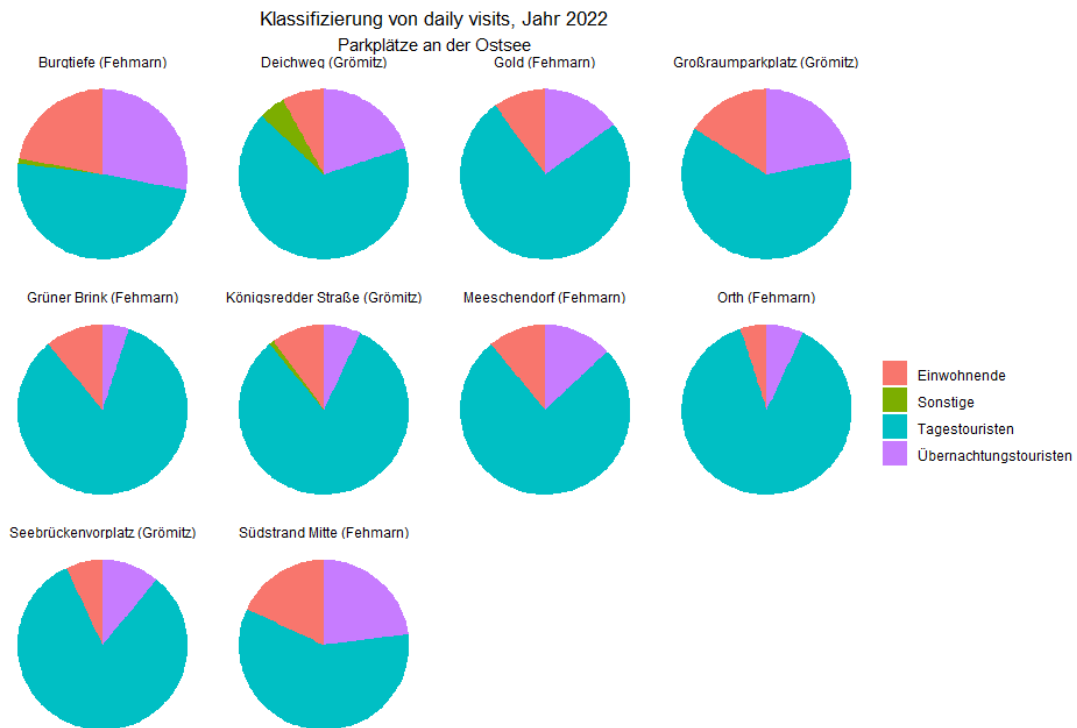


Abbildung 30: Parkplätze an der Ostsee: Daily visits nach Besuchergruppen

Auch in Scharbeutz wird die Bedeutung des Tagestourismus deutlich (vgl. Abb. 31). Auffallend hier ist, dass der Parkplatz P2, nördlich gelegen, unter den vier untersuchten Parkplätzen den höchsten Anteil an Signalen von Einwohnern hat (14%). Die Vermutung hier ist, dass die anderen Parkplätze eher von Gästen aufgesucht werden, da diese die näher gelegenen Parkplätze sind, wenn Gäste aus Hamburg anreisen, dem Hauptquellmarkt. Der Strand weist einen Anteil von rund einem Viertel Übernachtungsgäste aus.

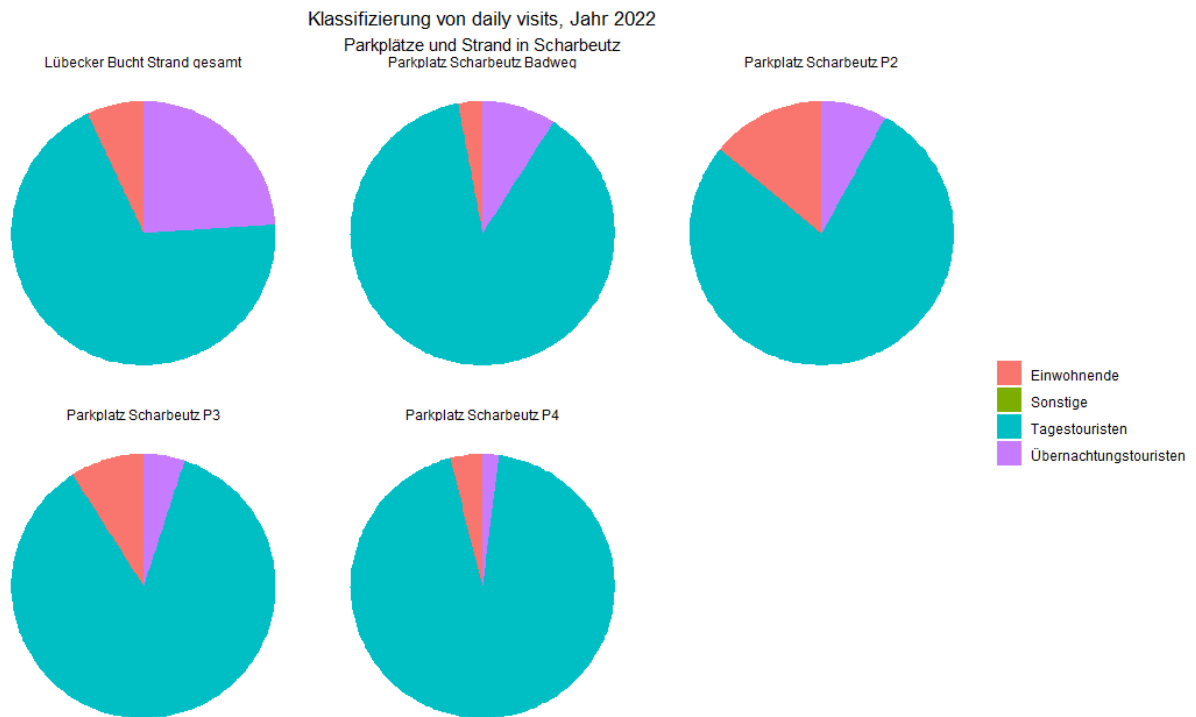


Abbildung 31: Parkplätze und Strand in Scharbeutz: Daily visits nach Besuchergruppen

Die Radwege stellen einen Sonderfall dar. Unter allen Polygonen sind hier die daily visits, die von Einwohnenden generiert werden mit am höchsten, beispielsweise am Standort Drecht (72%). Allerdings liegen auch hier die Fallzahlen am niedrigsten (vgl. Abbildung 32).

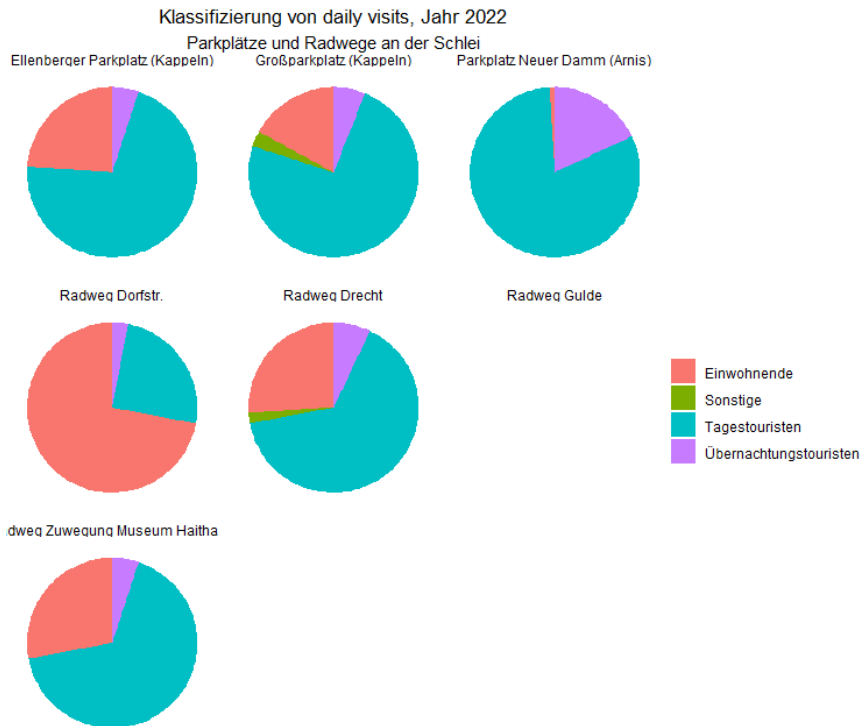


Abbildung 32: Parkplätze und Radwege an der Schlei: Daily visits nach Besuchergruppen

Auch an der Westküste zeigt sich die hohe Bedeutung des Tagestourismus (vgl. Abb. 33). Im Vergleich zu den bisher betrachteten Polygonen an der Ostküste sind die analysierten Polygone hier stärker vom Übernachtungstourismus geprägt. Dies liegt u. a. an der touristischen Attraktivität der Polygone. Während an der Ostsee eher funktionale POI (Parkplätze, Zuwege etc.) untersucht wurden, handelt es sich bei den analysierten Westküsten Hotspots um emotionale POI, an denen Gäste ihren Bedürfnissen nachgehen (Entspannen an der Perlebucht, Shopping am Museumshafen etc.). Auch ist es bemerkenswert, dass auf den Inseln, insbesondere auf Amrum, der Anteil der daily visits, die von Übernachtungstourist*innen generiert werden, deutlich höher liegt (62%) als bei den betrachteten POI auf dem Festland. Aber auch am Strand von Sylt oder St. Peter-Ording liegt der Anteil der Übernachtungstourist*innen bei den daily visits bei knapp 50%. Einwohnende spielen mit Ausnahme von Amrum (23%) bei den ausgewählten Use Cases an der Westküste keine bedeutende Rolle.

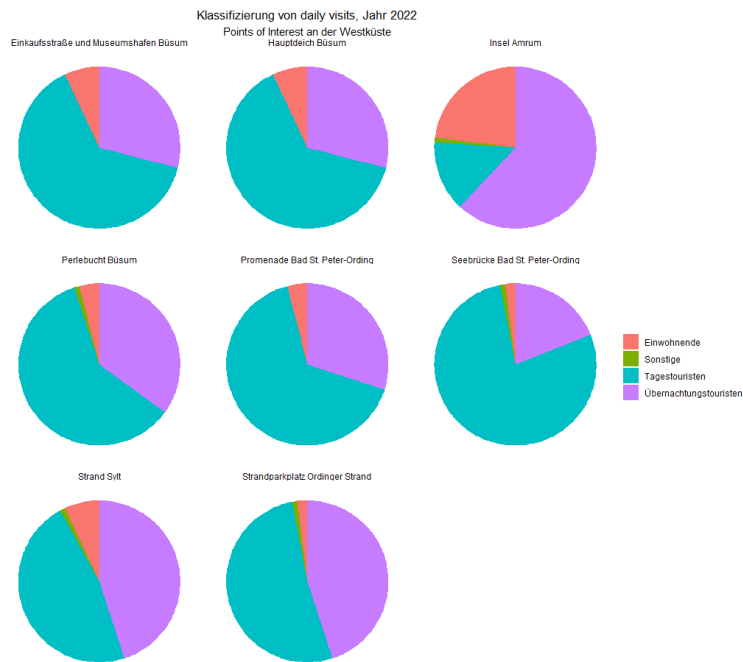


Abbildung 33: Hotspots an der Westküste: Daily visits nach Besuchergruppen

Ein anderes Bild zeichnet sich bei Betrachtung der Wohnmobilstellplätze in der Holsteinischen Schweiz (vgl. Abb. 34). Hier ist der Anteil der Übernachtungstourist*innen in den Fällen der Campingplätze Spitzenort (63%), Prinzenholz (80%) und an der Schwentine (66%) am höchsten. Der Womo-Stop in Malente wird vor allem von Tagestourist*innen und Einwohnenden genutzt. Bei der Analyse der Segmente nach Kalenderwochen (vgl. Abb. 35) wird nochmal besonders die Saisonalität der Übernachtungstourist*innen deutlich – gut zu erkennen beispielsweise bei den Campingplätzen Spitzenort und Naturcamping.

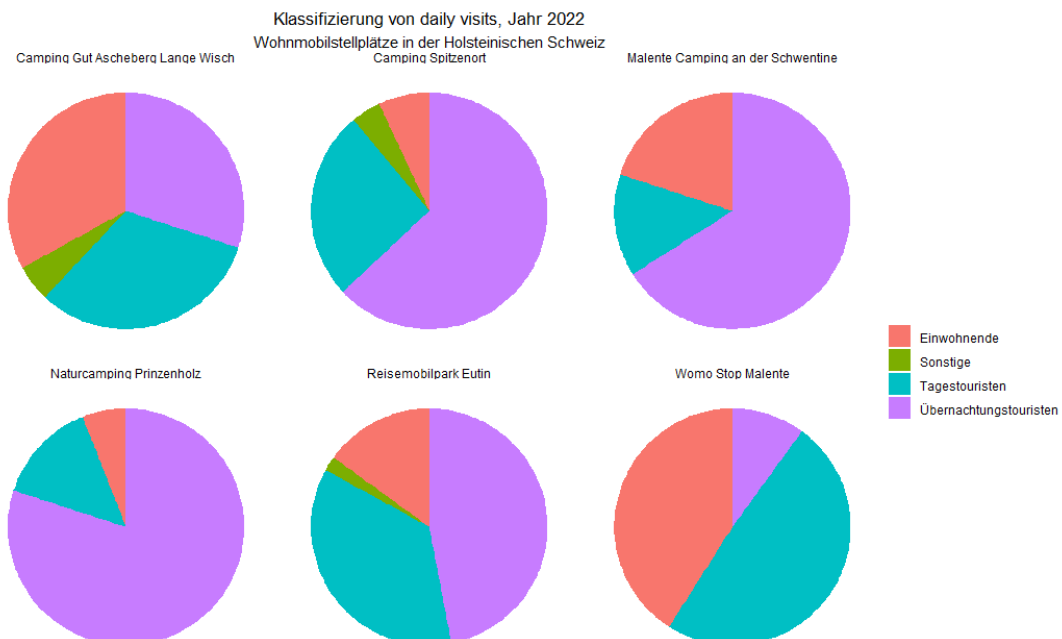


Abbildung 34: Wohnmobilstellplätze und Campingplätze in der Holsteinischen Schweiz: Daily visits nach Besuchergruppen



Abbildung 35: Wohnmobilstellplätze und Campingplätze in der Holsteinischen Schweiz: Daily visits nach Besuchergruppen und Kalenderwochen

Da der Tagestourismus eine bedeutende Rolle bei den Use Cases spielt, soll im Folgenden untersucht werden, woher die Tagestourist*innen kommen und wohin sie gehen (Quell-Zielgebietsbeziehungen). Dafür wird sich dem ersten sowie dem letzten Signal des Endgerätes am Tag des Besuches bedient. Betrachtungszeitraum ist in diesem Falle das gesamte Jahr 2022 auf Basis der daily visits. Dargestellt sind in den nachfolgenden Abbildungen lediglich die häufigsten Netzwerkbeziehungen, da die gesamten Verflechtungen wesentlich komplexer und nicht mehr darstellbar sind. Dabei muss die Definition der Tagestourist*innen aus den Daten beachtet werden. Im Falle des Großparkplatzes in Kappeln (vgl. Abb. 36) handelt es sich dabei um Personen (Endgeräte), die weder ihren gewöhnlichen Aufenthaltsort am Abend (Homezone) noch ihren gewöhnlichen Aufenthaltsort am Tag (Workzone) in Kappeln (oder Arnis) haben, aber in dem Polygon des Parkplatzes sichtbar waren, jedoch nicht an zwei aufeinanderfolgenden Tagen. Die Homezone eines Smartphones berechnet sich jedoch anhand des häufigsten Übernachtungsortes innerhalb der letzten 30 Tage. Daher kann die Homezone auch wechseln, beispielsweise bei längeren Aufenthalten an einem Ort über einen Monat hinaus. Es ist also durchaus möglich, dass es als Tagestourist*innen klassifizierte Signale gibt, die am Anfang und Ende eines Tages das erste und letzte Signal aus Kappeln senden. Das heißt nur, dass sie sich hier aufgehalten (bzw. übernachtet) haben, ihre Homelocation aber z. B. in Hamburg liegt und sie deshalb als Tagestourist*innen klassifiziert sind. Das würde beispielsweise auch auf ankommende bzw. abfahrende Übernachtungstourist*innen in Kappeln zutreffen. Das führt dazu, dass die Quell-Ziel-Gebietsbeziehung für den Ort Kappeln in der Analyse stark vertreten ist. Als Tagestourist*innen klassifizierte Personen haben also das erste Signal des Tages in Kappeln, tauchen dann im Polygon des Parkplatzes auf und senden dann abends wieder aus Kappeln, ihre Homezone liegt dort aber nicht. Neben diesem Umstand zeigt sich generell, dass Quelle und Ziel im Falle des Tagestourismus in der Regel identisch sind. So starten und enden die Besucher*innen des Parkplatzes beispielsweise in Waabs, Damp, Brodersby oder Süderbrarup. Die Hauptquellmärkte der Tagestourist*innen (Homezone) im Jahr 2022, die den Parkplatz besuchten, sind Süderbrarup, Grödersby und Hamburg (vgl. Tab. 9).

Quell-Ziel-Gebietsbeziehungen von Tagestouristen

Großparkplatz Kappeln, Jahr 2022, Daily visits >=20

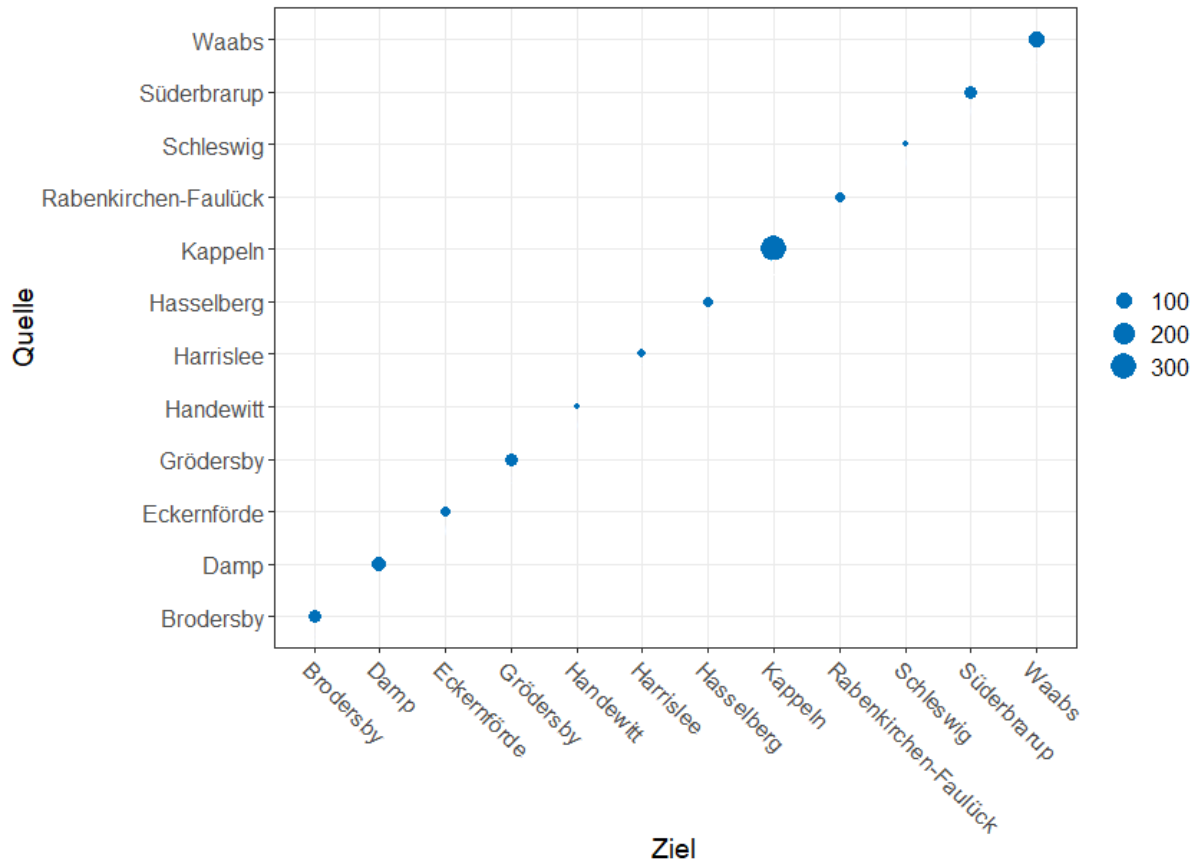


Abbildung 36: Quell-Zielgebietsbeziehungen von Tagestourist*innen am Großparkplatz Kappeln

Ein ähnliches Bild zeichnet sich auch für die Perlebucht in Büsum (vgl. Abb. 37). Hier tritt die Quell-Ziel-Gebietsbeziehung "Büsum" noch deutlicher auf. Weiterhin zeigen sich die ausgeprägten Quell-Ziel-Gebietsbeziehungen an der Westküste, beispielsweise Westerdeichstrich, Heide, St. Peter-Ording, aber auch aus Hamburg. Hamburger*innen starten und enden in der Stadt, es zeigen sich aber auch Quell-Ziel-Gebietsbeziehungen zwischen Hamburg und Büsum und vice versa. Hauptquellmarkt bezogen auf die Homezone im Jahr 2022 war Hamburg.

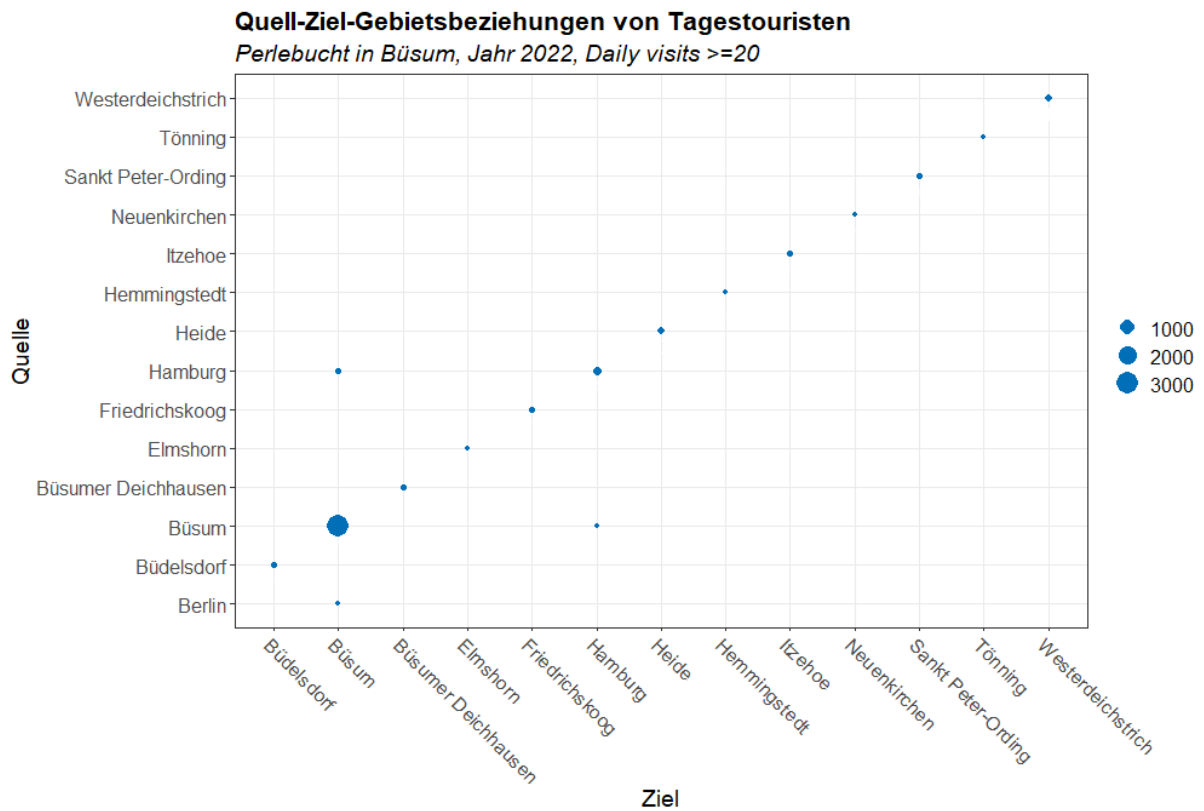


Abbildung 37: Quell-Zielgebietsbeziehungen von Tagestourist*innen an der Perlebucht, Büsum

Da bei den Wohnmobilstellplätzen der Übernachtungstourismus überwiegt, wurde für das Beispiel des Campingplatzes Spitzenort eine Analyse der Quell-Ziel-Gebietsbeziehungen der als Übernachtungstourist*innen klassifizierte Signale vorgenommen (vgl. Abb. 38). Hier werden Pings als Signale von Übernachtungstourist*innen klassifiziert, wenn Sie den gewöhnlichen Wohn- und Arbeitsort nicht in den Gemeinden Eutin oder Malente oder Plön oder Ascheberg haben und an zwei aufeinanderfolgenden Tagen ein Signal von dem Campingplatz gesendet haben. Der Campingplatz liegt in Plön, daher ist es plausibel, dass diese Quell-Ziel-Gebietsbeziehung in den Daten überwiegt. Es lassen sich aber auch anreisende bzw. abfahrende Übernachtungstourist*innen identifizieren, die aus Plön kommend das letzte Signal in Ascheberg senden oder auch aus Hamburg bzw. Ascheberg kommend abends in Plön das letzte Mal in den Daten zu finden sind.

Quell-Ziel-Gebietsbeziehungen von Übernachtungstouristen
Campingplatz Spitzenort, Plön, Jahr 2022, Daily visits >=10

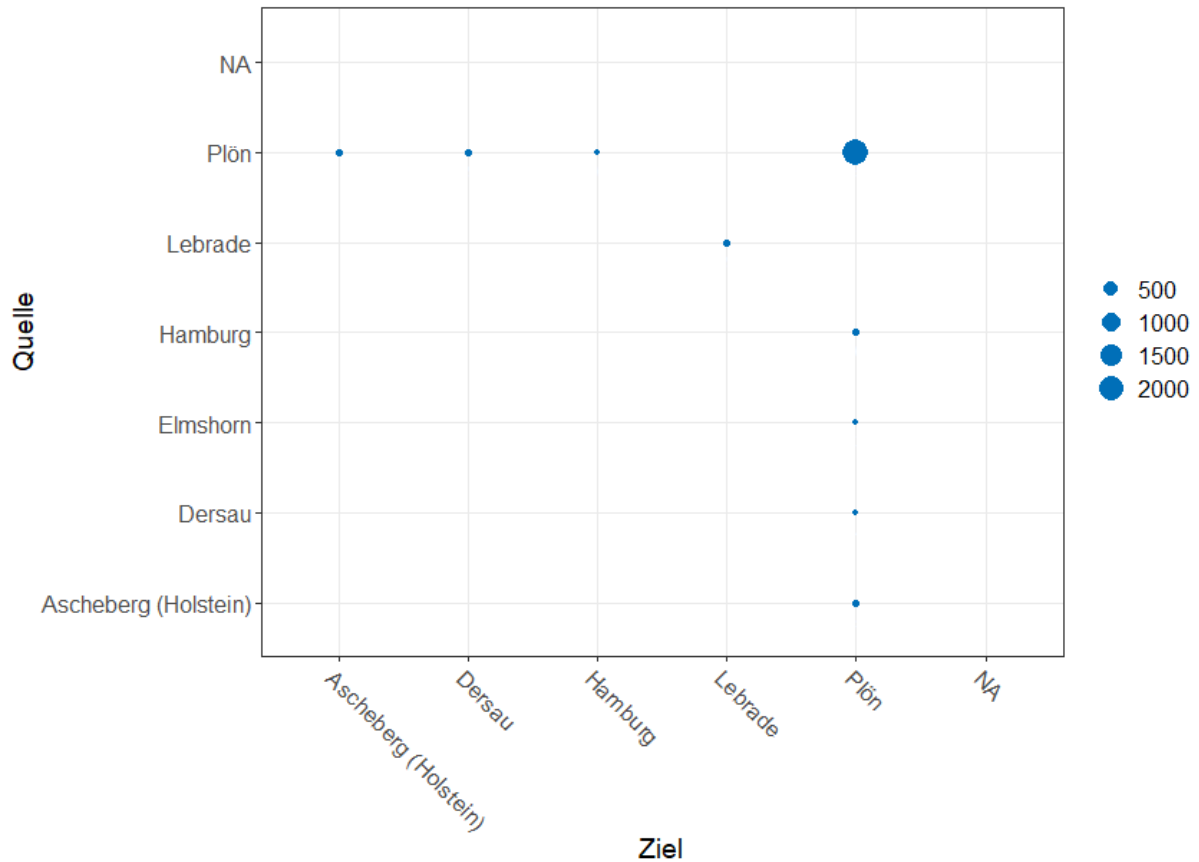


Abbildung 38: Quell-Zielgebietsbeziehungen von Übernachtungstourist*innen auf dem Campingplatz Spitzenort

Da sowohl auf den Campingplätzen und den Parkplätzen die Signaldichte und die Polygongröße relativ gering bzw. klein ist, wird exemplarisch der Strand in Scharbeutz analysiert (vgl. Abb. 39). Neben einer deutlich höheren Anzahl an daily visits ist erkennbar, dass neben Scharbeutz selbst vor allem die großen Quellmärkte Hamburg und Lübeck relevant sind. Auch hier sind Quelle und Ziel in der Regel identisch mit Ausnahmen von Scharbeutz, Lübeck, Hamburg und Timmendorfer Strand. Unabhängig von dem ersten und letzten Signal des Tages sind die Hauptquellmärkte der durch Tagestourist*innen generierten daily visits am Strand in Scharbeutz im Jahr 2022 Hamburg, Lübeck und Bad Schwartau.

Quell-Ziel-Gebietsbeziehungen von Tagestouristen

Strand in Scharbeutz, Jahr 2022, Daily visits >=40

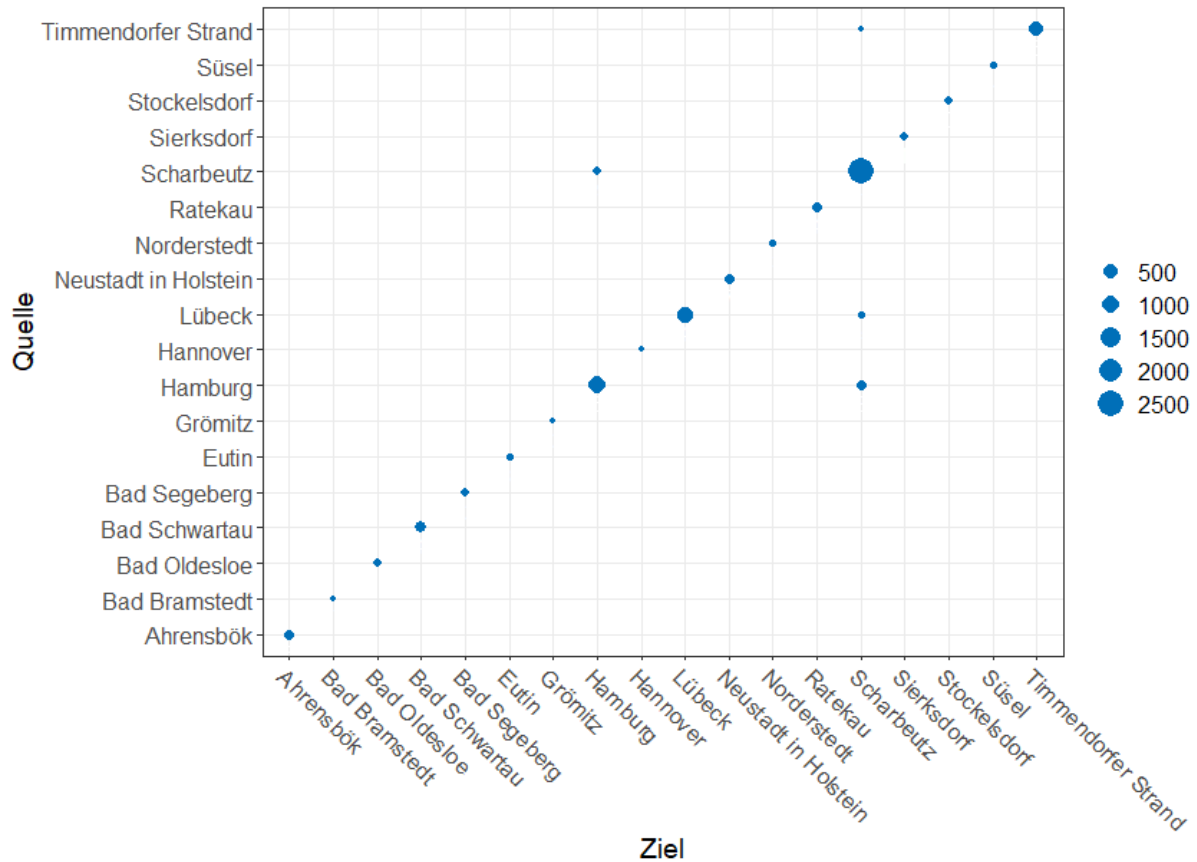


Abbildung 39: Quell-Zielgebietsbeziehungen von Tagestourist*innen am Strand in Scharbeutz

Tabelle 9 gibt abschließend eine Übersicht über die drei wichtigsten Quellmärkte je untersuchten Polygon nach Tages- und Übernachtungstourismus.

Use Case	Name des Zielgebiets (Polygons)	Top-3-Quellmärkte Übernachtungstourist*innen (daily visits)	Top-3-Quellmärkte Tagestourist*innen (daily visits)
Amrum	Insel Amrum	Hamburg Nordfriesland Berlin	Nordfriesland Hamburg Lüneburg
Büsum	Perlebucht	Hamburg Region Hannover Berlin	Dithmarschen Hamburg Pinneberg
Büsum	Hauptdeich	Hamburg Berlin Hildesheim	Hamburg Dithmarschen Pinneberg
Büsum	Einkaufsstraße und Museumshafen	Hamburg Dithmarschen Berlin	Dithmarschen Hamburg Pinneberg
Flensburg	Einkaufsstraße Holm	Schleswig-Flensburg Flensburg Rendsburg Eckernförde	Schleswig-Flensburg Nordfriesland Flensburg
Hotspots Schlei	Großparkplatz Kappeln	Schleswig-Flensburg Rendsburg Eckernförde	Schleswig-Flensburg Rendsburg Eckernförde

		Mettmann	Hamburg
Hotspots Schlei	Arnis Parkplatz Neuer Damm	Schleswig-Flensburg Bremerhaven Nordfriesland	Schleswig-Flensburg Hamburg Rendsburg Eckernförde
Hotspots Schlei	Kappeln Ellenberger Parkplatz	Steinfurt Weimar Berlin	Schleswig-Flensburg Hamburg Rendsburg Eckernförde
Kiel	Kieler Woche	Rendsburg Eckernförde Plön Kiel	Rendsburg Eckernförde Plön Hamburg
Lübecker Bucht	Parkplatz Scharbeutz P3	Stormann Ostholstein Hamburg	Ostholstein Hamburg Segeberg
Lübecker Bucht	Parkplatz Scharbeutz P2	Stormann Ostholstein Hamburg	Ostholstein Hamburg Segeberg
Lübecker Bucht	Lübecker Bucht Strand gesamt	Hamburg Ostholstein Region Hannover	Hamburg Ostholstein Lübeck
Lübecker Bucht	Parkplatz Scharbeutz Badweg	Bochum Würzburg Hamburg	Ostholstein Hamburg Lübeck
Lübecker Bucht	Parkplatz Scharbeutz P4	Ostholstein Segeberg Herzogtum Lauenburg	Ostholstein Lübeck Hamburg
Parkplätze Ostsee	Parkplatz Burgtiefe	Ostholstein Hamburg Berlin	Hamburg Ostholstein Region Hannover
Parkplätze Ostsee	Parkplatz Südstrand Mitte	Hamburg Herzogtum Lauenburg Magdeburg	Hamburg Ostholstein Kassel
Parkplätze Ostsee	Parkplatz Meeschendorf	Pinneberg Ostholstein Hamburg	Segeberg Recklinghausen Region Hannover
Parkplätze Ostsee	Parkplatz Gold	Hamburg Region Hannover Ostholstein	Hamburg Ostholstein Kiel
Parkplätze Ostsee	Parkplatz Orth	Konstanz Köln Stormann	Göttingen Hamburg Recklinghausen
Parkplätze Ostsee	Parkplatz Grüner Brink	Lippe Ostholstein Stormann	Hamburg Peine Stormann
Parkplätze Ostsee	Parkplatz Königsredder Straße	Ostholstein Bremen Hamburg	Ostholstein Hamburg Stormann
Parkplätze Ostsee	Parkplatz Seebrückenvorplatz Seestr. 31	Höxter Minden-Lübbecke Recklinghausen	Ostholstein Hamburg Region Hannover
Parkplätze Ostsee	Parkplatz Deichweg	Berlin Soest Vorpommern-Rügen	Hamburg Ostholstein Berlin
Parkplätze Ostsee	Großraumparkplatz Grömitz	Ostholstein Region Hannover Hamburg	Ostholstein Hamburg Region Hannover
Radwege Schlei	Standort 23 Reitanlage Dorfstr Radweg	Gelsenkirchen - -	Rendsburg Eckernförde Pinneberg Gelsenkirchen
Radwege Schlei	Standort 22 Gulde	-	-
Radwege Schlei	Standort 21 Drecht	Segeberg Schleswig-Flensburg Stormann	Pinneberg Schaumburg Schleswig-Flensburg

Radwege Schlei	Standort 17 Radweg 1 Zuwegung Museum Haithabu	Plön - -	Plön Nordfriesland Hamburg
St. Peter-Ording	Promenade Bad	Hamburg Region Hannover Rendsburg-Eckernförde	Hamburg Pinneberg Nordfriesland
St. Peter-Ording	Seebrücke Bad	Hamburg Region Hannover Darmstadt-Dieburg	Hamburg Pinneberg Nordfriesland
St. Peter-Ording	Strandparkplatz Ordinger Strand	Nordfriesland Region Hannover Hamburg	Nordfriesland Hamburg Pinneberg
Sylt	Strand Sylt	Hamburg Berlin Region Hannover	Hamburg Berlin Region Hannover
Wohnmobilstellplätze Holsteinische Schweiz	Reisemobilpark Eutin	Pinneberg Ostholstein Hamburg	Ostholstein Plön Segeberg
Wohnmobilstellplätze Holsteinische Schweiz	Naturcamping Prinzenholz	Peine Hamburg Gelsenkirchen	Ostholstein Plön Schleswig-Flensburg
Wohnmobilstellplätze Holsteinische Schweiz	Malente Camping an der Schwentine	Pinneberg Gütersloh Hamburg	Ostholstein Plön Eckernförde
Wohnmobilstellplätze Holsteinische Schweiz	Womo Stop Malente	Stormann Soest Schleswig-Flensburg	Ostholstein Plön Dithmarschen
Wohnmobilstellplätze Holsteinische Schweiz	Camping Spitzenort	Pinneberg Segeberg Stormann	Plön Segeberg Neumünster

Tabelle 9: Top 3-Quellmärkte der als Tages- und Übernachtungstourist*innen klassifizierten Daily visits je Untersuchungsregion im Jahr 2022

3.3 Ausgewählte Vergleiche von lokaler und globaler Sensorik

Für ausgewählte Standorte lassen sich die Frequenzen aus der lokalen Sensorik mit den daily visits vergleichen. Dies ist nicht für alle Standorte der Use Cases möglich, da Installationen erst ab August 2022 durchgeführt werden konnten und die globale Sensorik bislang erst für das Jahr 2022 vorliegt. Daher können lediglich die Kalenderwochen 31-52 im Jahr 2022 (01.08. – 31.12.2022) analysiert werden. Dabei wurden die summierten Messwerte für Einfahrten auf den Parkplätzen der entsprechenden Summe der daily visits gegenübergestellt. Tabelle 10 zeigt die Korrelationsanalyse der jeweiligen Messpunkte. Während bei einigen Parkplätzen moderate Korrelationen vorliegen (Burgtiefe, Südstrand) zeigt sich beim Parkplatz Meeschendorf eine stärkere Korrelation ($r=0,57$; $R^2=33\%$) (vgl. Abb. 40). Gleichwohl zeigen sich auch gering positive Korrelationen oder geringe negative Korrelationen. In diesen Fällen spielt es keine Rolle, ob die Einfahrten auf den Parkplätzen zunimmt oder abnimmt. Es hat keine Auswirkungen auf die Anzahl der daily visits oder umgekehrt. Der nicht vorhandene lineare Zusammenhang zeigt sich auch in einem am Beispiel des Großraumparkplatzes in Grömitz geringen Bestimmtheitsmaß von 0,4%.

Betrachtungsgebiet	Pearson Korrelation [®]	Bestimmtheitsmaß (R ²), gerundet
Parkplatz Burgtiefe Fehmarn	0,44	0,19
Parkplatz Meeschendorf Fehmarn	0,57	0,32
Parkplatz Südstrand Fehmarn	0,45	0,20
Parkplatz Brink Fehmarn	0,04	0,00
Parkplatz Orth Fehmarn	0,33	0,11
Parkplatz Gold Fehmarn	0,15	0,02
Parkplatz Königsredder Straße Fehmarn	-0,04	0,00
Großraumparkplatz Grömitz Fehmarn	-0,06	0,00

Tabelle 10: Pearson Korrelation und Bestimmtheitsmaß nach Betrachtungsgebiet

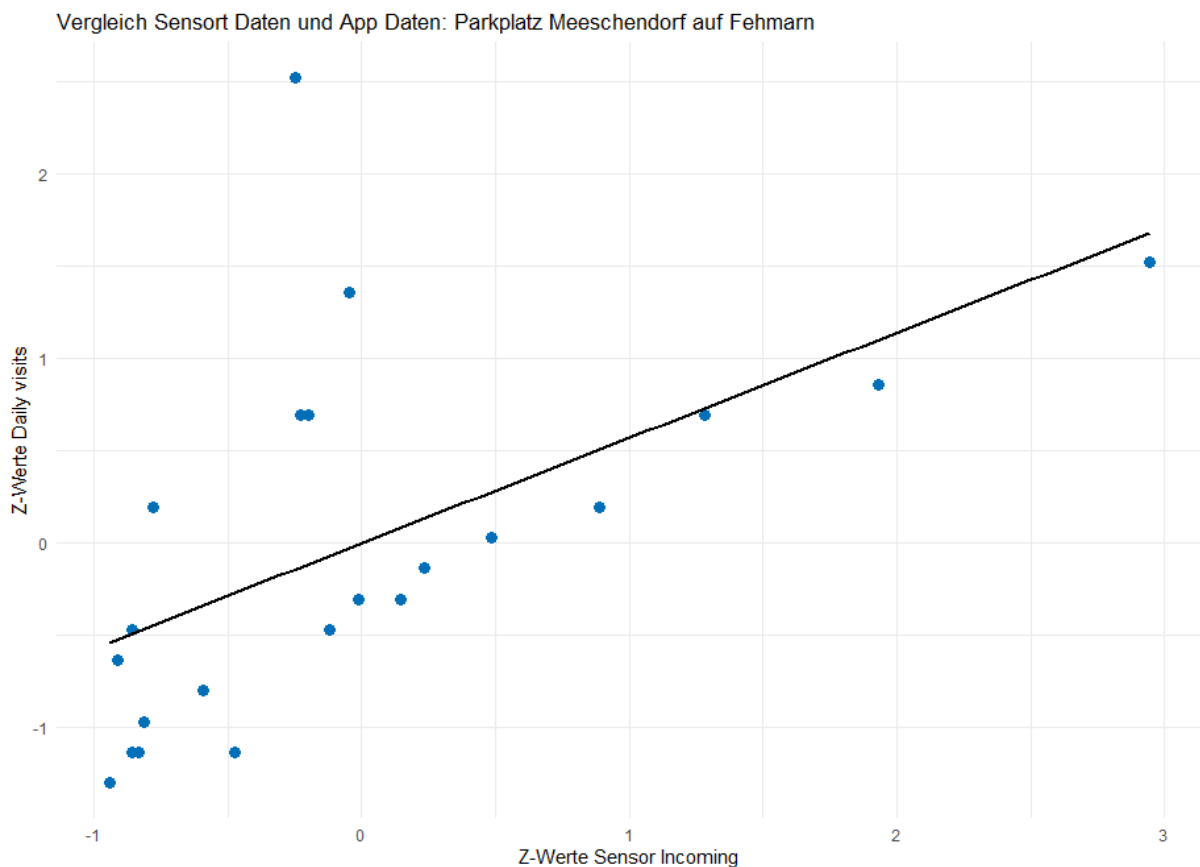


Abbildung 40: Z-Werte von Daily Visits und Incoming-Messwerten Parkplatz Meeschendorf, Fehmarn

Bei der Nutzung von Sekundärdaten, beispielsweise aus der Perlebucht in Büsum oder auch bei der Nutzung von Daten aus Scharbeutz, besteht jedoch die Möglichkeit der Analyse des gesamten Kalenderjahres 2022. Neben einer erhöhten Stichprobe von 52 Kalenderwochen gibt es hier zu bedenken, dass das Betrachtungsgebiet (Polygon) ein größeres Areal abdeckt (siehe **Anhang**) sowie die genutzte

lokale Sensorik eine andere ist. Im Falle vom Strand in Scharbeutz (vgl. Abb. 41) werden die Strandzugänge mit Lasersensorik gemessen. Ein ähnliches Verfahren wird auch in der Perlebucht in Büsum genutzt (vgl. Abb. 42). Hohe Korrelationen über den Jahresverlauf der beiden Messreihen zeigen sich sowohl in der Perlebucht in Büsum ($r=0,82$) als auch am Strand in Scharbeutz ($r=0,78$). Es besteht demnach ein linearer Zusammenhang zwischen den Messreihen, der sich in einem relativ hohen Bestimmtheitsmaß widerspiegelt. Dieses Maß gibt an, welcher Anteil der Varianz in den beobachteten Daten vorhergesagt werden kann. In Scharbeutz liegt dieser Wert bei 61%, in Büsum bei 67%. Dies bedeutet, dass in Büsum 67% der Varianz der Incoming-Messwerte durch die Location Events erklärt werden können und umgekehrt. Die restlichen 33% könnten durch andere Faktoren, wie beispielsweise Events oder Ferientage beeinflusst werden. Aufgrund des deutlichen Volumenunterschieds zwischen den Messwerten der daily visits und den gemessenen Eingängen wurden die Daten z-transformiert. Dies ermöglicht die Visualisierung eines Jahresverlaufs, jedoch auf Kosten der Sichtbarkeit von Nachfragespitzen in den Daten

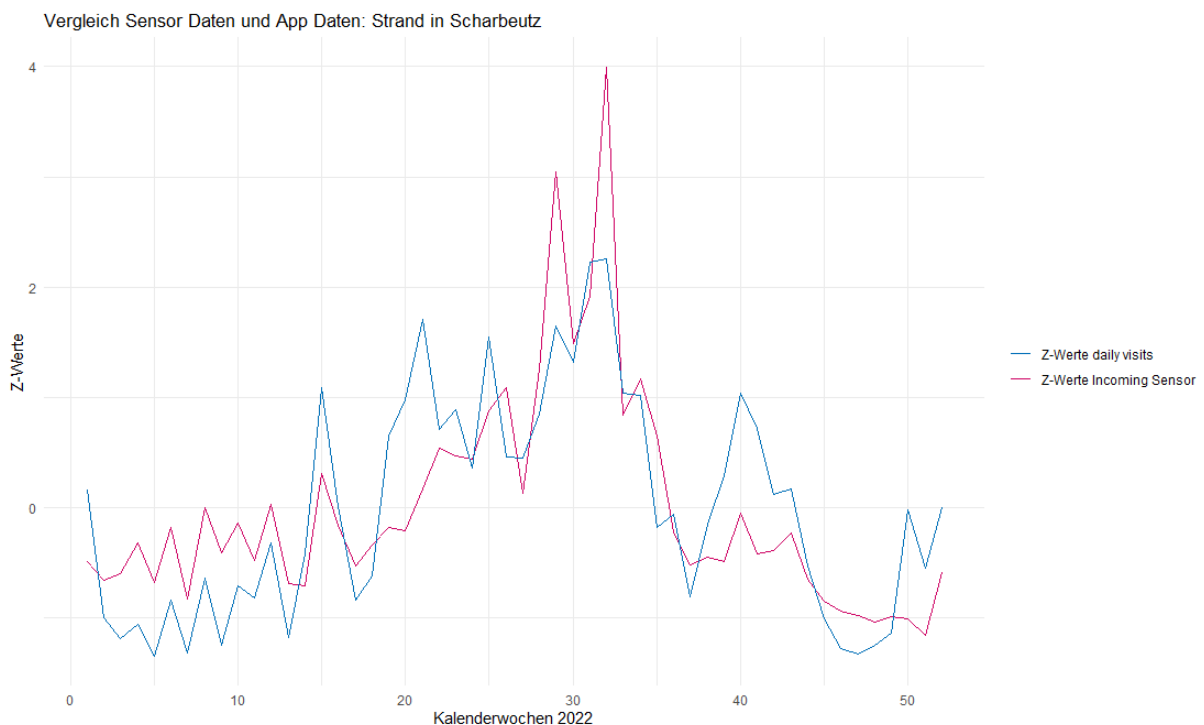


Abbildung 41: Z-Werte von Daily visits und Incoming-Messwerten am Strand in Scharbeutz

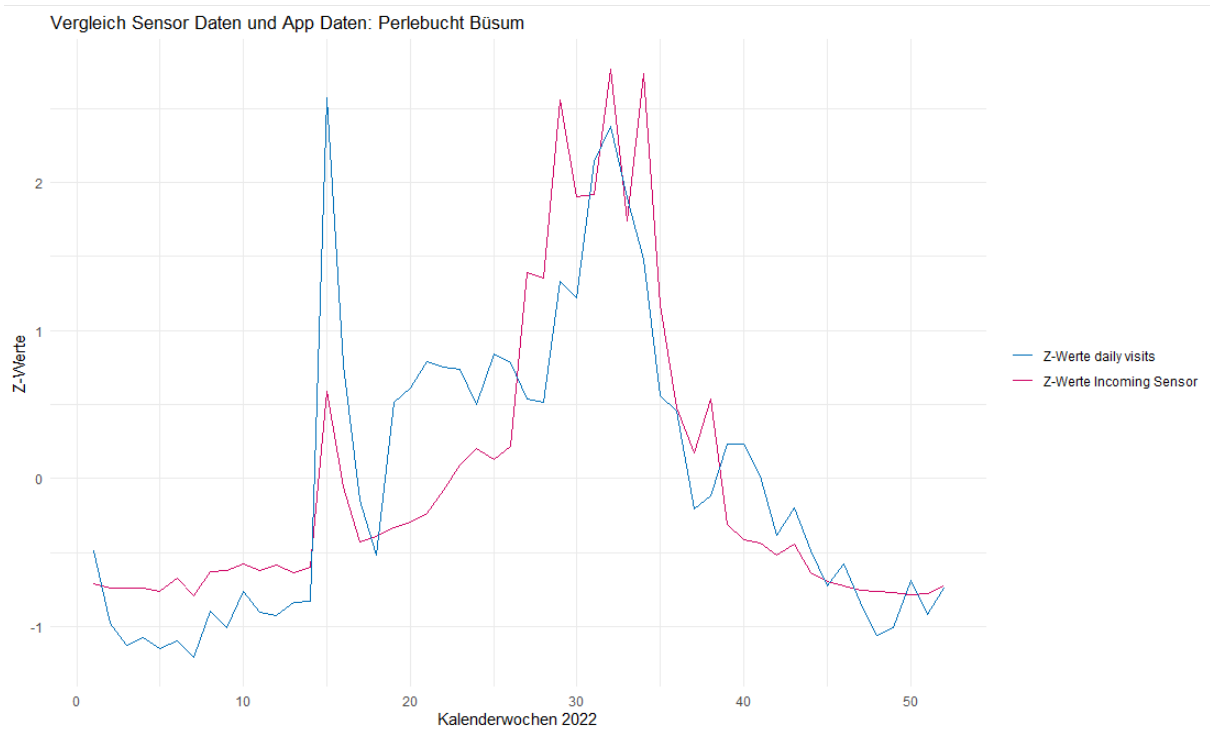


Abbildung 42: Z-Werte von Daily visits und Incoming-Messwerten in der Perlebucht Büsum

Sehr gut funktioniert der Vergleich auch mit Daten von der Seebrücke in St. Peter-Ording (vgl. Abb. 43). Hier wiederum liegen aufgrund von Sensorausfällen nur Daten zum Vergleich des zweiten Halbjahres 2021 auf Basis der Kalenderwochen vor. Auch hier zeigt sich eine hohe Korrelation der beiden Messreihen ($r=0,92$; $R^2=85\%$).

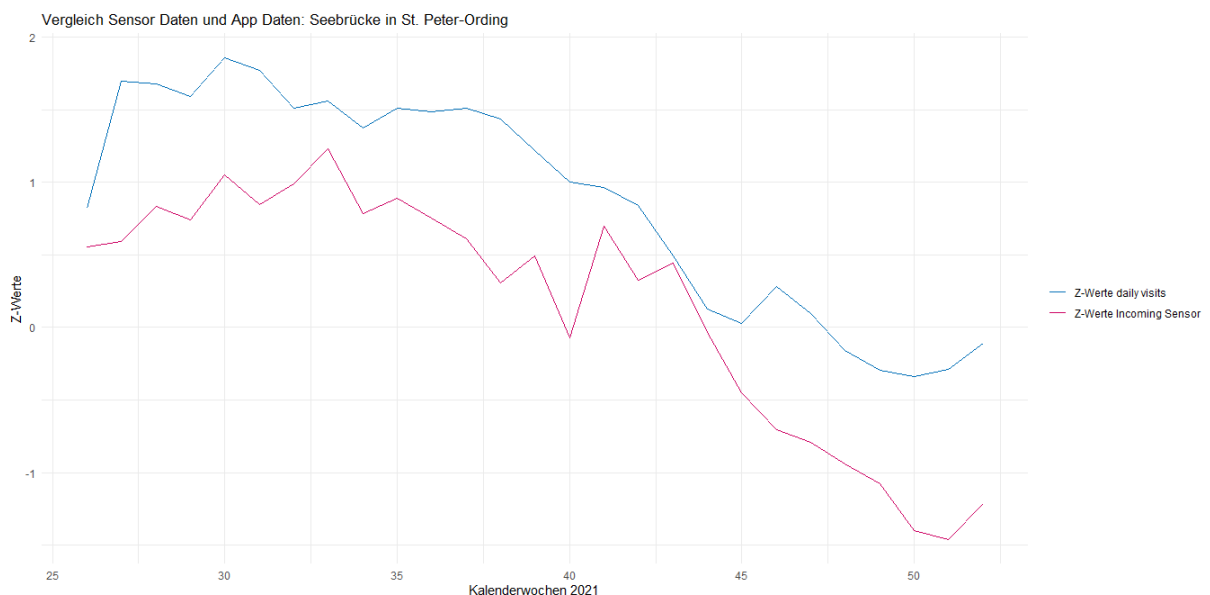


Abbildung 43: Z-Werte von Daily visits und Incoming-Messwerten an der Seebrücke in St. Peter-Ording

Ebenfalls sehr gut funktioniert der Abgleich zwischen den daily visits auf Amrum und den am Fährlager auf Amrum installierten Wifi-Zählgerät (Destination One) (vgl. Abb. 44 & 45). Hier liegen für die

Kalenderwochen 20-52 Daten vor, die den daily visits gegenübergestellt werden können. Bemerkenswert ist hier, dass die Volumenunterschiede im Vergleich zu den bisherigen Vergleichen nicht sehr hoch ausfallen. Das Volumen der Wifi-Signale von Endgeräten ist dennoch dreimal höher als das der daily visits. Auffällig sind bei den Wifi-Signalen die Einbrüche in KW 29 sowie KW 49, wobei letztere mit fehlenden Daten für drei Tage in dieser Woche zu erklären ist. Auch diese Zeitreihe korreliert sehr stark ($r=0,89$; $R^2=80\%$).

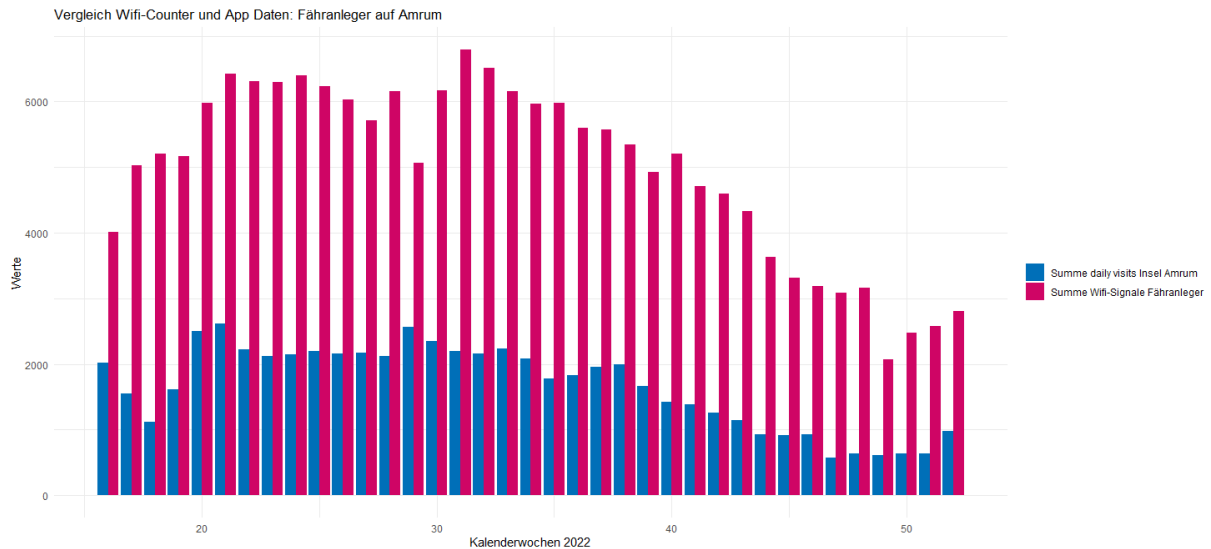


Abbildung 44: Volumen von daily visits (Insel Amrum) und Wifi-Signale (Fähranleger Insel Amrum)

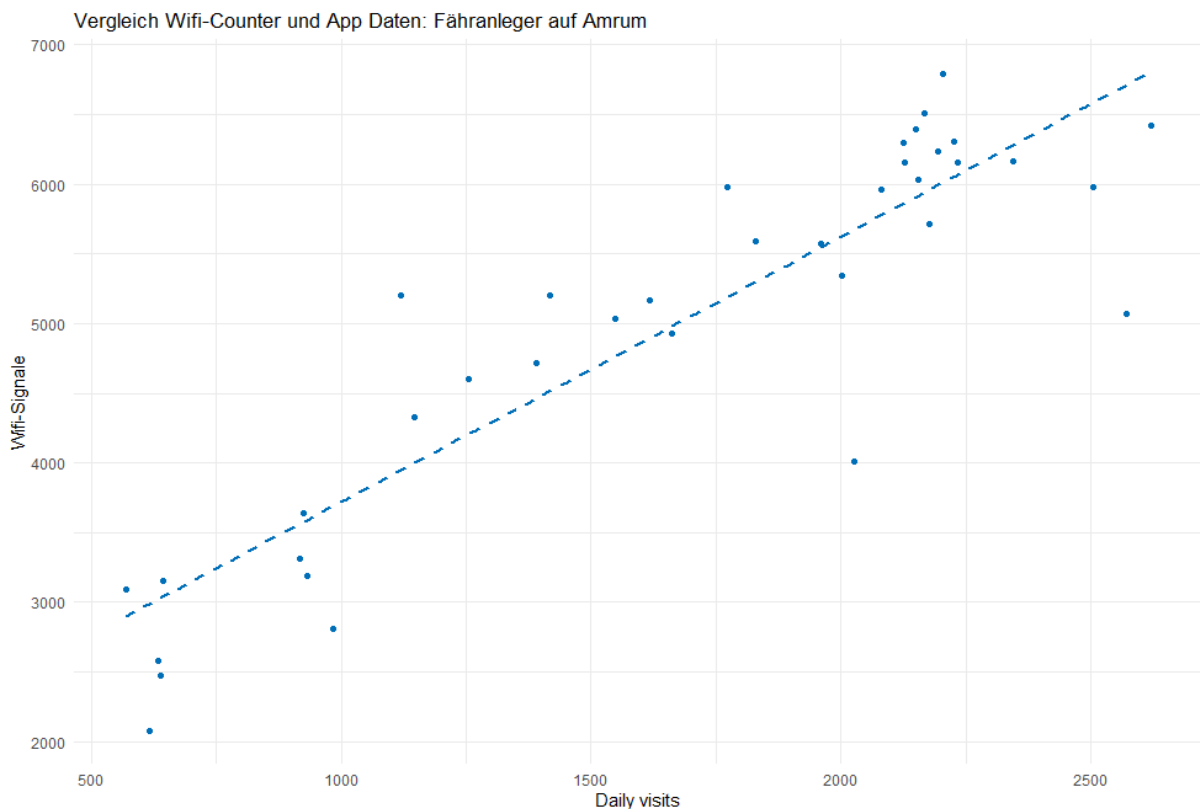


Abbildung 45: Korrelation von daily visits (Insel Amrum) und Wifi-Signale (Fähranleger Insel Amrum)

Um die Aussagekraft derartiger Vergleiche zu erhöhen, wäre es bei zukünftigen Projekten wichtig, dass die location events in einer täglichen Auflösung vorliegen. Die Ergebnisse der hier durchgeführten Analysen und der Abgleich von Big Data mit Referenzdaten zeigt sich auch bei der Nutzung von anderen Datenquellen, beispielsweise bei Daten aus App Events der App Wetter.com oder bei Mobilfunkdaten (Schmücker & Reif, 2022b).

Aber auch bei der Nutzung von Messreihen über ein Jahr und anderer Sensorik zeigen sich bisweilen keine guten Passungen zwischen App-Daten und lokaler Sensorik. Beim Vergleich der daily visits mit Hilfe von Laser-Sensoren generierten Passantenfrequenzen in der Einkaufsstraße Holm in Flensburg zeigt sich – ähnlich wie bei den Parkplätzen auf Fehmarn – eine geringe Korrelation von $r=0,11$ ($R^2=1,3\%$) (vgl. Abb. 46). Auffallend sind hier bei den App-Daten ein Einbruch in Kalenderwoche 18 sowie bei Hystreet in Kalenderwoche 51. Auch hier ist das Niveau der mittels Laser gemessenen Passanten um das 384 (!) fache größer als das der App-Daten. Zwar muss berücksichtigt werden, dass Passanten mehrfach gezählt werden, darunter beispielsweise auch Kinder, die kein Smartphone besitzen, jedoch sollte die Nutzung von Smartphone Applikationen in der Innenstadt sehr hoch liegen. Allerdings liegen in Summe für das Kalenderjahr 2022 nur 21.406 daily visits vor.

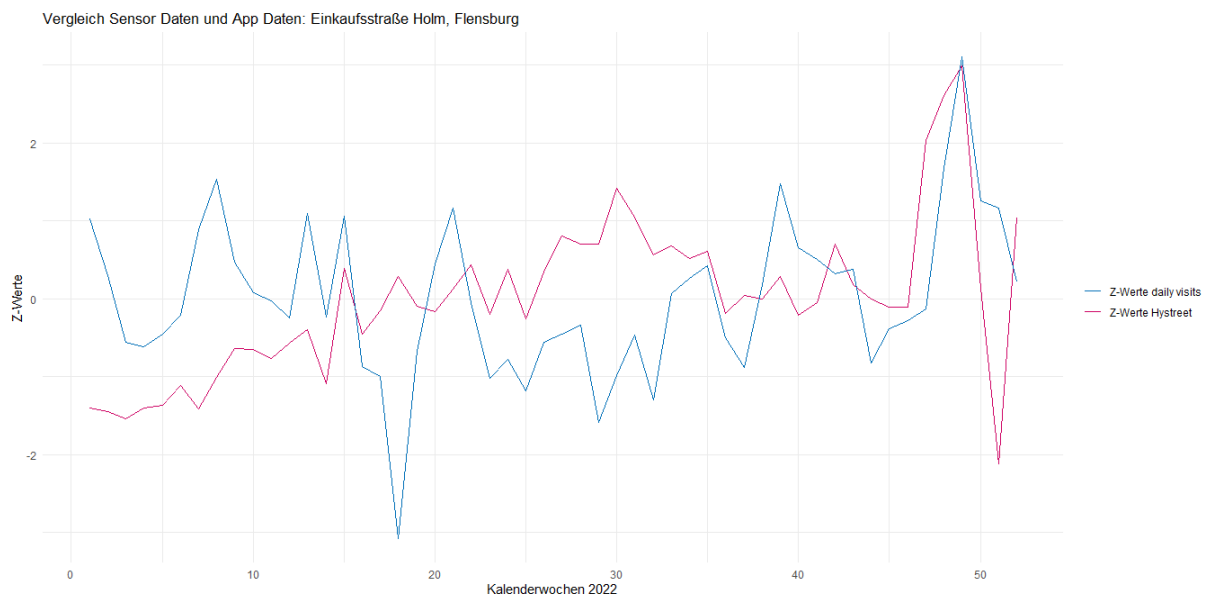


Abbildung 46: Z-Werte von Daily Visits und Incoming-Messwerten in der Einkaufsstraße Holm, Flensburg

Datenquelle Hystreet: www.hystreet.com

Zusätzlich wurden die Daten am Beispiel der Insel Amrum mit Referenzdaten der Übernachtungen aus der Kurabgabestatistik gegenübergestellt. Hierfür wurde die Hochsaison gewählt und für die Kalenderwochen 23-34 im Jahr 2022 die daily visits der Übernachtungstourist*innen den Übernachtungen laut Meldescheine gegenübergestellt. Die Messreihen korrelieren jedoch nur schwach miteinander ($r=0,26$).

3.4 Weitere Analysen

Im Rahmen des Projektes wurde auch überprüft, welchen Einfluss das Wetter und die Ferientagssumme auf die Einfahrten hat. Um diese Analyse durchzuführen liegen idealerweise so viele Datenpunkte wie möglich vor. Die Analyse erfolgt für den Parkplatz Burgtiefe auf Fehmarn. Da hier die längste Datenreihe vorliegt und der Parkplatz in unmittelbarer Nähe zum Strand liegt und davon ausgegangen werden kann, dass hier die Effekte des Wetters eine gewisse Rolle spielen, wurde dieser Parkplatz ausgewählt. Die Wetterdaten stammen aus der API von openweathermap; die Informationen zu den Ferien wurden auf Basis von Schmücker 2023 zusammengetragen. In Abbildung 47 repräsentiert ein Punkt die Summe der täglichen Einfahrten auf dem Parkplatz. Die Ferientagssumme ist die Anzahl der Bundesländer mit einem Ferientag in Deutschland an dem jeweiligen Tag und kann von 0-16 reichen. Wie schon oben in Abbildung 9 zu erkennen, geht die Summe der Einfahrten in den Wintermonaten stark zurück, um dann ab April wieder anzusteigen. Gleichzeitig ist erkennbar, dass die Nachfragespitzen i.d.R. an den Tagen sind, an denen auch viele Bundesländer Ferien haben (dunkelblaue Punkte). Bei der schwarzen Linie handelt es sich um einen Smoother (Loess-Methode: Locally Weighted Scatter-plot), der die Datenreihe glättet, um so den Verlauf besser zu visualisieren. Die Sensormesswerte korrelieren dementsprechend positiv mit der Ferientagssumme, wenn auch nur moderat ($r=0,51$). Ähnliche hohe Korrelationen der Einfahrten mit der Ferientagssumme zeigen sich beim Parkplatz Seestraße in Grömitz ($r=0,57$).

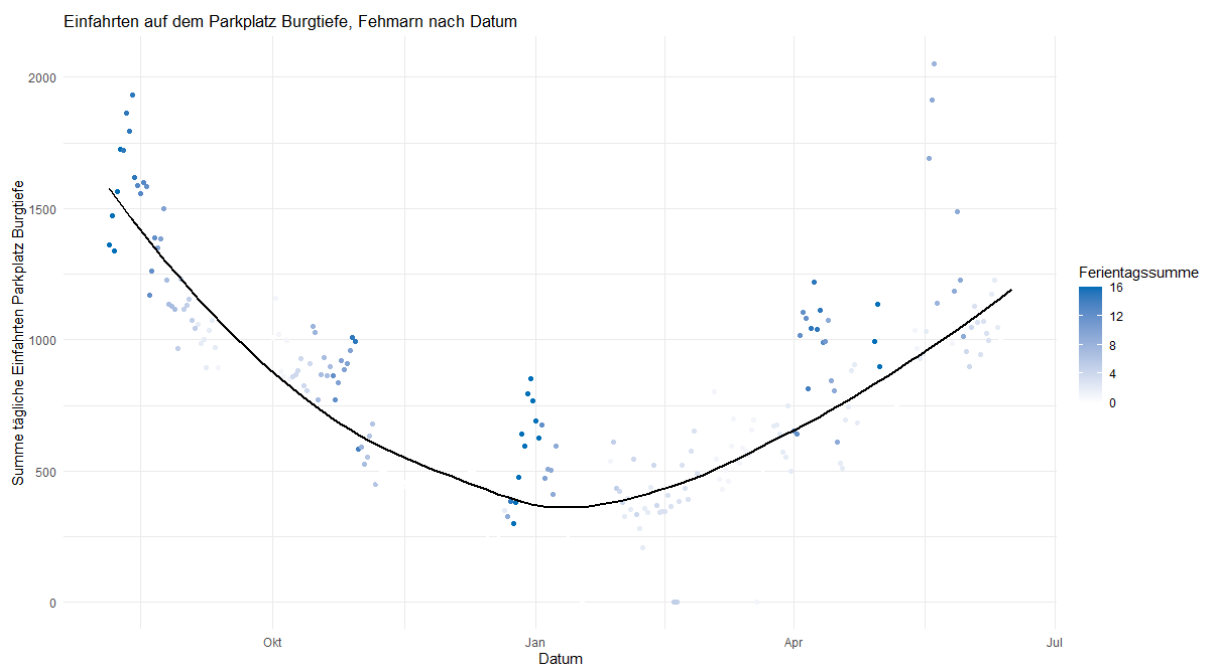


Abbildung 47: Einfahrten Parkplatz Burgtiefe nach Datum

Abbildung 48 zeigt nun wiederum die Einfahrten je Tag nach der Tagestemperatur. Eingefärbt ist die Anzahl der Bundesländer mit einem Ferientag. Die Summe der täglichen Einfahrten korreliert stark positiv mit der Tagestemperatur ($r=0,77$). Mit ansteigender Temperatur steigen also auch die Einfahrten auf dem Parkplatz Burgtiefe. Bei einer einfachen linearen Regression lassen sich demnach 60% der Varianz der Daten in den Einfahrten auf dem Parkplatz mit der Temperatur erklären. Ähnliche Korrelationen zeigen sich auch für den Parkplatz Seestr. In Grömitz ($r=0,63$).

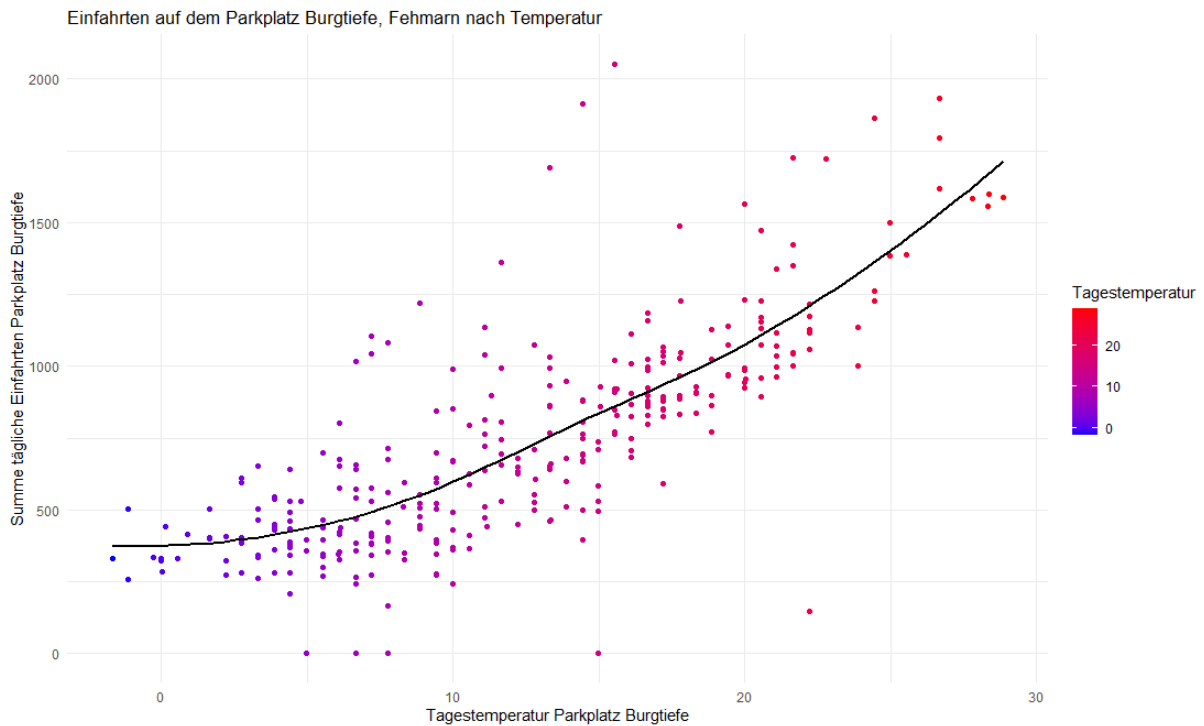


Abbildung 48: Einfahrten Parkplatz Burgtiefe nach Temperatur

Im Projekt wurde ebenfalls eine Analyse am Beispiel des Polygons des Veranstaltungsortes der Kieler Woche (siehe Anhang) vorgenommen. Dabei stand die Frage im Vordergrund, welchen Einfluss die Kieler Woche auf die Besucherzahl in Kiel hatte. Die Kieler Woche im Jahr 2022 fand im Zeitraum vom 18.6. – 26.6.2022 statt. Es wurden in diesem Falle Daten vom Datenlieferant near⁶, von Thaltegos (siehe Abschnitt 3.1.2) und vom #sh_wlan (siehe Abschnitt 3.1.3) genutzt. Bei den Daten von near handelt es sich ebenfalls um Daten aus globaler Sensorik, die nach dem identischen Prinzip erhoben werden, wie in Abschnitt 3.1.2 beschrieben. Die Daten liegen für das identische Polygon und Zeitraum vor und sind direkt mit den Daten von Thaltegos vergleichbar. Für die Analyse in Abbildung 49 wurden die drei Kennzahlen der anonymisierten Device Ids (near_id) (einzigartige Endgeräte/Besucher*innen), der Anzahl der clients (sh_clients) (einzigartige WLAN-Endgeräte/Besucher*innen) und der Anzahl der visitors (thaltegos_visitors) (einzigartige Endgeräte/Besucher*innen) gegenübergestellt. Dabei handelt es sich im Wesentlichen also um die aggregierte Anzahl der einzigartigen Wochenbesucher*innen pro in dem Polygon. Der Effekt der Kieler Woche auf die Anzahl der Besucher*innen lässt sich bei der Anzahl der visitors bei thaltegos und bei der Anzahl der #sh_wlan clients erkennen. Bei den Daten von near hingegen zeigt sich keine steigende Besucherzahl während des Events. Aber auch bei den Daten von thaltegos sind die Wochen nach der Kieler Woche auf einem ähnlich hohen Niveau. Bei der Interpretation ist zu berücksichtigen, dass beim #sh_wlan alle Access-Points genommen wurden, die sich in dem Polygon befinden, während bei den mobile location events eine flächendeckende Betrachtung des Polygons möglich ist. Dabei zeigt sich beim #sh_wlan die Bedeutung der Access-Points an der Kiel-Linie (vgl. Abb. 50), da sich hier die meisten Besucher*innen aufgehalten haben. Weitere Analysen, beispielsweise mit Hilfe von Daten von Hystreet (siehe Abschnitt 3.1.4) und auf Tagesbasis sind nötig,

⁶ www.near.com; Die Daten wurden im Rahmen des Projektes zu Testzwecken kostenfrei zur Verfügung gestellt. Derzeit ist eine weitergehende Analyse dieser Daten im Rahmen einer wissenschaftlichen Veröffentlichung im Gange.

um die räumlichen Auswirkungen der Kieler Woche zu verstehen. Eine Hypothese ist, dass die Kieler Woche viele Menschen bewegt, jedoch nur auf einem kleinen räumlichen Gebiet.

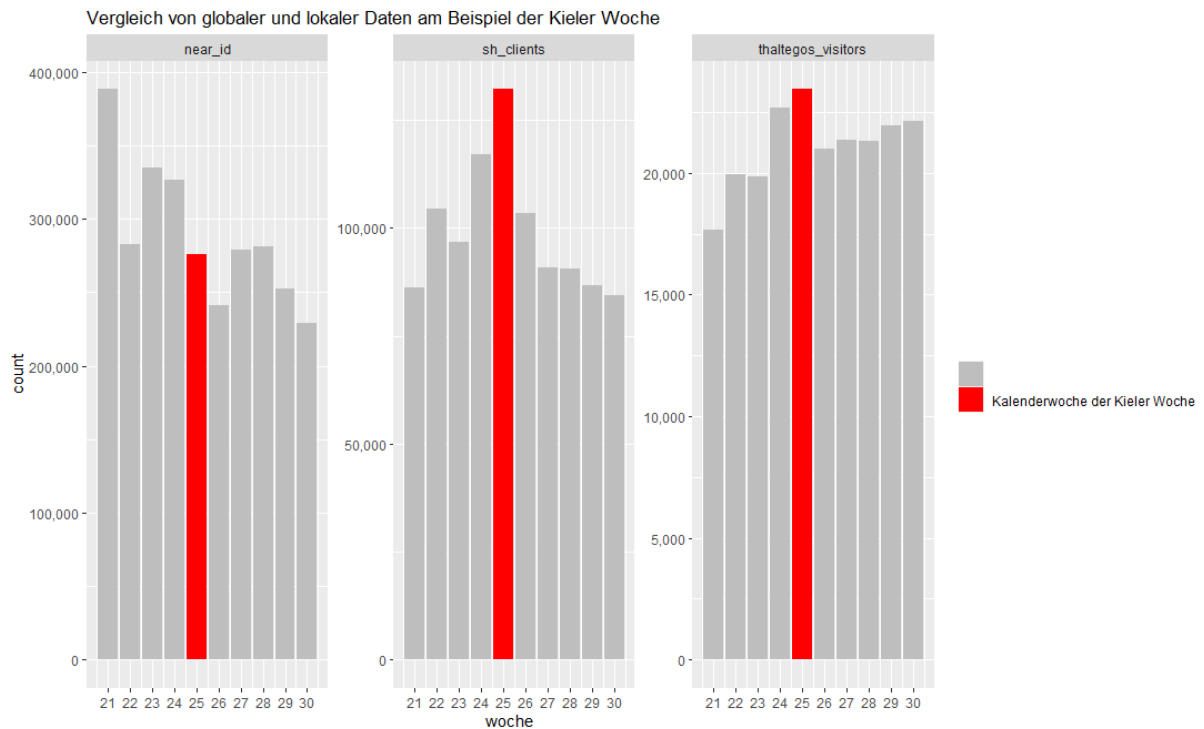


Abbildung 49: Anzahl Device Ids, #sh_wlan clients und visitors in Kiel vor, während und nach der Kieler Woche

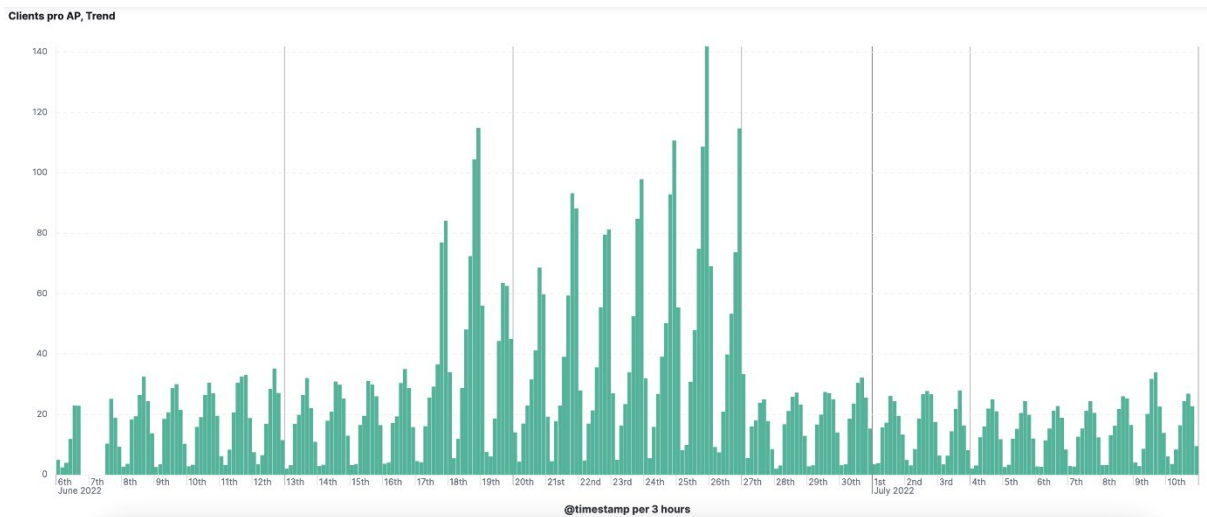


Abbildung 50: Anzahl unterschiedlicher Clients pro Access Point für einige ausgewählte APs an der Kiellinie vor, während und nach der Kieler Woche 2022

3.5 Zwischenfazit

Mit der in Arbeitspaket 2 installierten Sensorik sowie den weiteren Datenquellen, die beschafft oder hinzugezogen wurden, wurde die Grundlage für die Prognose geschaffen. Gleichzeitig erlauben die Daten ein besseres Verständnis für Besucherströme in den Use Cases. Für die jeweiligen Use Cases

erfolgt daher eine kurze inhaltliche Zusammenfassung wesentlicher Erkenntnisse aus der Analyse der lokalen und globalen Sensorik:

- **Parkplätze an der Ostsee:** Die Mehrheit der Standorte weist das gleiche Muster auf, in den Wintermonaten gibt es geringe Frequenzen, ab April steigen die Frequenzen und in den Sommermonaten kommt es schließlich zu Hochfrequenzen. Es sind zudem verschiedene Einflüsse erkennbar. Zum einen scheint der Wochentag einen Einfluss auf die Frequenzen zu haben, da bei allen Parkplätzen auf Fehmarn der Samstag am stärksten frequentiert ist. In Grömitz sind neben Samstag ebenfalls Freitag und Sonntag Spitzentage einiger Parkplätze. Zudem ist an Feiertagen, beispielsweise Ostern und Pfingsten, eine erhöhte Frequenz erkennbar sowie um den Jahreswechsel zu Silvester. Auch die Tagestemperaturen haben einen starken positiven Einfluss auf die Frequenzen der Parkplätze. Anhand der Daten der globalen Sensorik wird darüber hinaus deutlich, dass größere Parkplätze im Jahresverlauf bedeutend mehr angefahren werden als kleinere. In Grömitz sind anhand der globalen Sensordaten ebenfalls Saisonalitäten im Jahresverlauf erkennbar, mit Nachfragespitzen beispielsweise an Ostern. Durch die Aufschlüsselung der globalen Daten nach Zielgruppen ist zudem ersichtlich, dass der Tagestourismus insgesamt eine hohe Bedeutung hat. Übernachtungstourist*innen sowie Einwohnende nutzen auf Fehmarn hauptsächlich die großen Parkplätze. In Grömitz werden einige Parkplätze ebenfalls kaum von Übernachtungstourist*innen aufgesucht. Ein erhöhtes Aufkommen von Übernachtungstourist*innen gibt es vor allem in den Ferien, beispielsweise den Herbstferien. Ein Vergleich der Frequenzen der lokalen Sensorik mit den daily visits der globalen Sensorik ergibt für einige Parkplätze auf Fehmarn moderate Korrelationen sowie für den Parkplatz Meeschendorf eine stärkere Korrelation. In Grömitz liegen bei den beiden untersuchten Parkplätzen lediglich geringe Korrelationen vor.
- **Hotspots an der Schlei:** Da die Sensoren an der Schlei erst in den letzten Monaten des Jahres 2022 installiert werden konnten, liegt lediglich ein kurzer Messzeitraum vor. Ab April 2023 sind deutlich steigende Frequenzen zu verzeichnen. Der Großraumparkplatz hat dabei die meisten Einfahrten. Dass die Größe der Polygone von Bedeutung ist, zeigt sich ebenfalls daran, dass in Kappeln rund fünfmal so viele daily visits registriert wurden wie in Arnis. Auch an der Schlei zeigen sich zudem die Auswirkungen der Feier- und Ferientage. In Kappeln wurden am Wochenende nach Christi Himmelfahrt Höchstwerte gemessen und in Arnis an Pfingstsonntag.
- **Wohnmobilstellplätze in der Holsteinischen Schweiz:** Bei den meisten Wohnmobilstellplätzen zeigen sich starke Saisonalitäten. In den Wintermonaten gibt es teilweise kaum Frequenzen, da einige Campingplätze in den Wintermonaten geschlossen haben. Mit Öffnung der Campingplätze im März/April ist eine deutliche Steigerung der Frequenzen erkennbar. Um die Feiertage, beispielsweise am Pfingstwochenende bzw. am 1. Mai, kommt es zu Spitzenwerten. Der Gini-Koeffizient spiegelt ebenfalls die hohe Konzentration der Frequenzen wider. Der WoMo Stop in Malente und der Reisemobilpark Eutin hingegen weisen seit Messbeginn über das ganze Jahr einen etwa gleichverteilten Durchfluss auf. Bei allen Stellplätzen werden freitags die meisten Fahrzeuge registriert. Neben PKWs dominieren dabei die Kleinbusse. Zu den Kleinbussen zählen Wohnmobile oder auch Campervans. Die vergleichsweise hohe Anzahl an LKWs an einigen Stellplätzen lässt sich damit erklären, dass große Wohnmobile als LKW erfasst werden. Der Tagesverlauf der Campingplätze bildet eine bimodale Kurve mit zwei Tagesmaxima, beispielsweise in den Morgen- und Abendstunden, an denen sich die meisten Personen auf

den Campingplätzen aufhalten. Die globalen Sensordaten zeigen, dass der Anteil der Übernachtungstourist*innen teilweise am höchsten ist, mit einer hohen Saisonalität. Beim WoMo Stop hingegen wurden vor allem Tagestourist*innen und Einwohnende erfasst.

- **Radrouten an der Schlei:** Für die Radrouten liegen bisher nur für einen sehr kurzen Zeitraum Daten vor. Doch auch hier ist bereits der Einfluss von Feiertagen erkennbar. An Christi Himmelfahrt und Pfingsten wurde der höchste Radverkehr registriert. Kurz vor der Fähre Arnis, bei Sundsacker, gab es an Ostern den meisten Verkehr in Richtung Fähre. Die globalen Forschungsdaten für die Radrouten weisen eine sehr geringe Reliabilität auf. Insbesondere bei den Standorten Gulde und Drecht gibt es viele Datenlücken. Die fehlenden Daten vieler Kalenderwochen aber auch die Witterungsbedingungen führen zu einer erhöhten Saisonalität. Bemerkenswert ist, dass die daily visits, die von Einwohnenden generiert wurden, unter allen Polygonen am höchsten sind, beispielsweise bei der Zählstelle Drecht, allerdings sind auch hier die Fallzahlen am niedrigsten.
- **Weitere Hotspots:** An der Westküste zeigt sich ebenfalls die Relevanz der Polygongröße anhand der Anzahl der daily visits. Alle Standorte an der Westküste weisen zudem einen Peak an Ostern auf. Es zeigt sich an der Westküste zudem ein typischer tageszeitlicher Verlauf in Form einer Glockenkurve. Der Strand in Scharbeutz hat eine deutliche Nachfragespitze in den Sommerferien und weist dabei insgesamt deutlich mehr Signale auf als die Parkplätze in Scharbeutz. Erfolgt eine Zielgruppendifferenzierte Analyse, fällt auf, dass der Tagestourismus sowohl für Scharbeutz als auch für die Westküste von hoher Bedeutung ist. Der Übernachtungstourismus ist an der Westküste stärker vertreten als an der Ostküste. Bei einer Gegenüberstellung mit lokalen Forschungsdaten zeigen sich für die Perlebucht in Büsum, den Strand in Scharbeutz und die Seebrücke in St. Peter-Ording hohe Korrelationen. Für die Einkaufsstraße Holm in Flensburg liegt jedoch lediglich eine geringe Korrelation vor.

Mit Hilfe der lokal verbauten Sensorik ist es nun möglich, Aussagen zu Frequenzen in den betrachteten Use Cases zu geben. Erstmals lassen sich so beispielsweise auch der Effekt von Feiertagen, Events, Wetter etc. anhand von Daten faktenbasiert evaluieren. Selbst wenn sich Besucher*innen nicht lenken lassen würden, ist die Messung ein Mehrwert, den Destinationen in Schleswig-Holstein für sich nutzen können. Durch das Dashboard Eco Visio und die Möglichkeit der individuellen Analyse und Beantwortung von Fragestellungen der Destinationen können auch abseits der via API abrufbaren Frequenzen diese beantwortet werden. Gerade durch die in Schleswig-Holstein bisher einmaligen Kombination aus lokaler und globaler Sensorik lässt sich das Besucheraufkommen und die Aktionsräume von Besucher*innen besser verstehen.

Die von Reif et al., 2023 aufgeführten Herausforderungen bei der digitalen Messung von Besucher*innen zeigen sich auch in diesem Falle. Unter den ökonomischen Herausforderungen stellen insbesondere die Installation und die dafür benötigten Genehmigungen sowie zeitliche und personelle Ressourcen eine große Hürde dar. Auch zeigen sich sozio-politische Herausforderungen. Wenngleich es keine Probleme aus Sicht des Datenschutzes gibt, zeigt sich doch die Abhängigkeit von Datenprovidern und die Schwierigkeit, alle wichtigen Anspruchsgruppen in der Destination von dem Mehrwert solcher Installationen zu überzeugen. Vor allem zeigen sich aber technische bzw. datenbezogene Herausforderungen, allen voran die Reliabilität und Validität der Daten. So kann es durchaus vorkommen, dass Daten nicht oder später übertragen werden, was insbesondere bei geringen Übertragungsintervallen

der Fall sein kann. Gleichzeitig hat sich die Messung von Ein- und Ausfahrten zur Erfassung von Auslastungen als nicht uneingeschränkt geeignet erwiesen (siehe Abschnitt 5). Vor allem für solche Flächen, bei denen nicht alle Zu- und Abgangsmöglichkeiten klar voneinander getrennt sind und einzeln überwacht werden können, oder für solche Flächen deren Auslastung nachts nicht annähernd null ist, wären zukünftig andere Methoden zur Erfassung der Auslastung empfehlenswert. Für solche Flächen eignen sich vermutlich solche Methoden besser, die den Gesamtzustand direkt beurteilen können. Beispiele für derartige Messmethoden wären Belegungssensoren aller einzelnen Parkplätze oder Kamerasysteme, die den ganzen Parkplatz überblicken und daraus die Autos auf dem Parkplatz zählen können. Zusätzlich wird empfohlen, dass die Messdaten der Parkplätze mit weiteren Daten plausibilisiert werden. Hierfür könnten unterschiedliche Sensortypen am selben Standort installiert oder auch regelmäßige manuelle Zählungen durchgeführt werden.

Weiterhin kann es hilfreich sein, Modellregionen zu errichten, bei denen mit Hilfe von manuellen Zählungen oder weiteren Sekundärdaten validierte lokale Sensorik installiert wird, die in der Lage ist, den Auslastungsgrad zu bestimmen. Dies kann für einzelne POI gelten, die dann übertragen werden können, beispielsweise Strandparkplätze, Wanderwege, Radrouten etc.. Ziel wäre es hierbei, für bestimmte Orte ein idealtypisches Muster zu generieren und dies auf andere Regionen zu übertragen. Dies hätte weiterhin den Vorteil, dass nicht flächendeckend in Schleswig-Holstein Sensorik installiert werden müsste. Auch der Einsatz von WLAN-Accesspoints zur indirekten Auslastungsmessung könnte in so einer Modellregion noch einmal umfassend evaluiert werden. Dafür würde sich besonders der simultane Einsatz eines Flächensensors, beispielsweise per automatisierter Personenzählung auf Videodaten anbieten.

Bei der globalen Sensorik haben die Polygongröße, die Art des Standortes sowie der Zeitraum der Betrachtung einen entscheidenden Einfluss auf die Güte der Daten aus anonymisierten Geolocation Events. Während bei kleinen Parkplätzen und kurzen Betrachtungszeiträumen nur geringe Fallzahlen generiert werden können, zeigt sich bei der Analyse von Standorten, die größer sind (beispielsweise Strand in Scharbeutz) und bei denen davon ausgegangen werden kann, dass sich aufgrund der Attraktivität der Standorte Personen länger aufhalten (bspw. Seebrücke in St. Peter-Ording) eine hohe Anzahl an daily visits. Eine der Stärken der Daten ist die Identifizierung von Bewegungsmustern im Sinne von Quell-Ziel-Gebietsbeziehungen. Dies zeigt sich neben der Analyse der Quellmärkte vor allem an den Analysen des ersten und letzten Signals des Tages. Diese Stärke muss zukünftig noch weiter in den Vordergrund gestellt werden, indem neben einer breit angelegten Analyse der Daten für einen großen Zeitraum und möglichst viele Standorte, darüber hinaus konkrete Fragestellungen in den Use Cases beantwortet werden. Mögliche Fragestellungen wären beispielsweise: "Besuchen Personen, die mittags am Museumshafen waren, nachmittags auch die Perlebucht in Büsum?"; "Lassen sich anhand der Signale Muster erkennen, welche Parkplätze in Grömitz nacheinander angefahren werden?" usw.

Weiterhin wird empfohlen, dass bei dieser speziell auf einzelne Fragestellungen abzielenden Analyse die Definitionen der Segmente spezifiziert werden. Im Rahmen dieses Projektes wurde dies auf Basis von Regelmäßigkeiten der Besuche des Endgerätes definiert, also anhand der gewöhnlichen Wohn- und Arbeitsorte der Geräte. Denkbar wären weitere Indikatoren, wie beispielsweise Anzahl der Pings in einem gewissen Zeitraum, durchschnittliche Aufenthaltsdauer oder auch die Sichtbarkeit einzelner Geräte über eine gewisse Zeitspanne in einer Destination. Dieser Detailgrad kann weiteres Datenrauschen verhindern und so zur Verbesserung der Datenqualität beitragen.

Weitere Forschungen sollten sich zudem mit dem tiefergehenden Abgleich der lokalen Sensorik mit den Big Data Quellen beschäftigen. Hier ist es elementar, dass eine lange Zeitreihe vorliegt, um die Aussagekraft zu erhöhen. Erkennbar sind im Rahmen der Projekte gute Passungen an Standorten mit einer hohen Aufenthaltsqualität (bspw. Strand in Scharbeutz, Fähranleger Amrum) und dem Vorliegen

einer Jahresreihe. Tagesgenaue Daten wären hier die bessere Alternative als Kalenderwochen, was im Rahmen des Projektes aus Fallzahlengründen nicht machbar war. Zwar konnten die typischen Tages- und Wochenverläufe der location events am Beispiel des Strandes von Scharbeutz umgerechnet werden (siehe Abschnitt 5), die Nutzung dieser Daten für die Prognose muss jedoch noch weiter erforscht werden. Das Projekt kann hier erste Ansatzpunkte liefern. Der Vergleich von lokaler und globaler Sensorik bestätigt, dass die Volumina sich deutlich unterscheiden. Ein validierter Hochrechnungsfaktor für bestimmte POI, beispielsweise Strände, Parkplätze etc. könnte einen Weg aufzeigen, der es erlaubt, auf eine lokale Installation von Sensoren zu verzichten und mit Big Data Quellen zu arbeiten. Hier ist jedoch noch weitere Forschung angezeigt.

4. Datenmanagement und -infrastruktur

Bei LAB-TOUR SH handelt es sich um ein datengetriebenes Forschungsprojekt. Dabei soll mit Hilfe von Daten der Tourismus in Schleswig-Holstein an ausgewählten Orten erfasst, prognostiziert und möglichst gelenkt werden. Dies stellt keine einfache Aufgabe dar und erfordert an vielen Stellen grundlegende Arbeit. Denn nur wenn das Grundgerüst aus Daten und Infrastruktur passt, können am Ende sinnvolle Prognosen und Alternativen nach außen gespielt werden. Datenplattformen, Data Hubs und Datenmanagement sind daher elementare Bestandteile des Projektes. Es hat aber nicht nur Einfluss auf das Projekt während seiner Laufzeit, sondern auch die Verwendung der entwickelten Software nach Projektabschluss hängt von diesen grundlegenden Konzepten ab.

Die im Projekt entwickelten Prognosemodelle und der Alternativengenerierer brauchen eine solide Datengrundlage und Dateninfrastruktur, um zuverlässige und genaue Ergebnisse zu liefern. Dabei fällt es diesen Algorithmen schwer, mit inkonsistenten sowie unterschiedlich strukturierten Daten zu arbeiten. Wenn jedoch unterschiedliche Datenquellen angebunden werden sollen, kommen diese Daten immer in unterschiedlichen Formaten. Eine Standardisierung der Datenmodelle ist somit unerlässlich.

Neben den Datenmodellen ist auch eine Standardisierung von Schnittstellen wichtig, da diese eine einfache Anbindung neuer Data Hubs in die Integrationsplattform ermöglichen. Dadurch ergibt sich eine skalierbare Plattform, da einfach neue Data Hubs ergänzt werden können, wenn die Plattform wächst. Aber auch die Persistenz findet in den Data Hubs statt. Dies schafft aber auch neue Probleme. Die Integrationsplattform, die die Daten der Hubs zusammenbringt, ist damit abhängig von der Anbindung an die Data Hubs. Aufgrund dieser Abhängigkeit muss eine hohe Verfügbarkeit gewährleistet werden, damit die Daten jederzeit durch die Integrationsplattform abrufbar sind.

Es muss also eine standardisierte Datengrundlage geschaffen werden, welche nicht nur persistent, sondern gleichzeitig hochverfügbar und skalierbar ist. Nur wenn diese drei Punkte in Einklang gebracht werden, können die darauf aufbauenden Empfehlungen und Prognosen gute Ergebnisse erzielen.

4.1 Datenfluss

Das gesamte Konzept baut auf divergierenden Datenquellen auf, um die Vorhersagen und Alternativen zu generieren. Dies stellt eine große Herausforderung aus verschiedenen Perspektiven dar. Nachfolgend wird der Datenfluss durch die Gesamtarchitektur des Forschungsprojektes beschrieben, also wie die Daten von den Quellen über die Verarbeitung in den einzelnen Komponenten bis zu den Ausspielungskanälen gelangen. Dabei soll eine vereinfachte Zusammenfassung einzelner Komponenten genutzt werden, um einen Gesamtüberblick zu ermöglichen.

Im Folgenden wird der Datenfluss nun von links nach rechts anhand von Abbildung 51 beschrieben. Den Anfang bilden die Sensordaten, gekennzeichnet als gekoppelte Standortmessungen (Datenquelle A-C), welche von den Sensoren des Anbieters Eco Counter an den verschiedenen Messstellen generiert werden. Diese werden von dem Daten Hub bei ADDIX regelmäßig eingesammelt, transformiert und gespeichert. Hier werden die Daten vorgehalten und können über die Schnittstelle bei ADDIX von allen Komponenten abgerufen werden. Um möglichst schnell auf unvorhergesehene Trends reagieren zu können und die Prognosen anzupassen, ist es essenziell, dass die Daten zeitnah von den Sensoren im Data Hub bei ADDIX abrufbar gemacht werden.

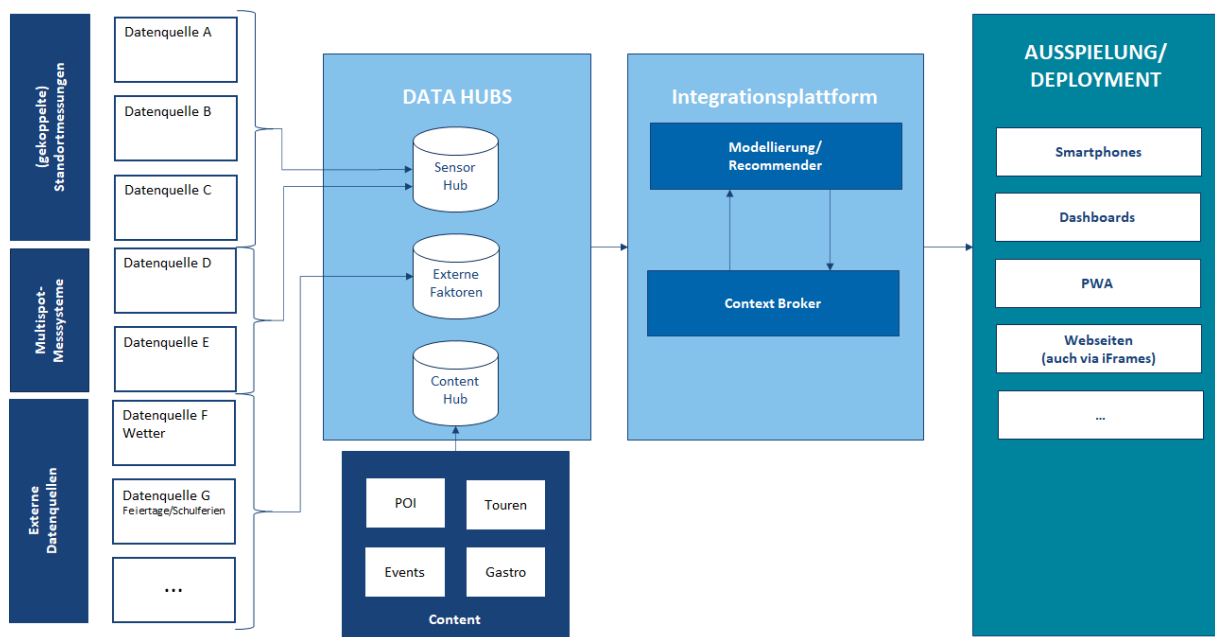


Abbildung 51: Datenfluss im Projekt LAB-TOUR SH

Wenn in Zukunft weitere Sensoren installiert oder andere Anbieter in die Plattform integriert werden, kann dies entweder über den Daten Hub von ADDIX geschehen, oder über neue Data Hubs, die problemlos in den Datenfluss von LAB-TOUR SH integriert werden können. In der Abbildung sind diese Datenquellen als A bis C repräsentiert und der zugehörige Daten Hub mit dem Namen Sensor Hub benannt. Dabei steht der einzelne Sensor Daten Hub repräsentativ für eine Vielzahl an Daten Hubs, die Sensor Daten für das Gesamtkonstrukt vorhalten könnten.

Eine weitere wichtige Datenquelle sind die Wetterdaten, welche einen maßgeblichen Einfluss auf die Auslastung der POI in Schleswig-Holstein haben (siehe oben). Hier wurde im Projekt auf den Anbieter OpenWeatherMap gesetzt, da hier neue Standorte zusätzlich erfasst werden können und die Datenquelle nicht von einzelnen Wetterstationen abhängig ist. Diese Daten werden von der Datenquelle F in den Daten Hub für externe Faktoren geladen und stehen dort für die Prognose zur Verfügung. Für jeden dieser Sensorstandorte wird dabei eine Zeitreihe gespeichert, beginnend mit dem Zeitpunkt der ersten Datenlieferung des Sensors, ab Aufnahme in die Plattform. Neben den Wetterdaten werden auch Daten zu Feiertagen und Schulferien in dem Daten Hub für externe Faktoren gespeichert. Dabei werden zwei öffentliche Open-Source APIs angesprochen, welche die Daten zu Feiertagen und Schulferien zur Verfügung stellen. Die Schulferien und Feiertage werden bundeslandspezifisch gespeichert.

Auf diese Daten kann dann das Prognosemodul zurückgreifen, um bessere Vorhersagen zu treffen. Diese Datenquellen sind in der Grafik als Datenquelle G repräsentiert.

Zusätzlich zu den lokalen Sensoren wurden Daten von sogenannten Multispot-Messsystemen eingekauft. Diese sind als Datenquellen D und E in der Grafik zu finden. Dabei handelt es sich um die globale Sensorik, die in Abschnitt 3.1.2 beschrieben wurde. Diese Daten sind statisch und werden ebenfalls in dem Sensor Hub gespeichert. Die Daten können so von den Prognosemodellen verwendet werden.

Ein weiterer Inputkanal ist die Datenbank der TASH. In der Grafik ist diese als das mit Content beschriftete Rechteck zu finden. Hier sind Daten zu den einzelnen POI und Gastro des Landes Schleswig-Holstein gespeichert. In Zukunft sollen auch Events in die Prognosen mit einbezogen werden. Die POI und Gastro Daten sollen im Projekt vor dem Speichern in dem Content Hub transformiert und gefiltert werden. Von dort können die Daten direkt in die Integrationsplattform fließen und anschließend dem Alternativengenerierer und der Prognoseplattform zur Verfügung gestellt werden.

In Zukunft wäre es auch denkbar Daten aus den Netzwerken der Sozialen Medien in die Integrationsplattform mit einfließen zu lassen. Denn wenn ein Event oder ein Ort über Social Media häufig erwähnt wird, lässt dies auf eine höhere Auslastung in naher Zukunft schließen. Es ergibt Sinn, einen weiteren Data Hub hierfür aufzubauen, da anzunehmen ist, dass intensive Transformations- und Filterarbeiten auf die Daten angewandt werden müssen. Von diesem Hub könnten die Daten dann wieder in die Integrationsplattform fließen. Allgemein lassen sich dank der modularen Strukturen sehr einfach neue Datenquellen jeglicher Art in das Gesamtkonstrukt integrieren.

Der Context Broker in der Integrationsplattform dient als zentrale Datendrehscheibe zum Zwischenspeichern und zum zur Verfügung stellen der Daten. Hier werden die Prognosen, ebenso wie die Alternativen zwischengespeichert und können dann über ein Gateway aus dem Internet abgefragt werden. Somit können die verschiedenen externen Plattformen zur Auspielung der Daten angebunden werden.

Mit all diesen gebündelten Daten auf der Integrationsplattform, können die Prognosen und Alternativen generiert werden. Zunächst sammelt das Prognosemodul alle relevanten Daten von der Integrationsplattform ein und trifft Vorhersagen zur Auslastung der einzelnen POI. Danach werden diese Vorhersagen auf den Context Broker gespielt, von wo aus diese dann im Anschluss für die Auspielung an die Deployments, bzw. für die Alternativengenerierung zur Verfügung gestellt werden.

Der Alternativengenerierer holt sich ebenfalls alle relevanten Daten von der Integrationsplattform, mit denen dann für die einzelnen POI Alternativvorschläge generiert werden. Dieses Ergebnis wird dann auf den Context Broker zurückgespielt, von wo aus diese den Deployments zur Verfügung steht.

Die Daten zur Prognose und der Empfehlung zu alternativen POI sind über eine Schnittstelle aus dem Internet abrufbar. Dabei besteht dann die Möglichkeit, die im Projekt generierten Daten über die Schnittstelle in Websites, PWAs, Dashboard oder Apps zu integrieren. Auch das Exemplarisch im Rahmen des Projektes programmierte Frontend greift über diese Schnittstelle auf die gesammelten Informationen und Ergebnisse zu.

Damit der Datenfluss zwischen den einzelnen Komponenten sichergestellt ist, mussten einheitliche Schnittstellen geschaffen werden. Diesem Thema wird sich der nächste Abschnitt widmen.

4.2 Datenschnittstellen und Datenhubs

4.2.1 Motivation

Für die Einbindung der geplanten Datenhubs von LH-IND und ADDIX sowie zukünftiger weiterer Data Hubs soll eine standardisierte Datenschnittstelle geschaffen werden. Das Ziel ist die Einführung einer Abstraktionsschicht zwischen der Sensorik bzw. der Datenhaltung von der Business-Logik, welche in der Integrationsplattform implementiert ist. Eine standardisierte Datenschnittstelle in Zusammenhang mit einem standardisierten Datenformat erlaubt das einfache Einbinden von zusätzlichen Datenhubs, d. h. die einfache Skalierbarkeit des Systems in Bezug auf zukünftige Erweiterungen. Ein weiterer Vorteil ist die strengere Trennung der Zuständigkeiten der Komponenten des gesamten IT-Systems. Grundsätzlich erhöht die Trennung der Zuständigkeiten die Komplexität des Gesamtsystems. Jedoch wird der Aufwand für die Umstrukturierung der Daten in das richtige Format zu den Datenhubs ausgelagert. Dadurch ist auf Seiten der Integrationsplattform nur ein Connector nötig, um alle Datenhubs anzubinden. Das erhöht einerseits die Wartbarkeit der Integrationsplattform, weil ein Teil des Wartungsaufwandes bei den Datenhubs liegt. Andererseits wird auch das Einbinden neuer Datenhubs erleichtert. Dies reduziert ferner die Abhängigkeit von einzelnen Datenanbietern durch einfachere Austauschbarkeit der einzelnen Datenhubs (Reduzierung des „Vendor-Lock-In“).

In den Datenhubs werden die Sensordaten persistiert, d. h. sie bleiben langfristig erhalten und abrufbar. Die Datenanbieter erlauben üblicherweise nur einen zeitlich begrenzten Zugriff, beispielsweise das vergangene Jahr. Außerdem können notwendige Schritte wie eine Säuberung, Vorprozessierung oder Aggregation der Daten bereits auf den Datenhubs durchgeführt werden. Dies führt wiederum zu einer weiteren Vereinfachung der Integrationsplattform, auch in Hinblick auf die Verteilung der Rechenlast. Außerdem wird die Menge der übermittelten Daten reduziert, wenn nur die aggregierten und vorprozessierten Daten abgerufen werden müssen.

Zusätzlich zu den Sensordaten sollen weitere externe Daten zur Auslastungsprognose herangezogen werden. Dies sind beispielsweise Wetterdaten, Informationen zu Events, Feiertage und Schulferien, die die Auslastung von POI beeinflussen können. Zusätzlich werden globale Sensordaten eingebunden, wie etwa historische Auslastungsdaten, die aus der Auswertung der daily visits geschätzt wurden. Teilweise liegen die Daten offen vor, wie im Portal der KielRegion. Für Wetterdaten oder Big Data Datenquellen hingegen müssen Lizenzen erworben werden. Die externen Daten liegen dabei in standardisierten Datenmodellen vor oder müssen dabei in diese transformiert werden.

4.2.2 Anforderungen

Damit die Schnittstellen im oben genannten Sinne funktionieren können, müssen sie den folgenden Anforderungen genügen:

1. **Freiheit:** Die Datenschnittstelle muss kostenlos und frei zugänglich sein, ohne versteckte Kosten oder Einschränkungen.
2. **Dokumentation:** Eine umfassende und gut strukturierte Dokumentation sollte vorhanden sein, um Entwicklern bei der Nutzung der Schnittstelle zu helfen. Die Dokumentation sollte klar, präzise und leicht verständlich sein.
3. **Standards:** Die Schnittstelle sollte den gängigen Industriestandards entsprechen, um eine nahtlose Integration in bestehende Systeme und Anwendungen zu ermöglichen. Dies umfasst zum Beispiel die Unterstützung von gängigen Datenformaten wie JSON oder XML.

4. **Skalierbarkeit:** Die Schnittstelle muss skalierbar sein, um steigende Datenmengen und wachsende Benutzeranforderungen effizient zu verarbeiten. Sie sollte in der Lage sein, sowohl kleine Anfragen als auch große Lastspitzen zu bewältigen, ohne dabei die Performance einzuschränken.
5. **Performance:** Die Schnittstelle sollte eine hohe Leistung bieten, um schnelle und effiziente Datenübertragungen zu ermöglichen. Dies beinhaltet eine minimale Latenzzeit sowie eine schnelle Verarbeitungsgeschwindigkeit der Anfragen.
6. **Verfügbarkeit:** Die Schnittstelle muss rund um die Uhr verfügbar sein, um kontinuierlichen Zugriff auf die Daten zu gewährleisten. Sie sollte eine zuverlässige Infrastruktur und Maßnahmen zur Fehlerbehandlung und Ausfallsicherheit bieten, um mögliche Unterbrechungen zu minimieren.

4.2.3 Umsetzung

Zu Beginn des Projektes gab es starke Verzögerungen, insbesondere bei der Installation und Inbetriebnahme der Sensoren. Anstatt wie geplant Anfang der Saison 2022 konnten lokale Sensoren erst ab August 2022 durch die Firma Eco Counter installiert werden und fingen an, Daten zu übermitteln. Außerdem kam es häufig zu Fehlern und Verzögerungen bei der Übermittlung der Daten von den Sensoren zu dem von Eco Counter bereitgestellten Datenhub.

Im Folgenden soll nun der Datenhub von ADDIX als Beispiel für einen Datenhub und die verwendete Schnittstelle näher betrachtet werden. Dabei werden die Schritte des Datenabrufs, der Datenaufbereitung, ebenso wie die Persistierung der Daten besprochen. Da die Daten, welche vom Datenlieferanten geliefert werden, nicht immer dem gewünschten Datenstandard entsprechen, benötigt der Daten Hub eine Logik, um Daten zu transformieren. Für diesen ETL-Prozess eignet sich Node-RED hervorragend. ETL steht dabei für Extract, Transform, Load, was im Folgenden näher beschrieben wird. In der Node-RED Nomenklatur wird ein solcher Prozess als Flow bezeichnet.

Mit einem oder mehreren Flows in Node-RED können Daten von beliebigen API-Schnittstellen unterschiedlichster Lieferanten gesteuert abgeholt (Extract), in die gewünschte Form gebracht (Transform) und dann in den Context Broker eingelagert (Load) werden. Die Steuerung kann hier so automatisiert werden, sodass auch eine zeitlich begrenzte Autorisierung der API beachtet werden und der Token bei Bedarf erneuert werden kann. Ein weiterer Vorteil von Node-RED ist eine schnelle und flexible Anpassung der Flows. Damit kann der ETL-Prozess mit wachsenden Anforderungen mitwachsen.

Der Datenanbieter Eco Counter bietet eine REST API an, welche einen einfachen Zugriff auf die Daten der Zählstellen ermöglicht. Die zurückgelieferten Daten im JSON-Format sind als Sites mit zugeordneten Channels organisiert. Eine Site ist ein Messpunkt, der zum Beispiel eine Einfahrt zu einem Parkplatz beobachtet. Die Channels sind dann die gemessenen Verkehrsströme dieser Messstelle. Die Verkehrsströme unterscheiden diverse Fahrzeugklassen, welche in Kombination mit der Richtung in den Channels verdatet werden. Damit die Daten in einer konsistenten Form zur übergreifenden Nutzung bereitgestellt werden können, wurde je Site und je Channel eine Entität vom Typ *TrafficFlowObserved* erstellt. Das genutzte Datenmodell ist bereits in den Smart Data Models enthalten und konnte somit für dieses Projekt verwendet werden.⁷

⁷ <https://github.com/smart-data-models/dataModel.Transportation/tree/master/TrafficFlowObserved>

Die Entitäten werden im nächsten Schritt mit zusätzlichen Attributen angereichert. Da jede Site und jeder Channel über eine Geokoordinate verfügt, ist es möglich, hierzu die Adresse über eine Anfrage an einen Nominatim Server zu erhalten. Dieser löst die Geokoordinate in eine Adresse auf. Zusätzlich zu der Adresse wird die codierte Richtung übersetzt, die Datenquelle angegeben und der Owner zugeordnet. So wird für eine Messstelle der nötige Context bereitgestellt. Mit diesen erzeugten Entitäten ist es nun möglich, zyklisch die Daten aus der Plattform auszulesen und für jede Messung den letzten Wert im Context Broker zu speichern. Hierzu werden je Site und je Channel die Daten für ein bestimmtes Zeitfenster in der gewünschten Auflösung abgefragt.

Exkurs: Datenmengen und Schnittstelle

Die Ausschreibungen für die Sensorik in diesem Projekt sah vor, dass die Daten alle 15 Minuten bereitgestellt werden, somit müssen für die 62 Sites (Sensorstandorte) mit bis zu 6 Channels 372 Anfragen je 15 Minuten an die Eco Counter API gesendet werden. Das bedeutet, dass an nur einem Tag 35.712 Anfragen generiert werden, die die Eco Counter API bedienen muss.

Dadurch ist die API des Datenlieferanten einer hohen Belastung ausgesetzt. Während der Entwicklung und bei Nachberechnungen müssen Daten auch über einen mehrtägigen Zeithorizont abgerufen werden, um die Daten von der Plattform des Datenlieferanten in den Daten Hub zu transferieren. Bei der Auswahl des Datenlieferanten ist auf eine leistungsfähige Plattform zu achten. Je enger der Zyklus und je mehr Messstellen angefragt werden müssen, umso höher ist die Anzahl der nötigen Anfragen gegen die Plattform. Die Grenzen der REST API von Eco Counter sind hier schnell erreicht. Push-Datenlieferungen sind daher einer Pull-Schnittstelle immer vorzuziehen. Eine REST API von Eco Counter ist eine Pull-Schnittstelle da immer angefragt werden muss, ob neue Daten vorhanden sind. Im Fall, dass keine neuen Daten vorliegen, entsteht somit eine Belastung, die vermieden werden sollte. Eben diese Einsparung von Ressourcen und Reduzierung von Netzwerkverkehr kann erreicht werden, wenn Datenlieferanten neue Daten selbstständig an das Data Hub senden würde. Auch Pub/Sub Services wie MQTT sind hier deutlich besser geeignet.

Da nun für die gewünschten Messstellen die Daten je Zeiteinheit abgefragt werden, können diese durch ein Update der Entität im Context Broker abgelegt werden. Der Context Broker repräsentiert jedoch immer nur den letzten übertragenen Datenstand, da aber mit Hilfe der Daten Auslastungsprognosen erstellt werden sollen, müssen die Daten in geeigneter Form historisiert werden. Um die Historisierung zu realisieren, bietet der Datenhub einen speziell dafür vorgesehenen Dienst namens Quantum Leap an. Der Quantum Leap übersetzt die NGSI-LD Daten in ein TimescaleDB Format, sodass je Entität eine Zeitreihe persistiert wird. Dazu bietet der Quantum Leap einen eigenen API-Endpunkt, an den die Daten gesendet werden können. Hier wird die Funktionalität des Context Brokers genutzt, der es ermöglicht Subscriptions zu erstellen die auf Änderungen des Datenbestands im Context Broker reagieren und damit eine Benachrichtigung an den Quantum Leap Endpunkt sendet, sobald sich eine definierte Änderung an einem beobachteten Datenmodell oder einer beobachteten Entität ereignet. Der Quantum Leap bereitet die Daten daraufhin auf, um sie an eine angeschlossene PostgreSQL TimescaleDB zu übermitteln. Dort werden die Daten je Tenant und Datenmodell dann in eigene Tabellen abgelegt und es entsteht eine Zeitreihe der Ereignisse. Diese lassen sich wieder über die NGSI-LD API abfragen und für weitere Berechnungen nutzen.

Exkurs: TimescaleDB Pflege

Die TimescaleDB zeichnet sich durch eine positive Eigenschaft aus: Durch eine kluge Konfiguration der Tabellen, interne Prozesse und kontinuierliche Abfragen lässt sich eine optimale Datenverwaltung erzielen. Wenn die Rohdaten aus dem Quantum Leap in die TimescaleDB geschrieben werden, entstehen Teilbereiche in der Hypertabelle, die nach einer bestimmten Zeit gefüllt sind. Diese sogenannten Chunks können „komprimiert“ werden. Das Komprimieren der Daten ist ein Umschreiben der internen Struktur auf ein platzsparendes und leseoptimiertes Format. Es hat jedoch den Nachteil, dass ein komprimierter Chunk nur gelesen und nicht bearbeitet werden kann. Aber dieser Chunk benötigt ca. 70-90% weniger Platz in der Datenbank und ist so optimiert, dass der lesende Zugriff deutlich schneller das Ergebnis liefert als ein unkomprimierter Chunk. Zusätzlich bietet die Datenbank die Möglichkeit, Daten nicht nur effizient zu speichern, sondern auch zuvor zu verarbeiten. So können zum Beispiel nach vier Wochen nicht mehr Werte für jede Viertelstunde gespeichert werden, sondern lediglich stündliche Werte mit den Minimal-, Maximal- und Durchschnittswerten. Nach einem Jahr könnten dann beispielsweise nur noch stündliche Werte vorgehalten werden. Damit ist eine TimescaleDB optimal geeignet große Datenmengen so verdichtet zu speichern, dass zwar eine lange Historie aber nicht jeden einzelnen Rohdatensatz aufbewahrt wird.

Da der von ADDIX implementierte Datenhub den FIWARE Context Broker nutzt, ebenso wie die Integrationsplattform, ist der Datenimport aus dem Datenhub direkt und problemlos möglich. Im allgemeinen Fall müsste ein Connector entwickelt werden, der die Daten aus dem Datenhub in das richtige Datenmodell transformiert. Dieser Aufwand erübrigt sich durch Nutzung derselben Technologie. Ein bereits genannter Vorteil der Nutzung des FIWARE Context Brokers ist die Möglichkeit, Subscriptions einzurichten. Dies kann ebenfalls für die Anbindung der Integrationsplattform ausgenutzt werden, da der Datenhub neue Daten direkt übermittelt. Dadurch entfällt die Einrichtung einer zeitlich getakteten Datenabfrage durch die Integrationsplattform.

Als weitere externe Daten wurden sowohl statische als auch dynamische Daten eingebunden. Diese Daten liegen jeweils in den Datenformaten vor, die der Datenanbieter für sinnvoll erachtet. In der Regel handelt es sich hierbei um Standard-Datenformate. Für diese Daten wird jeweils ein Connector implementiert, um die Daten in die Datenhubs zu importieren. Da dieser Vorgang dem Ablauf im Datenhub von ADDIX sehr ähnlich ist, wird im Weiteren auf eine detaillierte Beschreibung verzichtet. Es folgt lediglich eine Auflistung der zusätzlichen externen Datenquellen und Schnittstellen. Im Einzelnen wurden eingebunden:

- **Geodaten von OpenStreetMap (OSM):** Hierbei handelt es sich um eine offene Datenquelle, die von Freiwilligen erstellt und gepflegt wird. Die Daten sind erfahrungsgemäß sehr gut nutzbar, jedoch ohne Gewähr.
- **Daten über touristische Ziele (POI) der TA.SH:** Dies ist eine institutionelle Datenquelle, die Daten liegen ebenfalls frei verfügbar vor.
- **Sensordaten von LH-IND:** LH-IND hatte bereits vor Projektstart Sensorik an einem Strand und Parkplätzen bei Scharbeutz installiert (siehe Abschnitt 3) und diese Daten auch für dieses Projekt zur Verfügung gestellt. Durch die Nutzung dieser Daten war es zu Beginn des Projektes möglich, mit der Modellentwicklung zu beginnen, bevor die Daten von Eco Counter verfügbar waren.
- **Wetterdaten von OpenWeatherMap:** Für diese Daten wurde eine Lizenz erworben.

- **Ferien:** [Deutsche Ferientermine als JSON \(ferien-api.de\)](#). Hierbei ist zu beachten, dass sich die Ferientage im Januar bei den Weihnachtsferien des Vorjahres finden.
- **Feiertage:** [Feiertage Webservice \(feiertage-api.de\)](#)
- **Veranstaltungen:** Veranstaltungen sind ebenfalls von der Plattform der TA.SH abrufbar, beinhalten jedoch keine Information über die zu erwartende Besucherzahl.

In dem Projekt müssen aber nicht nur externe Schnittstellen angeschlossen werden. Die einzelnen Komponenten im Projekt müssen auch untereinander kommunizieren können. Dies wurde mit der Ausnahme des Context Brokers auf der Grundlage der REST-Prinzipien gelöst. Dies ist ein weit verbreiteter Standard und bietet sich aufgrund der Architektur der Plattform intern an. Da die interne Kommunikation in den folgenden Abschnitten detailliert behandelt wird, genügt an dieser Stelle eine einfache Darstellung.

Damit das Ziel der Lenkungswirkung von Tourist*innen erreicht werden kann, müssen die erzeugten Daten auch von außen zugänglich gemacht werden. Hierbei werden über ein API-Gateway bestimmte Routen auf den Context Broker freigegeben, damit die Daten von dort abgeholt werden können. Bei der Schnittstelle handelt es sich um die bereits zuvor eingeführte NGSI-LD Schnittstelle. Final soll nun noch einmal kompakt geprüft werden, ob die Schnittstellen den in dem vorherigen Abschnitt aufgestellten Anforderungen gerecht werden.

1. **Freiheit:** Sowohl der FIWARE Context Broker als auch das Python Package, welches zur Implementierung der REST-APIs verwendet wurde, sind OpenSource Projekte. Somit ist gesichert, dass diese Implementierungen frei zugänglich, transparent und kostenlos bleiben. Beide Projekte haben eine aktive Community, die zu den Projekten beitragen, wodurch auch eine langfristige Funktionsfähigkeit der Softwarekomponenten gewährleistet ist.
2. **Dokumentation:** Die Schnittstelle des Context Brokers greift dank der Pflege der FIWARE Foundation auf eine detaillierte und umfangreiche Dokumentation zurück. Somit ist für diese Schnittstelle der Punkt erfüllt. Bei den eigens implementierten Schnittstellen nach dem REST-Designpattern wird automatisch eine Dokumentation, dem OpenAPI Standard entsprechend, generiert und ist dann über einen Endpunkt der Schnittstelle abrufbar. Hier sind die Routen, ebenso wie die Datenmodelle der API dokumentiert.
3. **Standards:** Sowohl die NGSI-Schnittstelle als auch die REST-Schnittstellen sind Industriestandards, die beide das gängige Datenformat JSON zurückliefern.
4. **Skalierbarkeit:** Die internen Schnittstellen werden später noch im Detail erklärt. Deshalb soll hier nur eine kurze Ausführung folgen. Diese Schnittstellen wurden angelehnt an dem Designpattern von Microservices entwickelt, sind mit der Implementierung in Kubernetes leicht skalierbar und können sich an steigende Anforderungen problemlos anpassen. Für die Schnittstelle von FIWARE ist ebenfalls die Skalierbarkeit gegeben, da diese in Kubernetes implementiert ist und somit mit geringem Aufwand mit dem Cluster und steigenden Anforderungen mitwachsen kann.
5. **Performance:** Bei selbstentwickelten Schnittstellen wurde versucht, möglichst leichtgewichtige API-Softwarepakete zu verwenden, um die Codebasis möglichst schmal zu halten. Dies darf aber nicht die Flexibilität oder Leistungsfähigkeit der Schnittstelle selbst beeinflussen. All diese Anforderungen konnten mit dem Python Paket FastAPI erfüllt werden. Dieses wurde für die internen Schnittstellen im Projekt verwendet, wodurch alle Anforderungen für die internen Schnittstellen gut umsetzbar waren. Die von der FIWARE Foundation entwickelte Schnittstelle

ist ebenso performant und wird in Smart-City Projekten eingesetzt, um die Sensorik und Daten ganzer Städte zu bündeln.

6. **Verfügbarkeit:** Auch dieser Punkt ist sowohl für die eigenentwickelten Schnittstellen als auch für die Schnittstelle vom FIWARE Context Broker gegeben. Sichergestellt für beide Schnittstellen wird dies dadurch, dass das Deployment im Kubernetes Cluster geschieht. Eine der wichtigsten Eigenschaften eines Kubernetes Deployments ist es, dass eine hohe Verfügbarkeit und Ausfallsicherheit gewährleistet wird.

Somit konnten bei der Umsetzung der Schnittstellen innerhalb des Projektes alle Anforderungen erfüllt werden. Dennoch gibt es weiteres Verbesserungspotenzial, was im nächsten Abschnitt thematisiert werden soll.

4.2.4 Learnings und Potentiale

Generell zeigte sich während der Durchführung dieses Pilotprojektes, dass bei der Anbindung von Sensorik deutlich mehr Komplikationen auftreten können als ursprünglich angenommen. Die Sensoren sollten durchgehend Daten liefern und die Daten gängigen Standards entsprechen. So beispielsweise sollte eine Unterscheidung möglich sein, ob der Wert „0“ zurückgemeldet wird, oder ob ein fehlender Wert vorliegt.

Eine weitere Erkenntnis aus dem Projekt ist, dass es durchaus noch Potenziale im Open Data Portal Schleswig-Holstein gibt. Dies gilt insbesondere für den Bereich Tourismus. Es sind hier keine Feriendaten abrufbar. Feriendaten sind aber gerade im Bereich Tourismus von großer Bedeutung und deren Einbindung wäre unproblematisch möglich. Außerdem war es zunächst schwierig, die POI-Daten der TA.SH einzubinden. Diese Daten sollten ebenfalls im Open Data Portal offen zur Verfügung gestellt werden. Allerdings ist diese Übertragung derzeit nicht eindeutig geregelt, bzw. war es zum Zeitpunkt, an dem die Daten für das Projekt benötigt wurden, nicht, was zu Verzögerungen bei der Bereitstellung führte. Zu betonen ist hierbei, dass die TA.SH hier in Abhängigkeit zu einem privaten technischen Dienstleister steht, welchem die Prinzipien von offenen Daten erst mühsam erklärt werden mussten. An diesem Beispiel zeigt sich das Ausbaupotential des Open Data Portals, und die Wichtigkeit eines einheitlichen Vorgehens beim Ausbau der offen zur Verfügung gestellten Daten.

Ein weiterer Punkt sind die Veranstaltungen, die auf der Plattform der TA.SH abrufbar sind. Zum einen sind diese Daten nicht vollständig. Dies erscheint nachvollziehbar, da die TA.SH die Veranstalter nicht zwingen kann, ihre Veranstaltungen immer im Vorfeld der TA.SH zu melden. Zum anderen fehlt die erwartete Besucherzahl, was bei einer Darstellung der Daten in einem Veranstaltungskalender nicht zwingend ist, aber im Rahmen dieses Projektes erforderlich, da die Angabe der Besucherzahl natürlich bei der Gestaltung eines Recommenders elementar ist. Dieser Wert ist besonders interessant zur Vorhersage der Auslastung einer Lokalität bzw. der umliegenden Parkplätze. Hier erscheint eine automatisierte Klassifizierung von Veranstaltungen und eine Vorhersage von Besucherzahlen sehr vielversprechend und könnte künftig als ergänzende Angabe bei Veranstaltungen eingepflegt werden.

4.3 Datenmodelle

Um die Interoperabilität zwischen den Datenhubs und der Integrationsplattform zu ermöglichen, müssen die gespeicherten Daten standardisiert vorliegen. Nur so ist eine Verarbeitung der Daten in der Integrationsplattform möglich. Dies hat sodann den Vorteil, dass sie im Rahmen einer Analyse sowie zur Entwicklung eines Prognosemoduls miteinander in Beziehung gesetzt werden können. Zudem ist

es auf dieser Basis möglich, die in diesem Projekt erhobenen Daten mit anderen – außerhalb des Projektes – zu vergleichen. Die Einhaltung der Datenstandards ist die Voraussetzung dafür, dass zusätzliche Datenhubs einfach an die Integrationsplattform angeschlossen werden können. Gleichwohl ergibt sich für das Projekt LAB-TOUR SH die Herausforderung, dass es vergleichbare Projekte bisher nicht gibt und somit für das Recommender-Modul neue Datenmodelle entwickelt werden müssen.

Für ein zukünftiges Produktivsystem ist es notwendig, die tatsächlichen Auslastungen zusammen mit den prognostizierten Auslastungen sowie die Alternativen nach außen zur Verfügung zu stellen. Für den Zugriff auf diese Daten sind zwei unterschiedliche Ansätze vorgesehen. Um die Umsetzbarkeit des Konzepts zu überprüfen, soll zum einen während der Projektlaufzeit ein Frontend prototypisch entwickelt werden, das die Auslastungsprognosen und Alternativen in einer Web-Applikation darstellt. Zum anderen soll der Recommender als Webservice auch für externe Anwendungen zur Verfügung stehen. Hierfür müssen die vom Recommender bzw. der Integrationsplattform gelieferten Daten gängigen Datenstandards entsprechen.

4.3.1 Anforderungen

Um eine gute Integration der neu definierten Datenmodelle und die Verwendung der Daten in andere Forschungsprojekte und in die Praxis zu ermöglichen, müssen die Datenmodelle den folgenden Anforderungen genügen:

1. **Standards:** Datenmodelle sollten den gängigen Industriestandards entsprechen, um eine nahtlose Integration und Interoperabilität mit anderen Systemen zu gewährleisten. Es ist empfehlenswert, auf etablierte Standards wie schema.org oder FIWARE zurückzugreifen, da diese weit verbreitet und gut dokumentiert sind.
2. **Verwendung bestehender Datenmodelle:** Wenn möglich, sollten bestehende Datenmodelle genutzt werden, anstatt komplett neue Modelle zu definieren. Dies ermöglicht eine bessere Wiederverwendbarkeit und erleichtert die Integration mit bereits existierenden Datenquellen und -anwendungen. Durch die Verwendung etablierter Modelle können auch die Vorteile ihrer bereits bestehenden Strukturen und Konzepte genutzt werden.
3. **Erweiterung bestehender Modelle:** Falls ein bestehendes Datenmodell nicht alle erforderlichen Informationen abbilden kann, ist es ratsam, das vorhandene Modell zu erweitern. Dadurch wird die Konsistenz und Kompatibilität mit dem vorhandenen Ökosystem bewahrt, während die spezifischen Anforderungen erfüllt werden.
4. **Neues Modell nur bei Bedarf:** Die Definition eines komplett neuen Datenmodells sollte nur dann erfolgen, wenn es keine bestehenden Modelle gibt, die den Anforderungen gerecht werden können. Diese Situation tritt in der Regel auf, wenn spezifische Domänen oder Anwendungsfälle abgedeckt werden müssen, für die keine geeigneten Standardmodelle verfügbar sind. In solchen Fällen ist es wichtig, die best practices der Modellierung zu beachten und das neue Modell sorgfältig zu entwerfen.
5. **Vollständigkeit:** Schließlich müssen die Datenmodelle alle erforderlichen Attribute der Daten enthalten.

Durch die Einhaltung dieser Anforderungen an die Datenmodelle wird sichergestellt, dass die Modelle den gängigen Standards entsprechen und eine gute Integration in die gesamte Plattform ermöglichen. Die Verwendung bestehender Modelle und die Erweiterung dieser Modelle reduzieren Duplizierungen in der Entwicklung, fördern die Interoperabilität und erleichtern die Zusammenarbeit in einem breiteren Ökosystem.

4.3.2 Umsetzung

Bei der Umsetzung, also der Auswahl und ggf. Änderung/Erstellung der Datenmodelle, wurde sich an die vorher beschriebenen Anforderungen gehalten. Durch ein systematisches Vorgehen konnte die Anzahl der neu erstellten Datenmodelle minimal gehalten werden. Dabei wurde der Arbeitsablauf in zwei Schritte unterteilt. In einem ersten Schritt wurden Datenmodelle verschiedener Organisationen gesichtet und auf ihre Anwendbarkeit im Forschungsprojekt LAB-TOUR SH geprüft. Konnte in diesem Schritt kein passendes Datenmodell für die im Projekt verwendete Daten ermittelt werden, mussten in einem weiteren Schritt Datenmodelle angepasst und im Ausnahmefall neu erstellt werden. Dabei wurde stets darauf geachtet, die Konventionen der großen Organisationen einzuhalten. Dies ermöglicht eine nahtlose Integration in die bereits bestehenden Datenmodelle.

Sichtung bereits bestehender Datenmodelle

Es bietet sich an, diese Datenmodelle auf international verbreiteten und bereits bestehenden Sprachen zur Beschreibung von Daten (Ontologien) aufzubauen. Gemeinhin bekannt und oftmals auch als „Quasi-Standard“ bezeichnet, ist hierbei schema.org. Da diese Auszeichnungssprache jedoch eher generischer Natur ist, wurde im Projekt LAB-TOUR SH zusätzlich auf die im Smart City Kontext verbreiteten Smart Data Models zurückgegriffen, die ihrerseits mit der schema.org-Logik kompatibel sind. Da das Projekt im touristischen Kontext stattfindet, wurde als dritte Ontologie die Domain Specifications der Open Data Tourism Alliance (ODTA) berücksichtigt.

- **schema.org:** Die erste Anlaufstelle für Datenmodelle ist schema.org, eine kollaborative, gemeinschaftlich betriebene Website mit dem Ziel, Datenstandards vor allem im Internet zu verbreiten. Gegründet wurde schema.org von Google, Microsoft, Yahoo und Yandex. Entsprechend hoch ist die Verbreitung der hier spezifizierten Standards. Ein Nachteil von schema.org ist die Größe der Sammlung und die entsprechende Langsamkeit, mit der neue Modelle in die Sammlung aufgenommen werden. Daher wurden in diesem Projekt zwei weitere Sammlungen von Datenmodellen verwendet.
- **ODTA:** Die ODTA ist ein Zusammenschluss von rund 30 touristischen Länder- und Bundesinstitutionen aus Deutschland, Österreich, der Schweiz und Italien. Das Ziel ist die Etablierung von Datenstandards im Tourismusbereich. Bei der ODTA befasst sich derzeit eine Arbeitsgruppe mit Datenmodellen mit speziellem Fokus auf Tourismus. Ein Großteil der Arbeit der ODTA befindet sich zum Zeitpunkt des Projektes noch in der Konzeptionierungsphase. Um eine aktive Rolle in diesem Prozess zu spielen und spezifisches Fachwissen beizusteuern, wurde in Zusammenarbeit mit dem Projekt AIR Kontakt mit der Arbeitsgruppe aufgenommen und Feedback übermittelt. Es stellte sich heraus, dass die Arbeitsgruppe bereits ein RFC-Dokument (Request for Comments) erstellt hatte, das als Diskussionsgrundlage und Ideenaustausch diente. Dieses Dokument wurde gründlich überprüft und mit Anmerkungen, Vorschlägen und Ergänzungen bereichert. Die Zusammenarbeit mit der ODTA und die Beteiligung an der Entwicklung der Datenmodelle sind von besonderer Bedeutung, da dies die Möglichkeit bietet, die im Projekt gesammelten Praxiserfahrungen in die Diskussion einzubringen. Die Datenmodelle der ODTA finden sich auf <https://www.sti2.org/>.
- **Smart Data Models:** Eine weitere Sammlung für Datenmodelle sind die Smart Data Models der FIWARE Foundation, einer Non-Profit-Organisation, die sich unter anderem mit Datenstandards befasst. Die Verwendung der Smart Data Models von FIWARE bietet sich für LAB-TOUR SH besonders an, da bereits den Context Broker von FIWARE verwendet wird. Durch diese Synergie wird von einer nahtlosen Integration der Datenmodelle in die bestehende Infrastruktur profitiert. Die Smart Data Models zeichnen sich durch ihre ausgezeichnete Qualität und

Harmonisierung miteinander aus. Sie garantieren eine konsistente und umfangreiche Datenmodellierung, die den speziellen Anforderungen von LAB-TOUR SH gerecht wird. Mit ihrer breiten Palette an Modellen für unterschiedlichste Anwendungsbereiche, darunter Smart Cities, Industrie 4.0, Energie und Umwelt, stellen sie eine vielfältige Ressource dar. Diese Diversität ermöglicht die Auswahl relevanter Modelle für spezifische Vorhaben und deren nahtlose Integration in die entsprechenden Anwendungen. Ein weiterer Vorteil der Nutzung der Smart Data Models von FIWARE besteht in dem umfangreichen FIWARE-Ökosystem. FIWARE hat eine große und aktive Community, die kontinuierlich an der Weiterentwicklung der Datenmodelle arbeitet und zusätzliche Ressourcen und Tools zur Verfügung stellt. Dies ermöglicht einen einfachen Zugang zu bewährten Praktiken und nützlichen Erweiterungen, die bei der Implementierung und Nutzung der Datenmodelle unterstützen.

Es ist daher eine umfangreiche Auswahl an Datenmodellen verfügbar, die den Großteil der erforderlichen Attribute abdecken. Dabei wird im Einzelfall unterschieden, ob ein Datenmodell von schema.org, ein Smart Data Model oder eines der ODTA die beste Wahl ist. Beispielsweise wurde für die Ausflugsziele das Smart Data Model *Point of Interest* (POI) verwendet. Es bietet einerseits einen hohen Detailgrad, um alle benötigten Eigenschaften abzubilden, andererseits ist es allgemein genug, um alle Lokaltäten von Interesse zu haben. Im Gegensatz dazu unterscheiden die entsprechenden Modelle der ODTA zwischen beispielsweise *Beach*, *Aquarium* oder *IceCreamShop*, mit jeweils unterschiedlichen Attributen. Diese zusätzliche Genauigkeit wird im vorliegenden Fall nicht benötigt und würde das System unnötig kompliziert machen. Außerdem reicht die Datentiefe der TA.SH oftmals nicht aus, um diese sehr spezifischen Attribute im Datenmodell auszufüllen, die in den Modellen der ODTA zu finden sind. Ein Beispiel ist das Attribut „smokingAllowed“, das sich unter anderem im Typ „Beach“ findet. Ob beispielsweise am Strand das Rauchen erlaubt ist oder nicht ist oft nicht bekannt und auch für das Projekt nicht von Relevanz. Daher wurde die Entscheidung für den generischen POI getroffen.

Nach sorgfältiger Sichtung der vorhandenen Datenmodelle wurde festgestellt, dass diese nicht alle benötigten Attribute berücksichtigen. Daher wurde im Verlauf des Projektes eine kleine Anzahl neuer Datenmodelle entwickelt. Diese orientieren sich an den etablierten Standards, wie sie auch bei schema.org oder den Smart Data Models zur Anwendung kommen. Diese Datenmodelle wurden in enger Abstimmung mit der Arbeitsgruppe an der ODTA entwickelt und fügen sich nahtlos in die bestehende Modell-Landschaft ein. Im nächsten Abschnitt werden diese neuentwickelten Datenmodelle vorgestellt.

4.3.3 Technische Details

Die Art der Dokumentation orientiert sich stark an den Datenstandards von Schema.org. Die ODTA, die Datenstandards für den Tourismus als Domain Specification entwickelt, nutzt diese Art der Dokumentation in ihrer eigenen Arbeit und ist mit den Begrifflichkeiten und der Handhabung vertraut. Während der Projektlaufzeit wurden die in Tabelle 11 dargestellten Datenmodelle definiert. Diese im Projekt LAB-TOUR SH neu entwickelten Datenmodelle basieren auf den drei genannten Ontologien. Zusätzlich wurden die entwickelten Schemata bei den Smart Data Models eingereicht, damit sie dort akzeptiert und von anderen genutzt werden können. Der Link zu dieser Sammlung findet sich hier: <https://github.com/LAB-TOUR-SH/dvm>

Datenmodell	Beschreibung
Occupancy	Ein atomarer Auslastungswert. Kann verwendet werden, um die Auslastung einer Einheit zu einem einzelnen Zeitpunkt zu beschreiben. Diese Auslastung kann von einem Sensor gemessen oder durch ein Modell vorhergesagt sein.
OccupancySeries	Eine Reihe von Auslastungswerten für eine bestimmte Zeitspanne. Kann zur Beschreibung des Flusses oder der absoluten Auslastung einer bestimmten Einheit (z. B. eines Parkplatzes oder eines Strandes) verwendet werden.
TriggerRecommendation	Datenmodell für die Auslösung der Erstellung einer Empfehlung. Dies ist eine Anfrage an die Empfehlungsmaschine zur Erstellung einer Empfehlung für einen bestimmten Kontext.
TouristRecommendation	Eine Empfehlung, die für einen bestimmten Ort erstellt wurde. Die Empfehlung, die auf einem Ziel basiert, enthält eine Liste der empfohlenen Ziele und eine Liste der empfohlenen Parkplätze.

Tabelle 11: Im Projekt definierte Datenmodelle

Gemeinsame Attribute

Es gibt grundlegende Attribute, die in allen definierten Datenmodellen vorkommen (siehe Tab. 12- 16). In der Spalte „Kardinalität“ ist definiert, welche Felder Pflicht und welche optional sind. Eine Kardinalität von 1 definiert ein Feld als obligatorisch. Eine Kardinalität von 0..1 stellt ein optionales Feld dar, wobei es entweder keinen Wert (0) oder einen Wert (1) beinhalten kann. Schließlich lässt eine Kardinalität von 0..N zu, dass das Feld eine beliebige Anzahl von Werten annehmen kann, einschließlich keinem (0).

Attribut	Erwarteter Typ	Beschreibung	Kardinalität
id	Text	Aus dem Vokabular: Die Eigenschaft „id“ (eindeutige Kennung) steht für jede Art von Bezeichner für jede Art von Ding, wie ISBNs, GTIN-Codes, UUIDs usw. Schema.org bietet spezielle Eigenschaften für die Darstellung vieler dieser Kennungen, entweder als Textstrings oder als URL (URI)-Links.	1
dateCreated	DateTime	Das Datum, an dem die Entität erstellt wurde. Dieses Datum wird normalerweise von der Speicherplattform zugewiesen.	0..1
dateModified	DateTime	Das Datum, an dem die Entität zuletzt geändert wurde. Dieses Datum wird normalerweise von der Speicherplattform zugewiesen.	0..1
source	URL	URL der Quelle dieser Entität.	0..1
name	Text	Der Name der Entität.	0..1
alternateName	Text	Alternativer Name für die Entität, falls verfügbar.	0..1
description	Text	Eine Beschreibung der Entität.	0..1
dataProvider	Text	Name des Unternehmens, der Person oder der Organisation, die diese Entität bereitstellt.	0..1

owner	Text	Kennung des Eigentümers dieser Entität.	0..N
seeAlso	Text	Liste der zusätzlichen Ressourcen über die Entität.	0..N
location	GeoCoordinates	Geokoordinaten dieser Entität	1
address	PostalAddress	Physische Adresse der Entität.	0..1
areaServed	Text	Das geografische Gebiet, in dem eine Dienstleistung oder ein angebotener Artikel erbracht wird	0..1
measurementType	Text	Art der Messung. Z. B. beobachtet, vorhergesagt usw.	0..1
type	Text	NGSI Typ der Entität.	1

Tabelle 12: Gemeinsame Attribute in allen definierten Datenmodellen

[Occupancy](#)

Attribut	Erwarteter Typ	Beschreibung	Kardinalität
measurementType	Text	Art der Messung. Z. B. beobachtet, vorhergesagt usw.	0..1
period	Text	Periode der diese Auslastung zugeordnet ist. Kodiert als Dauer gemäß ISO8601.	0..1
refBasis	Text	Kennung, auf den sich diese Auslastung bezieht.	0..1
relativeValue	Number	Relativer Wert der Auslastung, bezogen auf die maximale Kapazität dieser Entität.	0..1

validFrom	DateTime	Datum und Uhrzeit, ab der die Entität gültig ist.	0..1
validTo	DateTime	Datum und Uhrzeit, bis zu der die Entität gültig ist.	0..1
value	Integer	Der Zahlenwert der Auslastung.	0..1

Tabelle 13: Attribute im Occupancy Datenmodell

OccupancySeries

Attribut	Erwarteter Typ	Beschreibung	Kardinalität
index	DateTime	Alle Zeitpunkte, die mit den Werten der Reihe verknüpft sind.	0..N
refBasis	Text	Kennung, auf den sich diese Auslastung bezieht.	0..1
relativeValue	Number	Relativer Wert der Auslastung, bezogen auf die maximale Kapazität dieser Entität.	0..N
validFrom	DateTime	Datum und Uhrzeit, ab der die Entität gültig ist.	0..N
validTo	DateTime	Datum und Uhrzeit, bis zu der die Entität gültig ist.	0..N
value	Integer	Die Zahlenwerte der Auslastungsserie.	0..N

Tabelle 14: Attribute im OccupancySeries Datenmodell

TriggerRecommendation

Attribut	Erwarteter Typ	Beschreibung	Kardinalität
maxDistance	Number	Maximale Entfernung der zu berücksichtigenden Entitäten.	0..1
refBasis	Text	Falls bekannt, Kennung der Entität, auf die sich diese Empfehlung bezieht.	0..1

Tabelle 15: Attribute im TriggerRecommendation Datenmodell

TouristRecommendation

Attribut	Erwarteter Typ	Beschreibung	Kardinalität
refBasis	Text	Kennung der Entität, auf die sich diese Empfehlung bezieht.	0..1
refRecommendedParking	Text	Kennung der empfohlenen Parkplätze, geordnet nach Wichtigkeit in absteigender Reihenfolge.	0..N
refRecommendendenPoi	Text	Kennung der empfohlenen POI, geordnet nach Wichtigkeit in absteigender Reihenfolge.	0..N
validFrom	DateTime	Datum und Uhrzeit, ab der die Entität gültig ist.	0..1
validTo	DateTime	Datum und Uhrzeit, bis zu der die Entität gültig ist.	0..1

Tabelle 16: Attribute im TouristRecommendation Datenmodell

4.3.4 Learnings und Potentiale

Während der Bearbeitung dieses Projektes traten eine Reihe unerwarteter Herausforderungen bei der Beschreibung der Daten auf. Diese beziehen sich auf die folgenden Aspekte:

- Die Taxonomie, die existierte, um den Typ eines POI nach Smart Data Model anzugeben, ist für die touristischen Attraktionen in Schleswig-Holstein unzureichend. Zudem wurde hier in den

letzten drei Jahren kein Typ mehr ergänzt, was letztendlich während des Projekts zu einer Stilllegung der Taxonomie geführt hat. Eine derartige Abhängigkeit wurde als nicht zukunftssicher identifiziert und bedarf einer sichereren Lösung.

- Die Form der Dokumentation bei schema.org ist stark verallgemeinert, sodass tourismusspezifische Aspekte nur ungenügend beschrieben werden können.
- Die Arbeitsgruppe ODTA entwickelt deshalb eine Domain Specification für den Tourismus auf Basis von den Datenmodellen von schema.org. Im Rahmen dieses Projektes konnten Use Case spezifische Aspekte zum Thema digitales Besuchermanagement beschrieben werden und bieten der ODTA damit Ansatzpunkte zur Weiterarbeit.
- OpenAPI erlaubt es, API-Endpunkte zu beschreiben. Über diese Endpunkte erwartete Daten sind im Optimalfall in einem standardisierten Datenformat, wie beispielsweise von schema.org. Eine Beschreibung der Datenmodelle im OpenAPI Format wäre demnach eine große Hilfe für die Entwicklung von Schnittstellen.
- Eine weitere Schwierigkeit liegt in der Definition des Attributes „Auslastung“. Diese kann als prozentualen Werten angegeben werden, oder in absoluten Besucherzahlen. Aus diesem Grund beinhalten die oben definierten Datenmodelle beide Werte. Bei einigen POI können Auslastungen von über 100% vorkommen. Bei einem Parkplatz mit einer festen Anzahl von Parkplätzen deutet eine Auslastung von über 100% beispielsweise auf einen Messfehler hin. Ein Strandabschnitt dagegen kann kurzfristig mehr Besucher*innen aufnehmen als vorgesehen. Eine in diesem Zusammenhang noch ungelöste Frage ist, wie die Auslastung bei Fahrrad- oder Wanderrouten angegeben werden kann. Hier sind weitere Forschungsarbeiten nötig.
- Als besonders herausfordernd gestaltete sich die Berechnung der Auslastung aus den Sensordaten. Die in diesem Projekt genutzten Sensordaten geben immer nur Werte über eingehende und ausgehende Personen, PKW oder andere Klassen weiter. Ein voller oder leerer POI ließe sich aber nur mit einer festen Auslastungszahl in Bezug zur maximalen Kapazität beschreiben. Auf diesen Punkt wird in Abschnitt 5 gesondert eingegangen.

4.4 Integrationsplattform

4.4.1 Motivation

Die Integrationsplattform bildet das Herzstück der IT-Infrastruktur. Hier werden die Daten aus den Datenhubs zentral gesammelt. Sie bilden die Grundlage für die ebenfalls in der Integrationsplattform erfolgende Auslastungsprognose sowie die Alternativengenerierung. Außerdem werden aus der Integrationsplattform die Empfehlungen per Webservice nach außen veröffentlicht. Es ergeben sich demnach vier Aufgabenfelder für die Integrationsplattform.

1. **Datenimport:** Aktuelle Daten müssen in Echtzeit von den Datenhubs bzw. den externen Datenanbietern importiert werden. Diese Aufgabe wurde bereits in Abschnitt 4.2 erläutert.
2. **Datenhaltung:** Sowohl historische als auch aktuelle Daten werden auf der Integrationsplattform dauerhaft gespeichert. Sie bilden die Datengrundlage für das Trainieren von Vorhersagemodellen. Es sollte ferner die Möglichkeit geben, aktuelle und prognostizierte Auslastung sowie die generierten Alternativen von außen abzurufen.
3. **Auslastungsprognose und Alternativengenerierung:** Auf Grundlage der historischen und aktuellen Daten wird eine Auslastungsprognose für alle POI erstellt. Die aktuellen und prognostizierten Auslastungen bilden im zweiten Schritt die Basis für die Generierung von Vorschlägen für alternative touristische Ziele. Auf diese Punkte wird in Abschnitt 5 detailliert eingegangen.

4. **Webservice:** Die aktuellen und zu erwartenden Auslastungen, gemeinsam mit den generierten Alternativen sowie ggf. weitere Informationen über die POI sollen per Webservice nach außen verfügbar gemacht werden. Der Webservice kann dann von einem Frontend genutzt werden, um die Prognosen und Vorschläge an die Tourist*innen auszuspielen.

4.4.2 Anforderungen

Aus den im vorigen Abschnitt definierten Aufgaben ergeben sich die technischen Anforderungen. Um sicherzustellen, dass die Integrationsplattform alle Aufgaben erfolgreich bewältigen kann, müssen diese erfüllt werden:

1. **Persistenz:** Es ist erforderlich, dass die Daten langfristig und persistent gespeichert werden und jederzeit verfügbar sind. Dadurch wird gewährleistet, dass kontinuierlicher Zugriff auf die Daten möglich ist.
2. **Verfügbarkeit:** Eine hohe Verfügbarkeit, insbesondere des Webservices, aber auch der anderen internen Komponenten, ist von großer Bedeutung. Die Plattform muss rund um die Uhr verfügbar sein, um einen kontinuierlichen Zugriff auf die Daten zu gewährleisten. Es sollte eine zuverlässige Infrastruktur und Maßnahmen zur Fehlererkennung und Ausfallsicherheit bieten, um mögliche Unterbrechungen zu minimieren.
3. **Standards:** Die Schnittstelle nach außen, insbesondere der Webservice, sollte den gängigen Industriestandards entsprechen. Dies beinhaltet die Unterstützung gängiger Datenformate wie JSON oder XML, um eine reibungslose Integration mit anderen Systemen zu gewährleisten.
4. **Skalierbarkeit:** Eine hohe horizontale Skalierbarkeit ist erforderlich, insbesondere für einzelne Komponenten wie den Recommender. Dies ermöglicht es, wachsende Benutzeranforderungen effizient zu bewältigen. Die Plattform sollte in der Lage sein, sowohl kleine Anfragen als auch große Lastspitzen zu verarbeiten, ohne dabei die Performance einzuschränken.
5. **Performance:** Die Plattform sollte eine hohe Leistung bieten, um schnelle und effiziente Datenübertragungen zu ermöglichen. Dies beinhaltet eine minimale Latenzzeit sowie eine schnelle Verarbeitungsgeschwindigkeit der Anfragen.
6. **Sicherheit:** Es ist wichtig, die Endpunkte der Plattform gegen unbefugten Zugriff abzusichern, um die Integrität und Vertraulichkeit der Daten zu gewährleisten.
7. **Erweiterbarkeit:** Das Gesamtsystem sollte von Anfang an so konzipiert werden, dass es im Nachhinein problemlos durch beliebige weitere Dienste oder Komponenten erweitert werden kann. Dadurch wird eine flexible Anpassungsfähigkeit an zukünftige Anforderungen ermöglicht.

Zusätzlich zu den technischen Anforderungen müssen auch Ressourcen und Rahmenbedingungen des Projektes berücksichtigt werden. Die folgenden Anforderungen sind dabei von besonderer Bedeutung:

1. **Vermeidung eines Vendor Lock-In:** Es soll maßgeblich auf Open-Source-Lösungen gebaut werden, um einen Vendor Lock-In zu verhindern. Dies garantiert die Fortführung des Projektes mit einer plattformunabhängigen Lösung und verhindert teure Lizenzgebühren im Anschluss an das Projekt.
2. **Minimierung des Entwicklungsaufwands:** Wenn möglich, werden bereits bestehende Lösungen bevorzugt, um den Entwicklungsaufwand zu minimieren und eine schnellere Umsetzung des Projektes zu ermöglichen.
3. **Modularer Aufbau:** Um den Anforderungen wie Skalierbarkeit, Verfügbarkeit, Performance und Erweiterbarkeit gerecht zu werden, sollte die Integrationsplattform intern über einen modularen Aufbau verfügen. Eine Microservice-Architektur wird dabei bevorzugt, um eine flexible und effiziente Integration verschiedener Komponenten zu ermöglichen.

4.4.3 Umsetzung

Die Integrationsplattform wurde im Rahmen des Projektes prototypisch umgesetzt. Das Ziel war dabei die Entwicklung eines lauffähigen Prototyps als *PoC*. Dadurch lassen sich Herausforderungen erkennen, die einer produktiven Nutzung im Wege stehen könnten und gegebenenfalls Lösungen entwickeln. Der vorliegende Abschnitt bietet einen detaillierten Einblick in die Architektur der Integrationsplattform, mit besonderem Fokus auf dem Zusammenspiel ihrer einzelnen Bestandteile. Die technische Ausführung wird anschließend ausführlicher in Abschnitt 4.4.4. diskutiert. Abbildung 52 veranschaulicht die Struktur der Integrationsplattform, unter Einbeziehung der eingesetzten Softwarelösungen. Die Plattform ist in einem eigens konzipierten Cluster implementiert, das speziell für LAB-TOUR SH entwickelt wurde. Hierbei wird OpenStack verwendet, um die diversen Server zu virtualisieren. OpenStack ist ein Softwareprojekt, das eine freie Architektur für Cloud-Computing zur Verfügung stellt. Dieser Ansatz eröffnet die Möglichkeit, sowohl virtuelle Maschinen als auch Kubernetes-Cluster zu erstellen. Bei der Realisierung wurde besonderes Augenmerk auf die Einhaltung von Best-Practice-Standards gelegt. Aus diesem Grund wurden die Datenbanken außerhalb des Kubernetes-Clusters auf dedizierten virtuellen Maschinen installiert. In ähnlicher Weise wurde der Central Logger in einer separaten virtuellen Maschine konfiguriert. Zusätzlich zu diesen Maßnahmen wurden zwei Kubernetes-Cluster eingerichtet: Ein Cluster fungiert als Betriebsumgebung für die restlichen Softwarekomponenten der Integrationsplattform, während das zweite ausschließlich in Verbindung mit Dask genutzt wird.

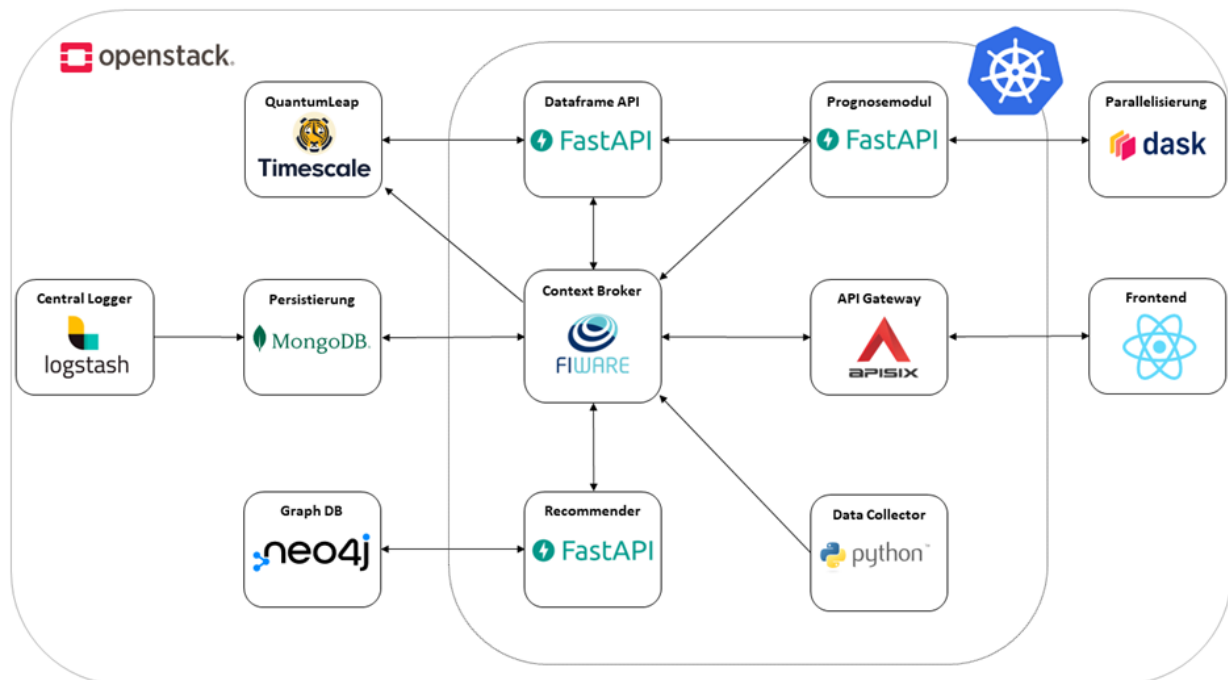


Abbildung 52: Gesamtübersicht über die Architektur der Integrationsplattform

Im Folgenden werden die jeweiligen Aufgaben der einzelnen Komponenten kurz vorgestellt.

1. **Context Broker:** Der Context Broker ist der Dreh- und Angelpunkt der Integrationsplattform. Fast alle Daten laufen in dem Projekt über diesen zentralen Knotenpunkt und werden hier für die verschiedenen Komponenten zwischengespeichert. Ebenso werden die Daten, welche nach außen freigegeben werden, auf dem Context Broker vorgehalten.

2. **Data Collector:** Der Data Collector besteht aus der Gesamtheit der Connectoren, die Daten von den Datenhubs oder den externen Datenquellen abrufen. Die gesammelten Daten werden auf den Context Broker gelegt, von wo aus diese für die anderen Komponenten zur Verfügung stehen.
3. **Persistierung:** In dieser Datenbank werden statische Daten abgelegt, insbesondere Daten zu den POI. Ebenso werden die Auslastungen und die Wetterdaten, welche als Zeitreihen vorliegen, in den Datenbanken persistiert. Zusätzlich werden die Daten des zentralen Loggers hier gespeichert.
4. **Central Logger:** Dieses Modul konsolidiert und formatiert die Logs, die alle Komponenten des Systems generieren. Diese Logs können zur Überwachung und Fehleranalyse genutzt werden. Sie werden zur späteren Auswertung in der Datenbank gespeichert.
5. **Graph-Datenbank:** In dieser Datenbank werden vor allem die Beziehungen zwischen den einzelnen POI gespeichert. Dies sind beispielsweise Ähnlichkeit zweier POI zueinander oder die räumliche bzw. zeitliche Entfernung zwischen zwei POI. Diese Informationen werden vor allem vom Recommender verwendet und bilden die Grundlage für die spätere Empfehlung.
6. **Dataframe API:** Die Dataframe API dient der Aufbereitung und Zusammenführung der Daten, wie diese im Prognosemodul benötigt werden. Dazu werden aus den verschiedenen Daten aus den Datenbanken und von dem Context Broker Datensätze generiert, welche zur Vorhersage, zum Training und zum Tuning der Prognosemodelle benötigt werden.
7. **Prognosemodul:** Dieses Modul erstellt Auslastungsprognosen auf der Grundlage von historischen und aktuellen Daten. In regelmäßigen Abständen erfolgt ein Nachtrainieren der Modelle, die das Modul zur Prognose nutzt. Dieses Modul wird in Abschnitt 5.1 ausführlich erläutert.
8. **Parallelisierung:** Das besonders rechenintensive Tuning der Modelle, die das Prognosemodul nutzt, erfolgt in einem weiteren orchestrierten Cluster. Dieser ist darauf optimiert, hochgradig parallel zu arbeiten, um so die vorhandenen Ressourcen optimal auszunutzen.
9. **Recommender:** Im Recommender werden, basierend auf aktueller und prognostizierter Auslastung der POI, Alternativvorschläge generiert. Dabei werden räumliche und thematische Ähnlichkeiten berücksichtigt, die in der Graph-Datenbank gespeichert sind.
10. **API Gateway:** Das API Gateway dient als Sicherheitsmechanismus für alle Zugriffe, die von außerhalb der Infrastruktur kommen. Dabei erfüllt das Gateway zwei Funktionen: Erstens übernimmt das Gateway die Aufgabe der Authentifizierung externer Anfragen. Es überprüft, ob eine eingehende Anfrage eine berechtigte Berechtigung zur Interaktion mit den internen Ressourcen hat. Zweitens agiert das Gateway als zentraler Zugangspunkt für alle externen Zugriffe. Es ermöglicht somit den sicheren und kontrollierten Zugriff auf interne Komponenten. Diese duale Funktion des API Gateways stellt sicher, dass die Sicherheit der Infrastruktur gewahrt bleibt, während gleichzeitig eine notwendige Interaktion mit externen Systemen ermöglicht wird.
11. **Frontend:** Das Frontend dient schließlich der Kommunikation mit den Endnutzer*innen. Hier werden die Informationen grafisch aufbereitet den Nutzer*innen präsentiert. Dabei ist das Frontend hier nur als ein Prototyp gedacht, um zu demonstrieren, wie eine Ausspielung der Daten erfolgen kann. Die Implementierung und Nutzung der Daten erfolgt im Idealfall durch die Plattformen der einzelnen Tourismusagentur.

4.4.4 Technische Details

Während der Fokus des vorherigen Abschnittes auf den Aufgaben der einzelnen Komponenten lag, soll in diesem Abschnitt auf die technische Umsetzung eingegangen werden. Dabei werden die verwendeten technischen Frameworks vorgestellt sowie Details der Implementierung erläutert.

1. **Cluster:** Der eigens für dieses Projekt aufgebaute Cluster besteht hauptsächlich aus zwei „Dell PowerEdge R750“ Servern, die als Compute Nodes in der OpenStack Cloud die benötigte Rechenleistung zur Verfügung stellen. Zusätzlich wurden zwei „Dell Smart Value PowerEdge R750“ Server bestellt, welche als Storage Nodes in ein Ceph Cluster verbunden wurden und so die entsprechende redundante Festplattenkapazität für Anwendungen bereitstellen.
2. **Cloud:** Um die Infrastruktur, auf der die Integrationsplattform läuft, als Service zur Verfügung zu stellen, wird OpenStack verwendet. In diesem Projekt wird OpenStack zur Virtualisierung genutzt, um die vorhandenen Ressourcen zwischen den Softwarekomponenten optimal aufzuteilen. Zu diesem Zweck laufen die Komponenten entweder direkt in der Cloud auf virtuellen Maschinen oder in orchestrierten Clustern, die wiederum auf der Cloud laufen.
3. **Kubernetes:** Kubernetes ist eine sogenannte Container Orchestrierungssoftware. Unter Orchestrierung wird in diesem Falle die Bereitstellung, Skalierung und Verwaltung von Containeranwendungen verstanden. Die Verwendung von Kubernetes ermöglicht eine einfache Skalierung und erhöht bei richtiger Implementierung die Ausfallsicherheit.
4. **Dask:** Zur Parallelisierung des besonders rechenintensiven Tunings der Modelle wird Dask verwendet. Dask erlaubt das hochgradig parallele Ausführen von Python Code. Das verwendete Daskcluster läuft in einem eigenen Kubernetes-Cluster auf der OpenStack-Cloud.
5. **Context Broker:** Es wird der Orion Context Broker der FIWARE Foundation verwendet. Der Vorteil dieser Software besteht zum einen darin, dass es sich um eine Open-Source-Anwendung handelt und zum anderen, dass sie in ein gut entwickeltes Ökosystem eingebettet ist. Der Context Broker läuft im Kubernetes-Cluster.
6. **Data Collector:** Der Data Collector ist, wie oben erwähnt, die Gesamtheit der Connectoren, die für den Datenimport zuständig sind. Der Data Collector ist in Python implementiert und läuft als Container im Kubernetes-Cluster.
7. **Dataframe API, Prognosemodul, Recommender:** Diese drei Komponenten sind ebenfalls in Python implementiert und laufen als Container im Kubernetes-Cluster. Als Schnittstelle kommt bei allen drei Komponenten FastAPI zum Einsatz. Diese drei Module werden ausführlich in Abschnitt 5 vorgestellt.
8. **API Gateway:** Das auf Apache APISIX basierte API-Gateway dient dem Zugriff von außen. Dafür wurden diverse API-Endpoints definiert, beispielsweise ein GET auf die aktuelle und prognostizierte Auslastung eines POI. Das API-Gateway läuft ebenfalls innerhalb des Kubernetes-Clusters.
9. **Graph DB:** Als Graphdatenbank wird Neo4j genutzt. Diese Datenbank ist die aktuell meist verbreitete Graphdatenbank, was ein Vorteil in der Entwicklung aufgrund der großen Nutzercommunity ist. Die Datenbank läuft auf einer virtuellen Maschine in der OpenStack-Cloud.
10. **Dokumentendatenbank:** Ein Großteil der statischen Daten werden in einer MongoDB persistiert. Diese läuft ebenfalls auf einer virtuellen Maschine in der OpenStack-Cloud.
11. **Central Logger:** Zur Sicherung der Logs wird die Open-Source-Software Logstash verwendet. Logstash sammelt die Logs der auf dem Kubernetes-Cluster laufenden Applikationen ein und speichert sie auf der MongoDB ab. Der zentrale Logger läuft ebenfalls in einer virtuellen Maschine in der OpenStack-Cloud.
12. **QuantumLeap:** Zur Persistierung der Zeitreihendaten wird die Software QuantumLeap von FIWARE verwendet. Als Datenbank für diese Zeit indizierte Daten wird Timescale DB genutzt. Diese läuft auch in einer virtuellen Maschine auf der OpenStack-Cloud.

13. **Frontend:** Das Frontend wurde prototypisch in React umgesetzt.

4.4.5 Learnings und Potentiale

Im Folgenden werden die Schwierigkeiten bei der Umsetzung der Integrationsplattform besprochen. Außerdem werden Lösungsvorschläge diskutiert, wie diese Probleme bei zukünftigen Projekten gelöst oder vermieden werden können. Gerade bei dem Aufbau des Clusters kam es zu Beginn des Projektes zu teils starken Verzögerungen. Maßgeblich waren hier die von der Corona-Pandemie beeinträchtigten Lieferketten während des Projektstarts im Jahr 2022. Es wurden die bestellten Server und Netzwerkgeräte erst deutlich verspätet geliefert. Die Geräte wurden im Februar 2022 bestellt, mit einer Lieferzeit von drei Monaten ergab das einen voraussichtlichen Liefertermin im Mai 2022. Jedoch wurden erst im August, also mit dreimonatiger Verzögerung, alle benötigten Teile geliefert. Die zusätzlich benötigten Netzwerktransceiver kamen sogar erst im Oktober 2022. Daher konnte erst mit sechsmonatiger Verspätung mit der Inbetriebnahme des Clusters begonnen werden. Beim Aufbau von Server-Infrastruktur sollte genügend zeitlicher Puffer eingeplant werden. Für zeitkritische Projekte empfiehlt es sich, wenn möglich auf bereits existierende Infrastruktur zurückzugreifen.

Des Weiteren zeigten sich bei der Umsetzung der Integrationsplattform zwei mögliche Schwachstellen der Architektur. Die erste ist spezieller Natur und bezieht sich auf die vom Recommender vorgeschlagenen Alternativen. Wünschenswert ist, dass bei jedem Aufruf des Recommenders leicht unterschiedliche Alternativen generiert werden. Diese orientieren sich zwar an der thematischen und räumlichen Nähe der POI zueinander. Gleichzeitig ist aber ein Zufallselement erwünscht, so dass bei einem zweiten Aufruf innerhalb von wenigen Minuten auch unterschiedliche Alternativen angezeigt werden. Dies ist in der aktuellen Architektur mit der Ausgabe der Recommendations über den Context Broker schwer umsetzbar. Der Fokus bei der Umsetzung war auf dem Zwischenspeichern der Ergebnisse für eine erhöhte Performance, was allerdings eine Zufallskomponente ausschließt.

Ein weiterer Fokus zur Optimierung betrifft die Komplexität der Integrationsplattform. In ihrer aktuellen Ausgestaltung besteht die Plattform aus einer Vielzahl von Einzelkomponenten, was eine komplexe Kommunikationsstruktur zwischen diesen hervorruft. Eine Option zur Reduzierung der Komplexität könnte beispielsweise die Konsolidierung der Daten in weniger Datenbanken sein. Allerdings geht dieser Ansatz auch mit eigenen Herausforderungen einher, insbesondere im Hinblick auf Aspekte der Datensicherheit und der Skalierbarkeit.

4.5 Zwischenfazit

In diesem Abschnitt wurde der Weg der Daten durch die in dem Projekt entwickelte Plattform untersucht. Dabei wurde zu Beginn der Fluss der Daten von einer Vogelperspektive betrachtet, um die verschiedenen Speicherpunkte der Daten, ebenso wie die benötigten Verbindungen der einzelnen Komponenten aufzuzeigen. Anhand dieser Grundlage wurden im zweiten Abschnitt die benötigten Schnittstellen beschrieben und deren Eignung auf die Probe gestellt. Diese sind essentiell für eine effektive Kommunikation zwischen den einzelnen Komponenten im Projekt. Die beste Kommunikation ist jedoch nicht hilfreich, wenn beide Enden unterschiedliche Sprachen sprechen. Daher befasste sich der dritte Abschnitt mit Datenmodellen. Die Wichtigkeit von Datenmodellen wurde dabei erläutert, ebenso wie die verschiedenen, relevanten Player vorgestellt. Anschließend wurde eine Strategie vorgestellt, um die passenden Datenmodelle für das Projekt zu identifizieren und ein Prozess definiert für den Fall, dass kein passendes Datenmodell existiert. Die in dem Projekt neu definierten Datenmodelle wurden eingeführt und ein potentiell Vorgehen zum Einfügen der Datenmodelle in die Sammlung der großen Player diskutiert. Nach der Erläuterung der Schnittstellen und Datenmodelle wurde die

zugrundeliegende Plattform im vorangegangenen Abschnitt beschrieben. Der Fokus hier wurde auf die Integrationsplattform gelegt, da diese die Daten aus den einzelnen Datenhubs zusammenbringt und das Herzstück des Projektes ist. Die einzelnen verwendeten Komponenten, Software und Plattform wurden dabei vorgestellt. Dabei wurden die Funktionsweisen und der Nutzen der einzelnen Bausteine erörtert. Zusammenfassend soll nun noch einmal aufgezeigt werden, dass das Ziel, eine solide Datengrundlage für das Projekt zu schaffen, erfolgreich umgesetzt wurde. Dies soll anhand der am Anfang definierten Ziele (Skalierbarkeit, Hochverfügbarkeit, Persistent) passieren.

Die Plattform bietet persistente und sichere Speicherung der Daten. Dies wird über die ausfallsicheren Speicherstrukturen in dem Cluster der Fachhochschule Kiel und den einzelnen Data Hubs realisiert. Die dezentrale Speicherung der Daten in verschiedenen Data Hubs sorgt dabei für zusätzliche Sicherheit, da im schlimmsten Fall bei dem Verlust eines Datenhubs und deren Backup nur ein Teil der Daten verloren gehen würde. Die Dezentralisierung der Datenhaltung führt aber zwangsläufig dazu, dass die Daten wieder zusammengebracht werden müssen. Diese Aufgabe übernimmt die Integrationsplattform, welche dank der Implementierung in Kubernetes nicht nur ausfallsicher, sondern auch gleichzeitig skalierbar gestaltet wurde. Die gezielte Auswahl und Standardisierung der Schnittstellen und Datenmodelle im Projekt tragen ebenfalls zur Skalierbarkeit der entwickelten Plattform bei. Dies wird unterstützt durch die Entwicklung unter der Hilfe des Microservice Designpattern. Dieses Design Prinzip ermöglicht die einfache Integration neuer Datenquellen und gleichzeitig eine einfache Skalierung mit steigender Nutzerzahl der Plattform.

In Summe konnten wertvolle Erkenntnisse für den Aufbau einer skalierbaren und ausfallsicheren Dateninfrastruktur geschaffen werden. Dabei konnte nicht alles, wie ursprünglich in der Vorhabensbeschreibung gedacht, auch umgesetzt werden. Es musste an manchen Stellen nachgearbeitet werden. Jedoch sind die vorgegebenen Ziele erfüllt. Es ist also eine sehr gute Grundlage geschaffen worden, die Skalierbarkeit, Persistenz und eine hohe Verfügbarkeit vereint. Diese prototypische Architektur bietet nun eine ideale Grundlage, auf welcher die Algorithmen, welche im nächsten Abschnitt vorgestellt werden, aufbauen können.

5. Recommender

Der LAB-TOUR SH Recommender präsentiert sich als integraler Bestandteil des in diesem Bericht skizzierten Konzepts. Er leistet einen bedeutenden Beitrag sowohl zur Ausarbeitung einer landesweiten Strategie für ein digitales Besuchermanagement als auch zur nachhaltigen Förderung des Tourismus in Schleswig-Holstein. Über ihn sollen sowohl Einheimische als auch Tourist*innen über die gezielte Bereitstellung von Informationen zu Auslastungen und Überlastungen möglichst so gelenkt werden, dass sich das Besucheraufkommen an touristischen Anziehungspunkten entzerrt. Außerdem soll der Recommender bislang weniger genutzte Sehenswürdigkeiten in Schleswig-Holstein in den Fokus rücken und so die Standortattraktivität steigern. Zusammengefasst hat der Recommender damit drei Hauptziele:

- **Erhöhung der Ausflugsattraktivität:** Die Bereitstellung von Informationen zu bislang wenig genutzten touristischen Attraktionen soll Nutzer*innen des Recommenders die Vielfalt der schleswig-holsteinischen Tourismuslandschaft näherbringen. Durch die intelligente Auswahl an vorgeschlagenen POI werden vor allem solche POI vorgeschlagen, die für eine/n Nutzer*in tatsächlich relevant sind. Dadurch wird die Standortattraktivität Schleswig-Holsteins im Bereich Tourismus erhöht.

- **Reduktion von Crowding:** Durch die Entfaltung der gezielten Lenkungswirkung des Recommenders können Überbelastungen von touristischen Attraktionen in ihrer Intensität und Frequenz potenziell vermindert werden. Dadurch wird für jeden Gast ein attraktiveres Besuchserlebnis erzeugt und Anwohner*innen und die Umwelt werden entlastet.
- **Reduktion von Parksuchverkehr:** Durch die intelligente Empfehlung von Parkplätzen, beruhend auf der momentanen Auslastung und durch die Lenkungswirkung des Recommenders, soll der häufig an beliebten Ausflugszielen auftretende Parksuchverkehr verringert werden. Dadurch werden Anwohner*innen entlastet, weniger klima-, umwelt- und gesundheitsschädliche Abgase ausgestoßen und Besucher*innen die Suche nach freien Parkplätzen erleichtert.

Um die erwähnten drei Ziele des Recommenders zu erreichen, soll der Recommender auf das Prinzip des „Nudgings“ zurückgreifen. Dabei sollen Gäste und Einheimische über die Bereitstellung von Informationen befähigt werden, nachhaltigere Reiseentscheidungen zu treffen, die das touristische Aufkommen in Schleswig-Holstein langfristig entzerren. Dabei soll die gezielte Selektion der zur Verfügung gestellten Informationen dafür sorgen, dass eine gewünschte Lenkungswirkung eintritt. Die zur Verfügung gestellten Informationen sollen vom Recommender über eine digitale Schnittstelle (API) in einem standardisierten Format (Datenmodell) an unterschiedliche Ausspielungskanäle übergeben werden können. So können die bereitgestellten Informationen bedarfsgerecht abgerufen werden und zum Beispiel an einer Informationsstele oder auf einer Website eingebunden werden. Um die Fähigkeiten des in diesem Forschungsprojekt prototypisch entwickelten Recommenders sichtbar zu machen, wurde eine Website entwickelt, die eine beispielhafte Ausspielmöglichkeit für interne Zwecke einsehbar macht. Über den Recommender sollen folgende Informationen an interessierte Gäste ausgegeben werden können:

- **POI-Stammdaten:** Detaillierte Informationen über eine Vielzahl an POI in Schleswig-Holstein, sodass sich Gäste vor einem Ausflug über potenzielle Ziele informieren können. Dies beinhaltet eine Beschreibung des POI, dessen Öffnungszeiten etc. Wenn verfügbar, sollen über die prototypische Website auch Bilder des POI angezeigt werden, um einen Eindruck der örtlichen Gegebenheiten zu vermitteln.
- **Auslastungsdaten:** Auskünfte über die Live-Auslastung an allen Standorten, an denen diese über eine Messung bestimmt werden kann, oder die im Umfeld einer Messung liegen. An Orten mit installierter Sensorik wird außerdem eine Prognose über die Auslastung der nächsten vier Tage zur Verfügung gestellt, so dass frühzeitig eine Einschätzung über die Auslastung eines POI vorliegt.
- **Alternative POI:** Recommendations zu alternativen Ausflugszielen und geeigneten Parkmöglichkeiten. Die Recommendations sollen nicht nur ausgegeben werden, wenn der ursprünglich angefragte POI überfüllt ist, sondern immer. Dies dient dazu, einem interessierten Gast konstant Informationen über andere, für ihn potenziell interessante Ausflugsziele zu bieten und so eine Lenkungswirkung auch dann zu entfalten, wenn (noch) gar keine Überfüllung eingetreten ist.

Der Recommender besteht daher aus zwei Sub-Modulen. Diese sind zum einen die Generierung der Auslastungsprognose (Prognosemodul) und zum anderen die Selektion von passenden alternativen POI (Recommendermodul). Diese beiden Module werden im Folgenden detailliert beschrieben.

5.1 Auslastungsprognose

Eine der beiden zentralen Ideen in LAB-TOUR SH ist es, zu untersuchen, wie gut sich Besucherströme über den Zeitraum von wenigen Tagen in die Zukunft vorhersagen lassen. Dafür muss zunächst her-

ausgefunden werden, welche Daten für solche Vorhersagen notwendig sind und weiterhin welche Modelle sich dafür eignen. Außerdem müssen Wege gefunden werden, die im Projekt generierten Rohdaten zu veredeln und anzureichern. Die über diese Rohdaten und Modelle generierten Auslastungsprognosen sollen dann die Grundlage für den Gesamt-Recommendier bilden, da diese Vorhersagen auch für den Alternativengenerierer (siehe Abschnitt 5.3) benötigt werden. Die Fachhochschule Kiel hat ein PoC des Vorhersagemoduls entwickelt, anhand dessen Entwicklung sich viele theoretische Punkte an einem praktischen Beispiel erklären lassen. Dieser Abschnitt wird sich sowohl mit der theoretischen Konzeptionierung als auch mit der praktischen Implementierung des Prognosemoduls beschäftigen. Im nächsten Abschnitt werden zunächst alle Anforderungen aufgeführt, die an eine im touristischen Kontext genutzte Auslastungsprognose existieren. Darauf folgt eine Erläuterung der aufgestellten Datengrundlage und aller notwendigen Schritte, um diese Daten zu bereinigen. Dann wird die Logik, die hinter den Modellen der Auslastungsprognose steht, erläutert, wobei deren technische Umsetzung im PoC im Vordergrund steht. Der Abschnitt schließt ab mit einer kritischen Reflektion der Vorgehensweise und einem Ausblick.

5.1.1 Anforderungen an das Prognosemodul

Die Anforderungen, die an ein Prognosemodul und die Auspielung seiner Ergebnisse im Kontext des Schleswig-Holsteinischen Tourismus gestellt werden, setzen sich zusammen aus praktischen Anforderungen, die sich während der konzeptionellen und prototypischen Entwicklung eines eben solchen Prognosemodells in LAB-TOUR SH ergeben haben, als auch aus theoretischen Überlegungen, die aber grundlegend für ein landesweites Besuchermanagementsystem sind. Die Anforderungen wurden in Form von Schlagworten festgehalten, um eine gezielte Analyse der einzelnen Aspekte im Abschnitt 5.1.6 zu ermöglichen, ohne eine implizierte Rangfolge der Wichtigkeit vorzunehmen.

- **Robustheit gegenüber Datenlücken:** Bei Messungen von Auslastungen kann es vorkommen, dass Sensordaten nicht verlässlich übermittelt werden und so Datenlücken entstehen. Diese mangelnde Reliabilität kann grundsätzlich auf zwei verschiedene Arten vorkommen. Zum einen kann es zu Schwierigkeiten bei der Übermittlung der Daten vom Sensor hin zur Datenplattform des Sensorbetreibers oder bei der Verarbeitung auf dieser kommen. Dadurch kann sich die Lieferung von Live-Daten verzögern. Sobald diese Problematiken allerdings nicht mehr die Lieferung der Daten verhindern, was oft recht kurzfristig der Fall ist, werden die in der Zwischenzeit gemessenen Daten nachgemeldet und stehen ab dann zur Verfügung. Dies stellt zusätzlich ein Problem der zeitlichen Verzögerung (Latenz) dar. Andererseits kann es auch zu einem Ausfall oder einer Funktionsschwierigkeit des Sensors an sich kommen, was dazu führt, dass bis zur Behebung des Problems keine Daten aufgenommen werden. Da diese Daten fehlen, können sie auch nicht nachgeliefert werden. Durch diese Art von Messstörung kann es vorkommen, dass eine langfristig aufgenommene Zeitreihe der Messdaten lückenhaft ist. Ein produktiv einsetzbares Prognosemodell muss in der Lage sein, mit beiden Fehlermustern in den Daten umgehen zu können.
- **Vorhersagequalität:** Die durch ein Prognosemodul gemachten Vorhersagen müssen von ausreichend guter Qualität sein. Das Verfahren zur Evaluation der Qualität sollte dabei die spezifischen Anforderungen an den Recommender aus der Perspektive des Tourismusmanagements in Schleswig-Holstein repräsentieren. Dazu gehören vor allem die zwei folgenden Anforderungen. Erstens sollte die Evaluation vor allem solche Datenpunkte berücksichtigen, die den Haupt-Use Case des Vorhersagemoduls ausmachen. Dies sind solche Zeitpunkte, in denen die Auslastung besonders hoch ist. Da dieser Zustand weniger oft eintritt als eine sehr geringe Auslastung, die jede Nacht eintreten sollte, muss ein Weg gefunden werden, die Evaluation

bei hohen Auslastungen stärker zu gewichten. Dies bezieht sich auch auf die Jahreszeiten, bei denen gesichert sein sollte, dass die Hauptevaluation zur Hauptsaison des jeweiligen POI-Typs stattfindet. Auf der anderen Seite dürfen die Zeitpunkte besonders hoher Auslastungen in der Evaluation nicht exklusiv sein, damit gewährleistet wird, dass das Prognosemodul in der Lage ist, die generelle Dynamik an einem POI zu verstehen, und zum Beispiel nachts nur sehr niedrige Auslastungen vorhersagt. Die zweite Anforderung an die Evaluation der Vorhersagequalität ist, dass eine Abweichung der Prognose vom tatsächlichen Wert unterschiedlich stark gewichtet werden sollte, je nachdem ob diese Abweichung zu einer größeren oder kleineren Prognose führt. Da das grundlegende Ziel des hier zu entwickelnden Recommenders darin besteht Crowding zu verhindern, sollte das Prognosemodul in seinem Fehler das Tourismusaufkommen eher ein wenig überschätzen als unterschätzen. Dadurch kann eine Art Puffer eingeführt werden, die dem Crowding zusätzlich zu allen weiteren Maßnahmen im Recommender entgegenwirken soll. Zudem sollte die Evaluation der Vorhersagequalität klar definiert sein und regelmäßig für alle Standorte bestimmt werden.

- **Cold-Start Mitigation:** Viele Machine-Learning Algorithmen haben das Problem, dass sie zunächst eine signifikante Menge an Trainingsdaten benötigen, bevor sie qualitativ hochwertige Aussagen treffen können. Dieses Problem wird grundsätzlich unter dem Titel „Cold-Start“ geführt (Moon et al., 2020), auch wenn diese Namensgebung vor allem durch den Bereich der Recommendation Systems geprägt ist. Das Problem des Cold-Starts könnte in einem Vorhersagemodul ebenfalls auftreten, wenn nicht geeignete Schritte getroffen werden, um dem zu begegnen. Neu an das System angeschlossene Sensoren sollen also möglichst schnell mit aussagekräftigen Prognosen bedient werden können. Nur so kann ein starker Anreiz für Kommunen gesetzt werden, in lokale Sensorik zu investieren und ihre Daten in einen Landes-Daten-Hub einzuspeisen, aus dem die Prognosen gebildet werden.
- **Skalenflexibilität:** Um einen möglichst großen Pool an Daten verarbeiten zu können, sollte das Prognosemodul in der Lage sein, mit Daten umzugehen, die auf verschiedensten zeitlichen Auflösungen beruhen. Je nach Art der Sensorik kann diese von einigen Minuten bis zu einigen Stunden variieren. Vorstellbar ist zum Beispiel, dass Kassensysteme ihre Daten an die Integrationsplattform nicht automatisiert übermitteln, sondern nur, wenn dies händisch initiiert wird. Solche Daten würden dann lediglich in einer geringeren und zudem nicht äquidistanten Auflösung vorliegen. Grundsätzlich scheint es allerdings auch sinnvoll, einige Mindestanforderungen an die Regelmäßigkeit von Datenlieferungen zu stellen. Basierend auf den Erfahrungen, die während der LAB-TOUR SH gesammelt wurden, sollten die gelieferten Daten in einer Weise strukturiert sein, um zumindest die zeitlichen Bereiche des Vormittags (8:00-11:00 Uhr), Mittags (11:00-14:00 Uhr) und Nachmittags (14:00-17:00 Uhr) separat zu beschreiben. Nur so kann ein Tagesgang interpoliert werden und damit das Prognosemodul eine einheitliche, auf einer stündlichen Auflösung beruhenden Prognose machen.
- **Aktualität:** Wenn Live-Messungen von Auslastungen an einem Standort vorhanden sind, sollten die Prognosen für diesen Standort möglichst stets auf Messungen beruhen. Ein solches Verfahren begünstigt eine hohe Qualität der Prognosen im Vergleich zu einer reinen Vorhersage der Auslastung anhand externer Faktoren oder Prämissen.
- **Effizienz:** In einem Projekt, das darauf ausgelegt ist, die Nachhaltigkeit des schleswig-holsteinischen Tourismus insgesamt zu steigern, sollten die verwendeten Algorithmen innerhalb des Prognosemoduls so effizient und energiesparend wie möglich sein. Dafür sollten Ansätze der Green-IT und effizienten Hyperparameteroptimierung angewandt werden. Außerdem sollte in allen Implementierungsentscheidungen der Trade-Off zwischen Genauigkeit von Vorhersagen

und Ressourcenaufwand berücksichtigt werden. Besonders sollte daher angestrebt werden, das aufwändige Tuning und Training der Vorhersagemodelle so selten wie möglich und nur so oft wie nötig durchzuführen.

- **Nachsteuerungsfähigkeit:** Um das Vorhersagemodul fortlaufend evaluieren und optimieren zu können, sollte es über die Fähigkeit verfügen, rückgemeldete Einzelobservationen in das Training von Modellen mit einzubeziehen und diese entsprechend stärker zu gewichten. So können Touristiker die tatsächliche, vor Ort beobachtete prozentuale Auslastung rückmelden und damit das Prognosemodul aktiv verbessern.
- **Verknüpfung:** In LAB-TOUR SH wurden Frequenzmessungen vor allem an Parkplätzen in touristisch relevanten Regionen durchgeführt. Außerdem wurden WLAN-Access-Points als Sensorik genutzt und Messungen an frequentierten Radwegen vorgenommen. Alle diese Standorte haben gemeinsam, dass sie nicht ohne weiteres einem touristischen POI zugeordnet werden können. Im LAB-TOUR SH Recommender ist allerdings vor allem die Auslastung genau dieser von Relevanz. Hier soll ein/e Nutzer*in sich vor allem für Strände, aber auch für andere POI-Typen informieren können, wie voll diese sind und werden. In einem Prognosemodul, das produktiv eingesetzt werden soll, muss also eine Verknüpfung von Messstationen zu touristischen POI stattfinden. Dabei kann es auch sein, dass ein POI von mehreren Messstationen beeinflusst werden sollte. Um dieses Thema zu verdeutlichen, soll folgendes Beispiel herangezogen werden: Ein Parkplatz mit 100 Stellplätzen, der mit Sensorik ausgestattet ist, sei so gelegen, dass er in unmittelbarer Nähe eines Strandes (200 m) und eines Spielplatzes (400 m) und in etwas weiterer Entfernung zu zwei Cafés (800 m) und einem Museum (600 m) liegt. Ein Café liegt dabei weiter landeinwärts und ein Café weiter in der Nähe des Strandes. Weiterhin hat das Museum einen eigenen Parkplatz, der allerdings nicht durch Sensorik erfasst wird. Aus der gemessenen Auslastung des Parkplatzes Aussagen darüber zu treffen, welcher dieser POI wie voll ist, ist nicht trivial. So ist im Sommer denkbar, dass eine Auslastung des Parkplatzes ungefähr proportional zur Auslastung des Strandes sein könnte, wobei wahrscheinlich eine Proportionalitätskonstante kleiner als null existieren würde. Dem liegt die Annahme zugrunde, dass der Strand vermutlich nicht überfüllt ist, sobald der Parkplatz voll ist. Dies scheint abhängig von den örtlichen Gegebenheiten. Außerdem ist es wahrscheinlich, dass das Café in der Nähe des Strandes stärker ausgelastet ist als das Café landeinwärts und das Museum annähernd komplett unberührt von der Auslastung des Parkplatzes bleibt. Diese Tendenzen würden sich im Winter noch verstärken, da dann beispielsweise davon ausgegangen werden kann, dass der Strand nicht überfüllt sein wird, selbst wenn 100 Fahrzeuge auf dem Parkplatz stehen und ihre Insassen am Strand spazieren sind. Daher sollte im Prognosemodul im Zweifelsfall ein POI nicht mit einer Messstation verknüpft werden, wenn nicht eine klare Funktion aufgestellt werden kann, die die Auslastung der Messstation auf die Auslastung des POI projizieren kann.
- **Verfügbarkeit:** Eine aktuelle Vorhersage muss von einem guten Prognosemodul schnell und ohne große Ladezeiten verfügbar sein. Nur dann kann ein externer Service wie zum Beispiel ein Web-Frontend diese sinnvoll einbinden. Außerdem müssen die Vorhersagen auch intern für die Weiterverarbeitung im Recommender schnell zur Verfügung stehen, um möglichst alle Informationen immer auf dem neuesten Stand zu halten.
- **Prozentuale Auslastungen:** Die von einem Prognosemodul ausgegebenen Werte sollten Auslastungen repräsentieren und dementsprechend schon durch die Kapazität eines POI skaliert worden sein. Die zurückgegebenen Werte können demnach als prozentuale Auslastung verstanden werden können. Diese ist limitiert durch 0% Auslastung am unteren Ende der Skala und hat nach oben keine Grenze, da Werte größer als 100% Auslastung bei Überbelastungen

durchaus vorkommen können. So müssen nachfolgende Applikationen keine Information über die Kapazität der Messstationen und POI haben, sondern können sofort mit den prozentualen Werten arbeiten.

- **Vorhersage-Horizont:** Eine ausgespielte Prognose sollte so weit in die Zukunft reichen, dass damit bereits deutlich im Voraus Ausflüge geplant werden können. Gleichzeitig sollte die Vorhersagequalität auch für den am weitesten entfernten Zeitpunkt ausreichend sein. Der optimale Vorhersage-Horizont eines produktiven Prognosemoduls sollte dafür genauer evaluiert werden, realistisch scheint allerdings etwas zwischen 3-7 Tagen als sinnvoller Zeitrahmen. Gleichwohl sind auch Szenarien denkbar, die eine längerfristige Prognose nötig machen, beispielsweise bei der frühzeitigen Planung für ein jährlich stattfindenden Events.
- **Anpassungsfähigkeit:** Das Prognosemodul sollte so konzipiert und technisch umgesetzt werden, dass es fähig ist, divergierende Besucheraufkommensdynamiken an verschiedenen Standorten zu erfassen und adäquat in die Zukunft zu projizieren. So könnte es zum Beispiel vorkommen, dass einige Orte bei besonders günstigen Wetterbedingungen ein hohes Besucheraufkommen verzeichnen, während andere Standorte gerade bei unvorteilhaften Wetterverhältnissen stärker frequentiert werden. Darüber hinaus kann das Besucheraufkommen standortabhängig einem klaren Jahresrhythmus folgen, wie es beispielsweise bei Stränden mit einem Maximum im Sommer der Fall ist. Es gibt jedoch auch Orte, die keinen ausgeprägten saisonalen Verlauf aufweisen. Das Prognosemodul muss daher in der Lage sein, automatisch auf diese unterschiedlichen Muster zu reagieren. Es darf nicht von einer bestimmten Art von Dynamik ausgehen, sondern muss flexibel und adaptiv sein, um den vielfältigen und sich ständig ändernden Anforderungen gerecht zu werden.
- **Umfassende Datengrundlage:** Um die örtlichen Gegebenheiten und Randbedingungen von POI möglichst gut abzubilden, sollte die Datengrundlage für verschiedene POI flexibel gehalten werden. So können theoretisch für Strände, an denen Surfen erlaubt ist, genauere Informationen über Wind- und Wellenverhältnisse in die Datengrundlage aufgenommen werden sowie genauere Informationen über Schneegrenze und -beschaffenheit in Skigebieten. Außerdem können so, falls vorhanden, Informationen zur Auslastung wichtiger Zubringerstraßen in die betreffende Region als Feature mit in die Vorhersagen integriert werden. Ein solches Verfahren erlaubt es, möglichst gut Use Case-spezifische Dynamiken eingehen zu können.
- **Einheitlichkeit:** Das Format der zurückgegebenen Prognose sollte immer einheitlich sein, ungeachtet der für diese Prognose notwendigen Preprocessing- oder Modellierungsschritte. Dafür sollte das eigens in LAB-TOUR SH entwickelte Datenmodell "OccupancySeries" genutzt werden. Nur so können externe Programme und Visualisierungen auf dieses Datenmodell ausgerichtet werden und reliabel die Ergebnisse des Prognosemoduls verarbeiten.
- **Einschränkungen kommunizieren:** Im Optimalfall hat ein produktives Prognosemodul probabilistische Fähigkeiten und kann somit einschätzen, wie wahrscheinlich es ist, dass eine Prognose tatsächlich eintritt und wie groß eine wahrscheinliche Abweichung von der Vorhersage sein könnte. Sollte dies nicht möglich sein, macht es Sinn, Einschränkungen an eine/n Nutzer*in mitzuteilen. So könnten zum Beispiel an Feiertagen an bekannten Hotspots zusätzliche Warnhinweise ausgespielt werden, dass die tatsächliche Auslastung die Prognose übersteigen könnte.

5.1.2 Datengrundlagen für die Prognose

Das Prognosemodul in LAB-TOUR SH basiert auf unterschiedlichen Datenquellen. Ein Großteil dieser ist bereits in den Abschnitten „Messung und Sensorik“ und „Datenplattform“ dieses Berichtes ausführlich diskutiert worden, weshalb an dieser Stelle bloß ein kurzer Rückblick stattfinden soll. Zusätzlich sollten alle angebotenen externen Datenquellen, wie die Wetter-, Ferien- und Feiertagsdaten kurz eingeführt werden. Danach soll der Fokus vor allem auf den logischen Implikationen liegen, die aus der Datengrundlage für die Prognose und den Recommender entstehen. Events können die Auslastungen von POI maßgeblich beeinflussen. In LAB-TOUR SH wurde jedoch kein Weg gefunden, Veranstaltungen in den Prognose-PoC mit aufzunehmen. Dies liegt maßgeblich daran, dass für eine standardisierte Verwendung eine ungefähre Schätzung der Besucherzahlen eines Events vorliegen müsste. Hierfür gibt es jedoch keine Datengrundlage in Schleswig-Holstein.

Auslastungsdaten

Kernelement der Auslastungsprognose sind die in den Datenhubs einlaufenden Messdaten. Die dafür in LAB-TOUR SH aufgebaute Sensorik wurde bereits ausgiebig eingeführt, weshalb an dieser Stelle darauf verzichtet wird. Neben dieser Sensorik fließen, gemäß dem Konzept der Integrationsplattform als Sammelbecken für Daten mehrerer Daten-Hubs, weitere Auslastungsdaten ein. Dafür stellten LHIND Frequenzdaten aus St. Peter-Ording und dem Strand und von Parkplätzen in Scharbeutz zur Verfügung. Beide Datenquellen durften mit Genehmigungen der jeweiligen Institutionen im Rahmen dieses Forschungsprojektes genutzt werden. In beiden Destinationen werden an mehreren Parkplätzen und Strandabschnitten in St. Peter-Ording und Scharbeutz Ein- und Ausgänge gemessen. Diese Daten waren für die Entwicklung des Prognosemoduls sehr wichtig, da sie über einen längeren Zeitraum vorliegen und damit ein umfassenderes Training und eine Evaluation der entwickelten Algorithmen ermöglichen. Diese Evaluation kommt allerdings mit der Einschränkung, dass sie mit den Jahren 2021 und in Teilen auch 2022 Zeiträume beinhalten, die unter einem starken Einfluss der Mobilitätseinschränkungen der Corona-Pandemie stehen. Die Pandemie hat zu einem veränderten Reiseverhalten geführt, was auch in Schleswig-Holstein deutliche Auswirkungen zeigte, bzw. im Rahmen eines verstärkten Ausflugsverkehrs an den Küsten (vgl. hierzu Eisenstein et al. 2021). Es ist wichtig zu beachten bei der Arbeit mit der hier im Projekt installierten Sensorik, aber auch mit den genutzten Daten aus St. Peter-Ording und der Lübecker Bucht, dass fast alle der insgesamt gemessenen Daten im Rohdatenformat keine Auslastungen repräsentieren, sondern für jeden einzelnen Sensor Bewegungsflüsse über eine virtuelle Linie beschreiben (also bspw. Ein- und Ausgänge bei einem Strandübergang oder Ein- und Ausfahrten bei einem Parkplatz). Aus der Zusammenfassung mehrerer Sensoren zu einer geschlossenen Fläche, in der möglichst alle Ein- und Ausgänge lückenlos überwacht werden, soll eine Auslastung im eigentlichen Sinne gebildet werden. Dafür wird eine kumulative Summe über alle Ein- und Ausgänge der so überwachten Fläche gebildet. Das hierfür genutzte Verfahren wird im Abschnitt 5.1.3 genauer beschrieben und dann beispielsweise für Parkplätze und Strände angewandt. Im Abschnitt 5.1.4 wird außerdem näher darauf eingegangen, wie die so bestimmten Auslastungen teilweise auch auf umliegende POI verteilt werden. Die Rohdaten liegen größtenteils in 60 min Intervallen vor, in denen die einzelnen Linienüberschreitungen bereits vom Sensorbetreiber kumuliert wurden, sodass nur die Summe aller Überschreitungen in die eine oder andere Richtung vorliegen. Für viele Parkplatzmessungen wurden die Sensoren so platziert, dass sie laut Datenlieferant die Ein- oder Ausfahrt eines Parkplatzes beschreiben. Für solche Sensoren wurde nur eine Richtung von Überschreitungen in die Auslastung des Parkplatzes mit einbezogen. Messwerte, die anzeigen, dass sich ein Fahrzeug beispielsweise über den *Eingang* zum Parkplatz vom Parkplatz entfernt hat, wurden ignoriert, da solche Events vom Datenlieferanten als wahrscheinliche Messfehler eingestuft wurden. Durch unbekannte Störungen sind viele der Rohdaten lückenhaft, sodass immer wieder kürzere oder längere Datenintervalle

fehlen. In St. Peter-Ording ist ein Sensor beispielsweise nur im Sommer aktiv und liefert im Winter keine Daten. Außerdem gibt es divergierende Datenbestände für die Sensoren, was durch die unterschiedlichen Aufstellungszeitpunkte der Sensorik bedingt ist.

Auch schwanken bisweilen die Latenzzeiten in denen die Daten zur Verfügung stehen und liegen bisweilen deutlich über den Zeitraum von 15 Minuten.

Zusätzlich zu diesen Sensoren wurden WLAN-Accesspoints, die von ADDIX im Rahmen des #sh_wlan betrieben werden, an das System angebunden. Für diese Sensoren wurde allerdings keine eigene Prognose gebildet, sondern nur der Live-Wert der Auslastung ausgespielt. Dies hat den Hintergrund, dass die über das #sh_wlan bestimmten Auslastungen noch wenig interpretierbar sind und dafür keine Prognosemodelle aufwändig entwickelt und trainiert werden sollten.

Neben den Auslastungsdaten, die die Zielvariable des Prognosemoduls darstellen, wurden auch sogenannte Feature-Variablen abgerufen, die maßgeblich zur Beschreibung der Umstände einer Auslastungsmessung beitragen. Dies sind vor allem Wetterdaten, aber auch Daten über Ferien- und Feiertage (siehe hierzu Schmücker 2023). Gemeinsam mit dem Zeitstempel einer Messung, über den Informationen zur Uhrzeit und zum Wochentag vorliegen, kann so der Kontext einer Messung recht ausführlich beschrieben werden.

Wetterdaten

Die Wetterdaten, mit denen die Messdaten im Prognosemodul angereichert werden, werden von dem kommerziellen Anbieter OpenWeatherMap bezogen. Im Vergleich zu den Wetterdaten des Deutschen Wetter Dienstes (DWD), die in der Anfangsphase der Entwicklung genutzt wurden, hat OpenWeatherMap einige Vorteile. Der größte davon ist, dass OpenWeatherMap ein Modell entwickelt hat, das unter Berücksichtigung von physikalischen und meteorologischen Gesetzmäßigkeiten eine Interpolation des Wetters zwischen tatsächlichen Messstationen flächendeckend durchführt. Das bedeutet, dass über OpenWeatherMap tatsächlich für jede beliebige Geolocation Wetterdaten angefordert werden können, auch wenn nicht an jeder eine Messstation steht. Dies ist besonders attraktiv, wenn eine semi-automatische Anbindung neuer Besuchermessstationen an das Prognosemodell ermöglicht werden soll. Bei einem Bezug der Daten über den DWD können im Vergleich dazu nur validierte Messdaten angefordert werden, die allerdings auch an einer Station aufgenommen worden sein können, die viele Kilometer entfernt von der Auslastungsmessung liegt. Zudem liefert OpenWeatherMap eine Wettervorhersage, die für eine Woche in der Zukunft recht zuverlässig ist und auch wieder für jede ausgewählte Geolocation zur Verfügung steht. Über OpenWeatherMap werden so automatisiert Daten zur Temperatur, Windgeschwindigkeit, relativen Luftfeuchtigkeit, Dichte der Wolkendecke und zum Niederschlag gesammelt. Diese fünf Datenpunkte beschreiben in diesem PoC die Wetterbedingungen einer jeden Auslastungsmessung.

Ferien und Feiertage

Um über das Wetter und die Wochentage hinaus auch Feiertage und Ferienzeiten mit in die Beschreibung einer Messung einfließen lassen zu können, wurde die Ferien-API unter www.ferien-api.de genutzt. Dort können per REST-API Informationen zu den Ferienzeiten und Feiertagen aller Bundesländer abgefragt werden. In diesem PoC wurden ausschließlich Feiertage aus Schleswig-Holstein eingebunden, bei den Ferienzeiten allerdings die aller Bundesländer. Diese wurden in einem Wert zusammengefasst, der ausführlicher im Abschnitt 4.2.4 erklärt wird.

Globale Sensorik

Weiterhin wurden in dem Prognose-PoC Auslastungsdaten der globalen Sensorik genutzt, um zu testen, ob mit der globalen Sensorik dem Cold-Start-Problem der Algorithmen entgegengewirkt werden kann. Dafür wurden die im Rahmen des Projektes bezogenen GPS-Location Events genutzt, die je Saison typische Tages- und Wochenverläufe an den Messstandorten im Projekt beinhalten. Eine genauere Beschreibung des dafür genutzten Verfahrens folgt in Abschnitt 4.2.5.

5.1.3 Data Cleaning

Ein großer Teil eines jeden KI-Projektes ist es, die zur Verfügung stehenden Daten zu bereinigen, zu bearbeiten und zu verknüpfen, um eine möglichst gute Grundlage für die darauf aufbauenden Modelle und Logiken bieten zu können. Für den Prognose-PoC musste dies ebenfalls umfangreich durchgeführt werden. Dabei haben sich zudem einige Limitationen der Datengrundlage aufgetan, die sowohl in diesem Abschnitt als auch im Abschnitt 5.2.6 thematisiert werden sollen. Der aufwändigste Teil des Data Cleaning war dabei die Transformation der gemessenen Ein- und Ausfahrten in Auslastungsdaten, die Fehler im Messkonzept und in den finalen Auslastungsdaten sichtbar macht. Im ersten Schritt des Data Cleanings wurden Lücken in den Sensordaten mittels Interpolation aufgefüllt.

Interpolation

Um in diesem PoC kleinere Lücken in den historischen Auslastungsdaten aufzufüllen, wurde für jeden einzelnen Sensor ein einfaches Regressionsmodell trainiert, das auf Grundlage der externen Faktoren wie Uhrzeit, Jahreszeit, Temperatur usw. eine Auslastung schätzt. Der Vorteil eines solchen Regressionsmodells ist, dass es im Gegensatz zu Zeitserienmodellen nicht darauf angewiesen ist, dass eine Zeitserie lückenlos vorliegt. Über die so bestimmten Auslastungen wurden dann die vorhandenen Lücken in den Sensordaten aufgefüllt. Es ist zwingend nötig, diese Interpolation auf Ebene der einzelnen Sensoren durchzuführen, bevor eine Zusammenfassung mehrerer Sensoren zu einer Fläche stattfindet, für die dann Auslastungen bestimmt werden. Sollte diese Korrektur nicht durchgeführt werden, könnte die Auslastung der Fläche stark verfälscht werden. Wenn zum Beispiel ein Parkplatz genau eine Ein- und eine Ausfahrt hat und der Sensor an der Ausfahrt für zwei Stunden keine Daten geliefert hat, würde die Auslastung des Parkplatzes wahrscheinlich als zu hoch eingeschätzt werden, da für zwei Stunden alle Einfahrten, aber keine Ausfahrten gezählt würden. Generell ist eine Interpolation von Zeitseriendaten immer dann notwendig, wenn Modelle verwendet werden sollen, die die letzten Werte der Zeitreihe selbst mit in die Vorhersage der nächsten Werte mit einbeziehen sollen. Da solche Modelle, vor allem auf kurzen Zeitskalen, sehr gute Ergebnisse erzielen, ist die Interpolation unabdingbar, um diese Art Modelle auf den historischen Daten trainieren zu können. Im Live-Betrieb findet eine solche Interpolation bei verzögerten Datenlieferungen allerdings nicht statt. Hier übernimmt bei fehlenden Live-Daten ein weiteres Regressionsmodell die Vorhersage. Mehr dazu im Abschnitt 5.1.4.

Auslastung von Parkplätzen

Durch die Interpolation der Sensordaten liegen als Grundlage für die Bestimmung von Auslastungen nun kontinuierliche Flussdaten an jedem einzelnen Sensor vor. Während der Aufstellung der Sensorik wurde außerdem bereits vermerkt, welche Überschreitung der virtuellen Messlinie des Sensors als Betreten und welche als Verlassen der überwachten Fläche gelten. Diese Ströme werden im Folgenden Ins und Outs genannt. Die Ins eines Parkplatzes sind also die summierten Linienüberschreitungen an

allen Einfahrten eines Parkplatzes, die auf den Parkplatz geführt haben. Die Outs sind die summierten Linienüberschreitungen, die an allen Ausfahrten des Parkplatzes vom Parkplatz weggeführt haben. Solche Linienüberschreitungen, die über eine *Ausfahrt* auf den Parkplatz, oder über eine Einfahrt vom Parkplatz weg stattgefunden haben, wurden nicht berücksichtigt, da es sich hier wahrscheinlich um Messfehler handelt (siehe Tab. 17).

	Als Einfahrt gemessen	Als Ausfahrt gemessen
Einfahrt	True Positive (hier berücksichtigt)	False Negative
Ausfahrt	False Positive	True Negative (hier berücksichtigt)

Tabelle 17: Klassifikation von Ein- und Ausfahrten

Die Summe der Linienübertretungen als Proxy für die Auslastung zu nutzen, hat sich als nicht valide herausgestellt, da diese im mathematischen Sinne nur die erste Ableitung der Auslastung nach der Zeit, also deren Veränderungsrate mit der Zeit darstellt. Anders ausgedrückt: Die Ein- und Ausfahrten auf dem Parkplatz geben einen Hinweis auf die Nutzung des Parkplatzes über die Zeit und deren Veränderung, aber nicht unmittelbar über die Auslastung. Hierfür wären weitere Variablen, wie beispielsweise die durchschnittliche Aufenthaltsdauer von Autos auf Parkplätzen nötig, um eine Berechnung durchzuführen. Gerade bei einer oftmals auftretenden glockenförmigen Auslastung im Tagesverlauf ergibt sich als absolute Summe der Linienübertretung eine Kurve, die zwei Höhepunkte hat, wenn sich der POI füllt oder leert und ein lokales Minimum aufweist, wenn die Auslastung besonders hoch ist. Abbildung 53 visualisiert diese Dynamik beispielhaft.

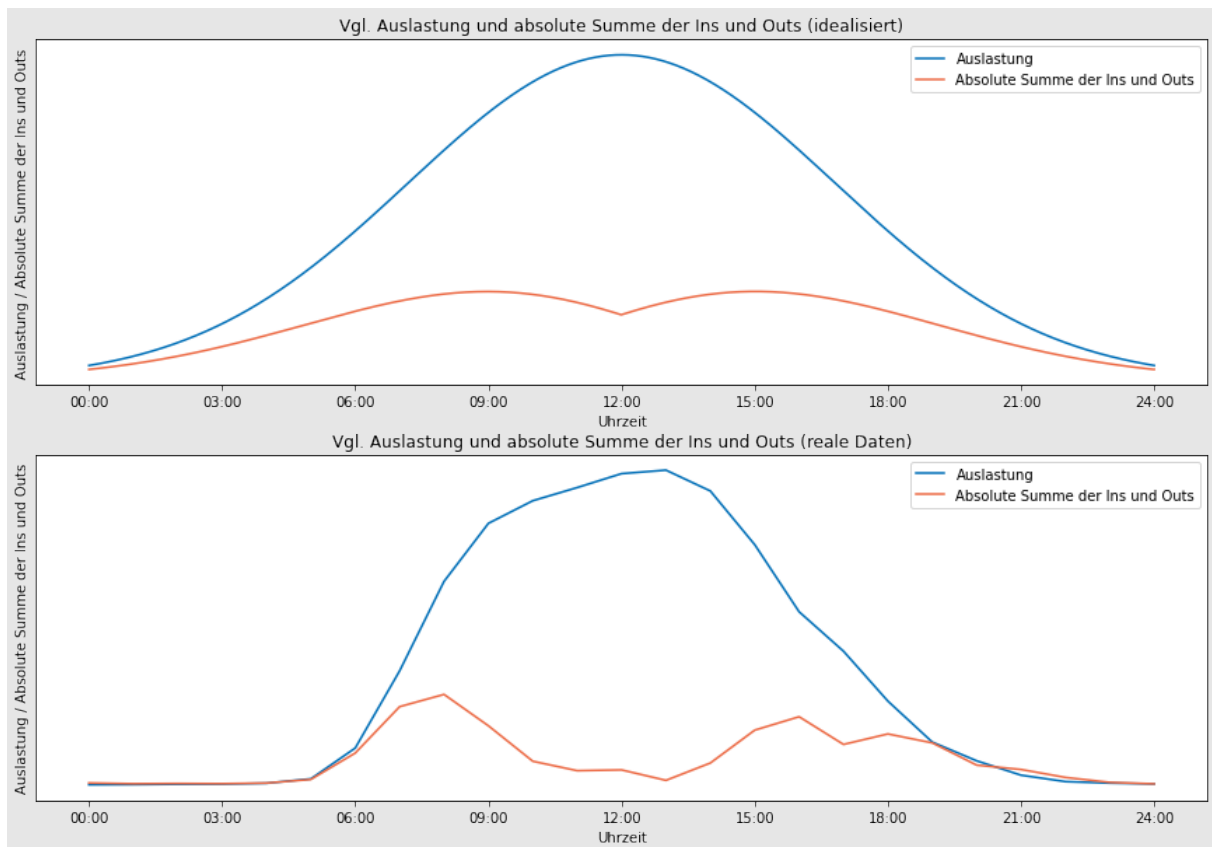


Abbildung 53: Dynamik von Ein- und Ausfahrten als Proxy für die Auslastung eines Parkplatzes

Aus diesem Grund können die Ins und Outs nicht unbearbeitet als Proxy für die Auslastung für das Prognosemodul genutzt werden. Dafür müssen die Ins und Outs zu einer Auslastung transformiert werden. Der Grundgedanke dieser Transformation ist, diese kontinuierlich gegeneinander zu rechnen und damit zu kumulieren. Dieses Verfahren hat sich in der Praxis allerdings ebenfalls als fehleranfällig herausgestellt. Wider Erwarten scheint es bei keinem der überwachten Parkplätze über lange Zeit zu einer gleichverteilten Anzahl an gemessenen Einfahrten und Ausfahrten zu kommen. Dies führt dazu, dass die kumulative Summe über längere Zeit entweder ins Positive oder Negative abdriftet. Eine negative Auslastung ist dabei in der Realität nicht möglich und eine gemessene Auslastung deutlich oberhalb der Parkplatzkapazität ebenfalls nicht. In Abbildung 54 ist der jeweilige Verlauf der kumulativen Summe der Ins und Outs über die Zeit für zwei ausgewählte Parkplätze abgebildet (Stand 04.05.2023).

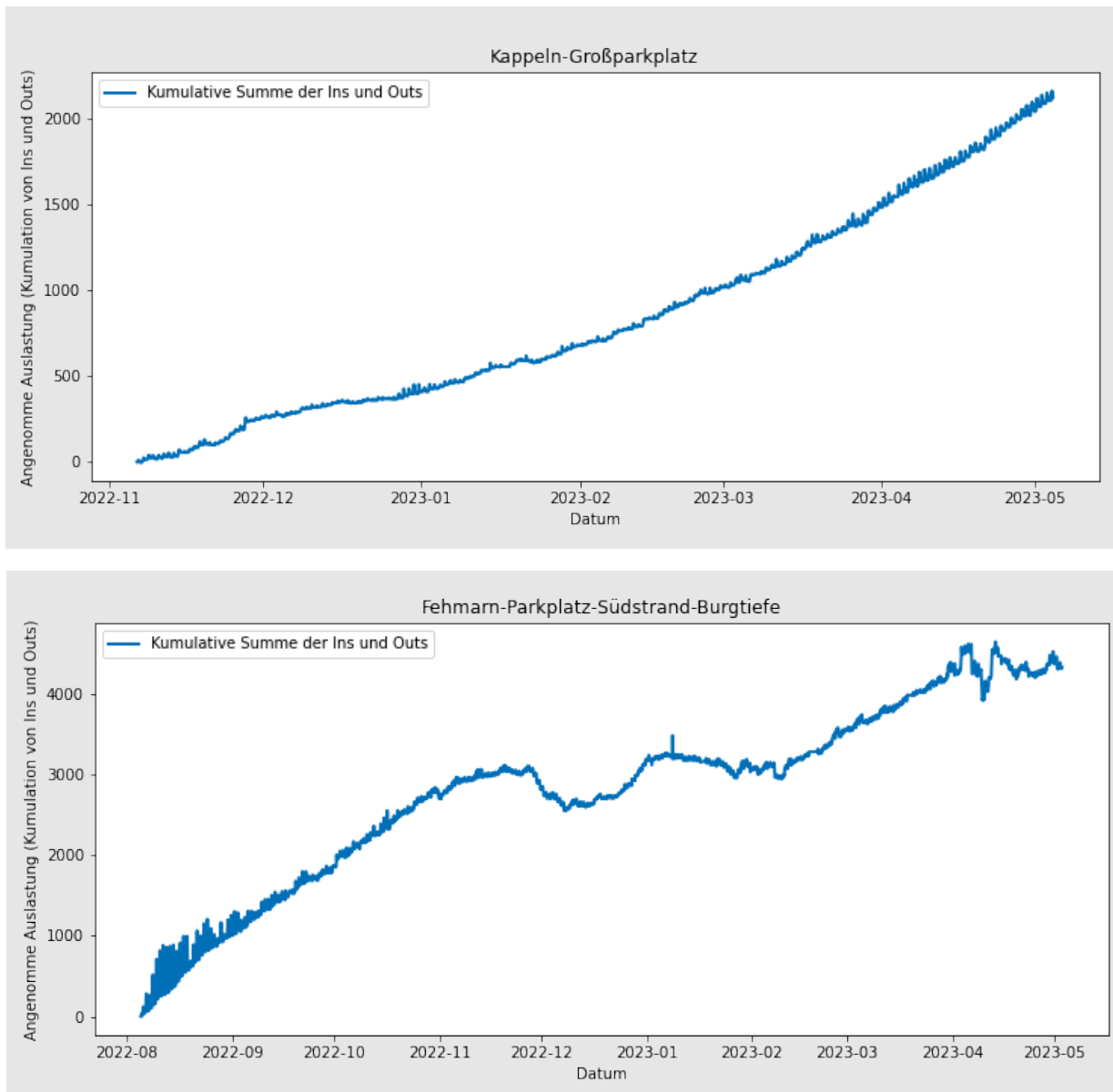


Abbildung 54: Drifts in den lokalen Sensordaten am Beispiel von Kappeln und Fehmarn

Dabei ist in beiden Verläufen klar die tatsächliche Messgröße, die Auslastung der Parkplätze zu sehen. Sie stellen sich als „Zacken“ auf der gezeigten Linie dar. In Kappeln sind diese besonders gegen Ende sichtbar, während sie in Fehmarn besonders am Anfang der Messung erkennbar sind. In beiden Grafiken lassen sich allerdings auch Drifts über den gesamten Messzeitraum beobachten. Die Gründe für diese Drifts sind auf den ersten Blick nicht klar ersichtlich. In Frage kommen dafür wahrscheinlich Ungenauigkeiten der Sensoren (bspw. kann der In-Sensor eine Messgenauigkeit von 95% und der Out-Sensor eine Genauigkeit von 98% haben), potenziell fehlerbehaftete Messaufbauten (bspw. die falsche Anlage der Ein- und Ausfahrten bei der Nutzung von Sensoren, die bidirektional messen), menschliches Verhalten, das in der Realität anders ist als erwartet (bspw. die Nutzung von „Schleichwegen“ zur Auf- und Abfahrt auf einem Parkplatz und nicht die offizielle Ein- oder Ausfahrt) sowie weitere noch unbekannte Fehlerquellen. Da die Drift-Dynamik an jedem Parkplatz ein wenig anders ist, scheint die Kombination dieser Fehlerquellen weiterhin für jeden Parkplatz individuell zu sein und lässt sich dementsprechend nicht über eine allgemeine Korrektur begleichen (vgl. hierzu auch Schmücker, Reif und Stau-

bert, 2023). Um den Messfehler weiter zu untersuchen, wurden an zwei Parkplätzen händische Zählungen über jeweils einen ganzen Tag durchgeführt. Deren Analyse hat allerdings nicht dazu geführt, den Fehler klar identifizieren zu können. Allerdings ist in der Analyse der Daten aufgefallen, dass der tägliche Fehler meistens, aber nicht bei jedem Parkplatz, klein ist im Vergleich zur Messgröße, der Auslastung. Über einzelne Tage hinweg liegt der Messfehler der Auslastung also bei den meisten Parkplätzen mit großer Wahrscheinlichkeit im Toleranzbereich und stellt kein großes Problem dar. Nur durch die Aufsummierung der kleinen Tagesfehler wird der Gesamtfehler irgendwann groß im Vergleich zur Messgröße.

Alle im Projekt LAB-TOUR SH untersuchten Parkplätze bieten die Möglichkeit des Overnight-Parkings, sodass durchaus die Möglichkeit besteht, dass sich auch nachts Fahrzeuge auf den Parkplätzen befinden. Um die Ins und Outs der Parkplätze nutzen zu können und damit eine Näherung der aktuellen Auslastung zu erhalten, wurde dennoch ein regelmäßiger Reset der Auslastung eingeführt. Damit wird verhindert, dass sich die ursprünglich kleinen Tagesfehler aufsummieren können. Über diesen Reset wird die Auslastung eines jeden Parkplatzes um 02:00 Uhr nachts auf 0% gesetzt. Dies folgt der Erwartung, dass die Parkplätze nachts beinahe leer sind. Ab diesem Zeitpunkt werden die Ins und Outs wieder kumuliert, um die Auslastung zu beschreiben, bis in der nächsten Nacht der nächste Reset erfolgt. Es wird erwartet, dass dieses Verfahren eine sehr gute Einschätzung der Auslastung von Parkplätzen bietet, die ländlich gelegen sind, sodass nachts wenige Autos auf den Parkplätzen stehen. Bei Parkplätzen in städtischen Gebieten, die für eine nächtliche Nutzung freigegeben sind, ist der Fehler des Resets deutlich größer. Sollte der Parkplatz bei einer 50% Belegung nachts resetet worden sein, muss der Fehler der Einschätzung der Auslastung im Folgenden ebenfalls bei mindestens 50% angenommen werden. Um diese Problematik abzuschwächen, wurde eine Korrektur eingeführt, die den gesamten Tagesverlauf rückwirkend so erhöht, dass keine negative Auslastung angenommen wird. Wenn beispielsweise nachts ein Reset auf eine 0%-Belegung vorgenommen wurde, aber noch 30 Fahrzeuge auf dem Parkplatz standen, von denen bis 09:00 morgens 15 den Parkplatz verlassen haben, während nur 5 Fahrzeuge auf den Parkplatz gefahren sind, wäre die gemessene Auslastung -10 Fahrzeuge. Da dies nicht möglich ist, wird die gesamte Tagesauslastungskurve um 10 Fahrzeuge nach oben verschoben. So ist in der korrigierten Variante der Tagesauslastungskurve dann um 02:00 eine Auslastung von 10 Fahrzeugen und um 9:00 eine Auslastung von 0%. So wird eine kleine Annäherung an die echten Auslastungsgrade geschaffen. Dieses Verfahren wird für alle Parkplätze angewandt, die in städtischen Gebieten liegen und nachts nicht entweder gesperrt sind oder Parkgebühren verlangen. Abbildung 55 zeigt ein Beispiel dieser Korrektur.

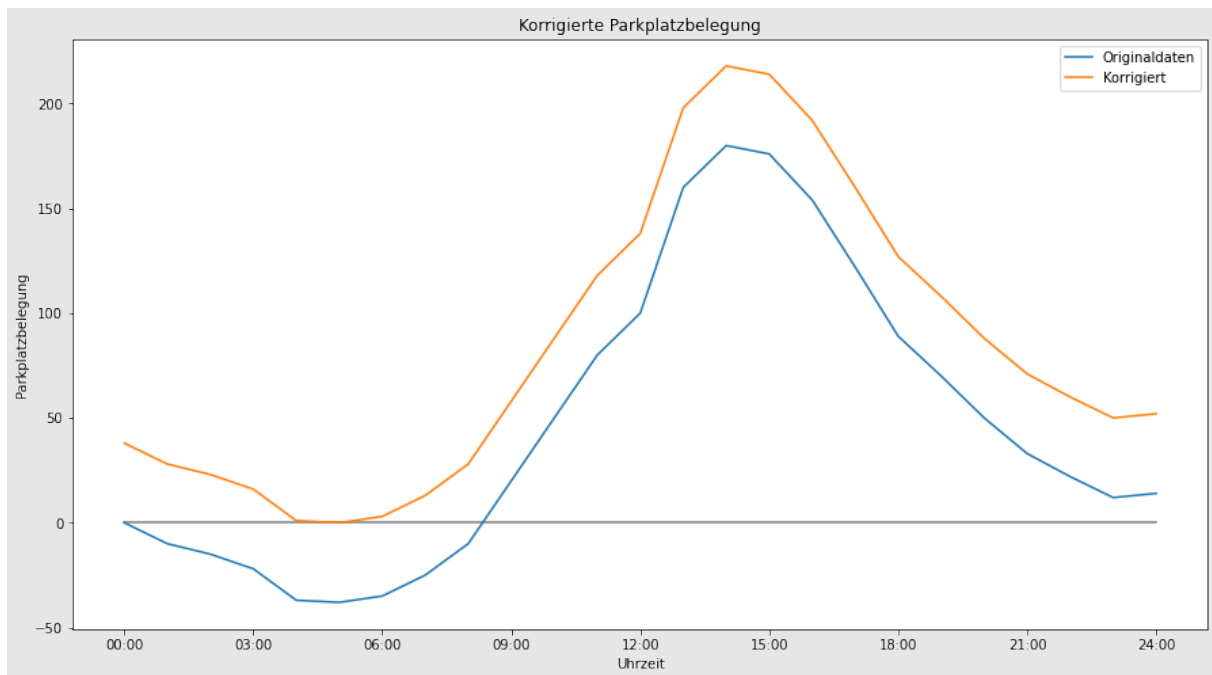


Abbildung 55: Beispielkurve einer korrigierten Parkplatzbelegung

Auf der anderen Seite wurde für alle Flächen, bei denen davon ausgegangen wird, dass die nächtliche Auslastung nahe null liegt, eine andere Korrektur eingeführt. Diese Korrektur bezieht sich allerdings ausschließlich auf historische Daten. Die Grundlage dieser Korrektur ist, dass eine über den Tag entstandene Auslastung, die sich bis zum Reset um 02:00 Uhr hält, als Fehler im Messaufbau interpretiert wird. Ein Beispiel, das dies gut verdeutlicht, ist der Messaufbau in Scharbeutz. Dort werden an mehreren Zugängen zum Strand die Ins und Outs gemessen. Allerdings ist es möglich, am Strand eine etwas weitere Strecke zurückzulegen und dann die Promenade an einem unüberwachten Strandabschnitt wieder zu betreten – oder andersherum. Personen wurden hier also beim Betreten oder Verlassen der Fläche registriert, nicht aber beim Ausgang. Wenn die Auslastung des Strandes nach der Kumulation der gemessenen Ins und Outs über einen ganzen Tag um 02:00 Uhr nachts hier noch 1.000 Personen beträgt, kann davon ausgegangen werden, dass diese beim Betreten des Strandes erfasst wurden, beim Verlassen allerdings nicht. Dementsprechend muss das Verlassen des Strandes dieser 1.000 Personen im Nachhinein in die Daten integriert werden. Dafür wurde der von den Sensoren gemessene “Out-Traffic” über den jeweiligen Tag analysiert. Dieser ergibt eine Einschätzung darüber, zu welcher Zeit wie viele Menschen den Strand verlassen haben. Anhand dieser Verteilung wurden dann die 1.000 nicht registrierten Ausgänge zusätzlich über den Tag verteilt. Dieser Ansatz geht davon aus, dass die Fehler gleichmäßig über den Tag verteilt sind und immer ein konstanter Prozentsatz aller Menschen, die den Strand verlassen, nicht registriert werden. Diese Korrektur kann zu einer erheblichen Veränderung des Tagesverlaufes der Auslastung, vor allem in den Abendstunden führen, wie Abbildung 56 zeigt.

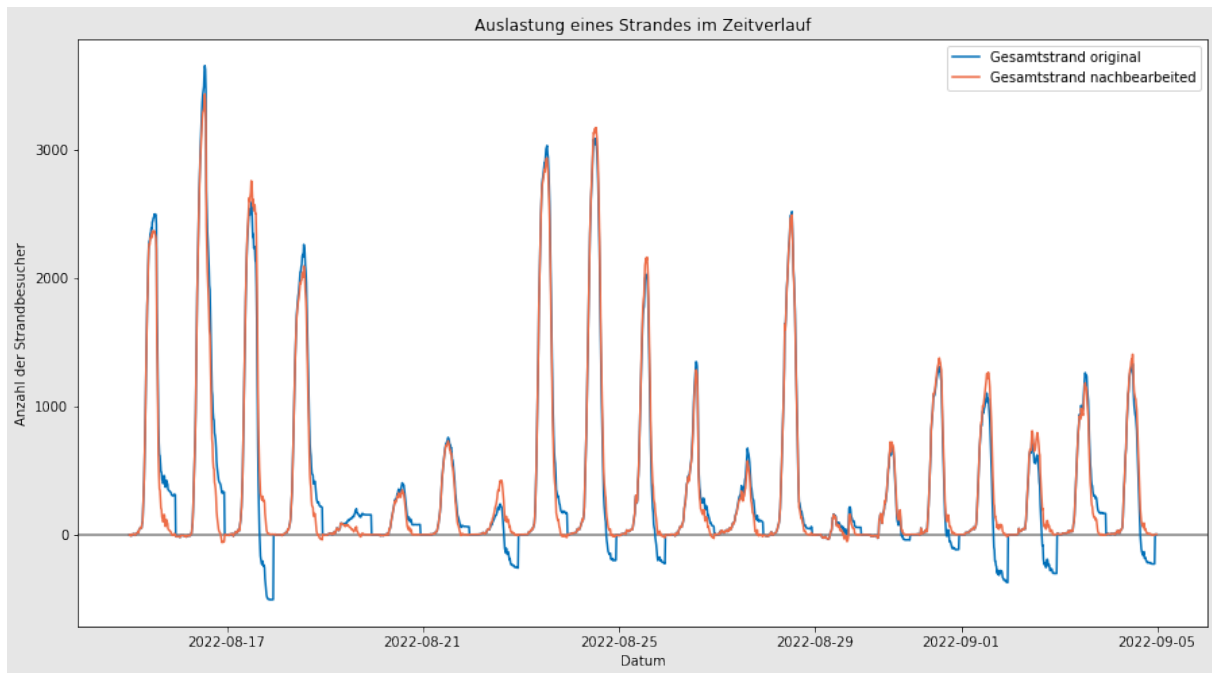


Abbildung 56: Korrektur von Auslastungsdaten im Zeitverlauf: Original und Bearbeitung

Ein erheblicher Nachteil dieser Korrektur ist, dass sie nur rückwirkend von einem Reset zum nächsten angewandt werden kann. Dementsprechend können auch nur die Trainingsdaten auf diese Weise korrigiert werden, während die Live-Daten so nicht korrigiert wurden. Trotz der beobachteten Diskrepanz, die gegen Abend im Vergleich zur Messgröße signifikant wird, wird diese nicht als kritisch angesehen. Im spezifischen Kontext vom Strand wird nicht erwartet, dass die Auslastung in den Abendstunden die Kapazitätsgrenzen übersteigt.

In einem letzten Schritt wurden die Kapazitäten aller überwachten Flächen bestimmt. Bei Parkplätzen ist diese entweder gezählt worden, wenn klare Park-Markierungen vorlagen, oder geschätzt worden, wenn dem nicht so war. Die Kapazitäten der Strände wurden geschätzt. Anhand dieser Kapazitäten konnten die bestimmten absoluten Auslastungen in prozentuale Auslastungen umgerechnet werden. Diese haben als untere Grenze eine Auslastung von 0%, jedoch im Allgemeinen keine Grenze nach oben. Auch eine 120% Auslastung wäre, zumindest für Strände, denkbar. In einem solchen Fall läge klar eine Überfüllung des Strandes vor. Bei Parkplätzen mit klar markierten Stellplätzen sollte die Obergrenze der relativen Auslastung allerdings sehr nah an 100% liegen, da kaum weitere Autos auf den Parkplatz passen sollten. Mit Hilfe dieser Korrekturen können die Sensordaten annehmbar gut zu Auslastungen transformiert werden. Allerdings schwankt dieses Verfahren für unterschiedlich überwachte Flächen in der Qualität. Für manche Parkplätze wurde mit diesen Korrekturen eine regelmäßige Auslastung von 200% beobachtet. Weitere Vorschläge zum Umgang mit solchen Messungen folgen in Abschnitt 5.2.4.

#sh_wlan-Accesspoints

Da auch die WLAN-Accesspoints des #sh_wlan in diesem PoC als Auslastungssensoren aufgenommen wurden, wurde ein Verfahren getestet, dass die momentane Auslastung eines Accesspoints in eine relative Auslastung übersetzt. Da ein solcher Accesspoint misst, wie viele Geräte über ihn mit dem #sh_wlan verbunden sind, ist es nicht möglich, von außen eine Kapazität festzulegen, wie bei Stränden

oder Parkplätzen. Eine solche Kapazität ist zu individuell und abhängig vom Standort des Accesspoints und dem Prozentanteil der umliegenden Geräte, die im #sh_wlan angemeldet sind. Um trotzdem eine Auslastung bestimmen zu können, wurde die Durchschnitts-Auslastung der 99. Perzentile aller an einem Standort gemessenen, absoluten Auslastungen als 100% Auslastung festgelegt. So lassen sich relative Auslastungen anhand dieses Wertes bestimmen. Dieses Verfahren hat allerdings den Nachteil, dass nicht klar ist, wie sehr die so angenommene 100%-Auslastung auch tatsächlich die Kapazitäten der umgebenden POI repräsentiert. Damit die Anwendung dieses Verfahrens valide ist, muss es im Messzeitraum mehrfach eine tatsächliche Überbelastung des umliegenden Gebiets gegeben haben, damit sich die theoretische Kapazität des Accesspoints mit der realen Kapazität des umliegenden Gebietes deckt.

5.1.4 Logik und Modelle

In den vorherigen zwei Abschnitten wurde bereits eingeführt, welche Rohdaten in diesem PoC genutzt wurden und wie die Auslastungsmessungen im Data Cleaning behandelt wurden, um sie in relative Auslastung zu transformieren. Dieser Abschnitt soll nun einen Überblick darüber geben, mit welchen Algorithmen und Logiken diese Daten genutzt wurden, um Prognosen über die zukünftige Auslastung zu machen. Dafür stehen den Algorithmen konkret in diesem PoC die folgenden Features einer jeden Messung zur Verfügung (siehe Tab. 18):

Feature	Beispiel
Zeitstempel	12:15:00 06.06.2023
Auslastung	50%
Temperatur	21.2°C
Gefühlte Temperatur	20.4°C
Windgeschwindigkeit	12m/s
Niederschlag	0mm
Niederschlagsform	-
Luftfeuchtigkeit	50%
Ferien in SH	Ja
Gesamtferienindex	0.76
Feiertag in SH oder HH	Nein

Tabelle 18: Features des PoC

Die nun skizzierte Logik hat dabei den Anspruch, den meisten der im Vorfeld genannten Anforderungen zu entsprechen.

Bildung eines Ensembles

Der Grundgedanke des Prognose-PoCs ist die Entwicklung eines Ensemble-Modells, in dem mehrere Algorithmen gemeinsam eine Auslastungsprognose abgeben. Diese Art von Modellen ist sehr gängig in der wissenschaftlichen Praxis und bei der Prognose von Zeitserien im Allgemeinen (Allende und Valle, 2017). Ein Anwendungsgebiet, in dem diese Art von Modellen oft genutzt wird, ist die Klimaphysik. Hier werden beispielsweise Ensemble-Modelle genutzt, um die Erhöhung der globalen Durchschnittstemperatur über die nächsten Jahrzehnte einschätzen zu können. Solche Ensembles haben oftmals die

positive Eigenschaft, dass die tatsächliche Dynamik der beobachteten Größe gut abgebildet wird, während viele Artefakte, die nur bei der Nutzung einzelner Modelle auftreten, durch die Mittelung der Vorhersagen der Ensemblemitglieder herausgefiltert werden.

In diesem PoC wurde das Ensemble um eine weitere Funktion erweitert, die in dem Bereich des automatisierten maschinellen Lernens (Auto-ML) bekannt ist. Dieses Fachgebiet beschäftigt sich mit der annähernd vollautomatisierten Entwicklung von ML-Modellen für vordefinierte Problemstellungen. Dabei wird eine Vielzahl an Algorithmen auf das vorgegebene Problem angewandt und evaluiert. Dann wird der Algorithmus verwendet, der in der ebenfalls automatisierten Evaluation die besten Ergebnisse produziert hat. In LAB-TOUR SH führt dieser Ansatz zu einer Aufteilung des Ensembles in einen aktiven und einen passiven Teil. Für jeden Standort wurde entsprechend des Verfahrens im Auto-ML ausgewertet, wie gut die Algorithmen funktionieren. Dafür wurde ein Pool von sechs verschiedenen Algorithmen zur Verfügung gestellt, die das Gesamtensemble bilden. Aus diesen wurden die drei leistungsfähigsten Algorithmen in das aktive Ensemble übernommen, die anderen drei in den passiven Teil des Ensembles. Die Aufteilung darüber, welche Algorithmen Teil des aktiven bzw. passiven Ensembles sind, ist also individuell für jeden Messstandort. Für die Prognosen werden nur die Modelle genutzt, die im aktiven Teil des Ensembles liegen. Über diese Funktionalität soll gewährleistet werden, dass die Algorithmen des Prognosemoduls eine möglichst große Anpassungsfähigkeit an verschiedene Use Cases haben. Die Aufteilung in das aktive und passive Ensemble wird im Abstand von einigen Monaten unter Berücksichtigung der in der Zwischenzeit neu gewonnenen Auslastungsdaten wiederholt automatisiert durchgeführt. Damit sollte gewährleistet sein, dass immer die Modelle ins aktive Ensemble aufgenommen werden, die mit den kompletten zur Verfügung stehenden Trainingsdaten und dem umfangreichen Evaluationszeitraum am besten performen.

Eine in der Entwicklung vorgegebene Einschränkung der automatischen Selektion von Modellen wurde dennoch gemacht. Diese Einschränkung baut auf folgendem Hintergrund auf. Der Pool der zur Verfügung stehenden Modelle umfasst in diesem PoC zwei Arten von Modellen. Dies sind zum einen Modelle mit einer autoregressiven Komponente und zum anderen reine Regressionsmodelle. Der Unterschied dieser beiden Modellklassen liegt dabei darin, dass die autoregressiven Modelle neben externen Daten den Auslastungszustand an kurzfristig zurückliegenden Zeitpunkten in einer Prognose berücksichtigen, während die reinen Regressionsmodelle ausschließlich externe Faktoren wie das Wetter, die Uhrzeit etc. zur Vorhersage nutzen. Rein autoregressive Modelle, die komplett unabhängig von externen Faktoren nur unter Berücksichtigung der vergangenen Entwicklung der Auslastung eine Prognose abgeben, haben in der Explorationsphase des PoC keine guten Ergebnisse geliefert und wurden deshalb nicht in das Ensemble aufgenommen. Die Vermutung liegt nahe, dass vor allem die Dynamik an den Stränden Schleswig-Holsteins, auf der dieser PoC maßgeblich entwickelt wurde, so abhängig vom Wetter ist, dass diese Information einem Modell zwingend vorliegen muss, damit es gute Vorhersagen machen kann. Die Prognosemodelle, die eine autoregressive Komponente besitzen, zeichnen sich dadurch aus, dass sie beispielsweise für eine Prognose zur Auslastung um 10:00 Uhr (bei einer stündlichen Auflösung) die Auslastung um 9:00 Uhr kennen müssen. Generell produziert diese Art von Modellen sehr gute Ergebnisse, da sehr viel Information zur Auslastung um 10:00 Uhr bereits durch die Nennung der Auslastung um 9:00 Uhr gegeben ist. Allerdings ist es schwieriger, mit dieser Art von Modellen Prognosen zu machen, die weit in die Zukunft gehen. Dafür müssen die Vorhersagen iterativ vorgenommen werden, sodass der vorhergesagte Wert für einen Zeitpunkt in die Vorhersage des nächsten Zeitpunkts mit aufgenommen wird und so weiter. So pflanzt sich der Fehler der Prognose allerdings immer weiter fort, sodass die Prognose schlechter wird, je öfter dieses iterative Verfahren angewendet wird. In mehreren Experimenten mit diesen Modellen wurde außerdem beobachtet, dass sie der vergangenen Auslastung so viel Wert beimessen, dass die Modelle Schwierigkeiten haben, den Start eines Tages vernünftig zu beschreiben. Da dort die vorangegangenen Werte sehr niedrig sind,

bleiben auch die prognostizierten Werte sehr niedrig. Um diese beiden Probleme zu adressieren, wurden die Modelle mit autoregressiver Komponente in ihrem Vorhersagehorizont auf sechs Stunden begrenzt. Innerhalb dieses Zeitraumes produzieren sie signifikant bessere Ergebnisse als die reinen Regressionsmodelle. In der Entwicklung wurde nun vorgegeben, dass in das aktive Ensemble genau ein Modell mit autoregressiver Komponente eingeht. Diese Modelle wurden von uns als Short-Term Modelle titulierte, während die reinen Regressionsmodelle als Long-Term Modelle bezeichnet werden (siehe Tabelle 19). Im Algorithmus Pool stehen dem Ensemble zwei Short-Term Modelle und vier Long-Term Modelle zur Verfügung, deren Performance jeweils untereinander automatisiert verglichen wird. Der Fehler der Regressionsmodelle steigt mit fortschreitendem Vorhersagehorizont nicht aufgrund der Algorithmik an sich, sondern nur weil der Fehler der Wettervorhersage größer wird und sich somit auf die Modelle auswirkt. Von den Regressionsmodellen werden im PoC zwei Modelle ins aktive Ensemble aufgenommen.

Durch diese Kombination der beiden Modelltypen ist außerdem eine Ausfallsicherheit gegenüber verzögerten Datenlieferungen gegeben. Sollten die Lieferung von Live-Daten zu einem Standort sich aus irgendeinem Grund verzögern, melden die Modelle mit autoregressiver Komponente aufgrund der fehlenden Daten einen Fehler und werden nicht berücksichtigt. Dann werden nur noch die beiden Regressionsmodelle genutzt, um die Vorhersagen zu machen.

	Modelle mit autoregressiver Komponente (Short-Term Modelle)	Reine Regressionsmodelle (Long-Term Modelle)
Erläuterung	Externe Daten zzgl. Auslastungsdaten	Ausschließlich externe Daten wie Wetter etc.
Anzahl	2	4
Vorteile	<ul style="list-style-type: none"> ● Berücksichtigen vergangene Auslastungszustände für genaue Prognosen ● Gute Ergebnisse für kurzfristige Vorhersagen ● Hohe Informationsdichte durch Kenntnis vorheriger Auslastungszustände 	<ul style="list-style-type: none"> ● Nutzen externe Faktoren wie Wetter und Uhrzeit für Vorhersagen ● Bessere Performance bei langfristigen Vorhersagen ● Nicht abhängig von Live-Daten
Nachteile	<ul style="list-style-type: none"> ● Schwierigkeiten bei langfristigen Vorhersagen ● Fortschreitender Fehler bei iterativen Prognosen ● Schwierigkeiten bei Start des Tages aufgrund niedriger vorangegangener Werte ● Abhängig von Live-Daten 	<ul style="list-style-type: none"> ● Steigender Fehler bei zunehmendem Vorhersagehorizont aufgrund von Wettervorhersage-Fehlern

Tabelle 19: Vergleich von Regressionsmodellen

Etwaiges Feature Engineering, also die gezielte Manipulation der Trainingsdaten mit dem Ziel, die Prognose zu verbessern, wurde in diesem PoC an die individuellen Modelle im Ensemble-Pool geknüpft. Dementsprechend konnten mehrere Feature-Engineering Pipelines für die Modelle entwickelt werden, sodass die individuellen Modelle für verschiedene Use Cases möglichst gut performt haben. Eine genaue Auflistung der im Feature Engineering vorgenommenen Schritte entzieht sich dem Scope dieses Berichtes. Einige erfolgreiche Kernelemente des Feature Engineering waren jedoch eine Glättung des Tageswetters durch Mittelung, das Erzeugen neuer synthetischer Datenpunkte und die Normalisierung der Daten im Jahresverlauf. Auf die beiden letztgenannten Feature-Engineering Schritte und deren Intention soll in den nächsten beiden Abschnitten gesondert eingegangen werden.

Fokus auf Momente hoher Auslastungen und Mitigierung des Cold-Start-Problems

Das Prognosemodul muss in der Lage sein, Momente mit außergewöhnlich hoher Auslastung besonders sicher vorherzusagen, obwohl diese nur einen Bruchteil der Trainingsdaten ausmachen. Um diese Diskrepanz zu adressieren, wurden im Feature-Engineering der meisten Modelle Schritte unternommen, die den Fokus der Modelle auf diese Momente lenken. Dafür wurden unter anderem synthetische Datenpunkte erzeugt, die hohe Auslastungen abbilden. Dafür wurde der SMOTE-Algorithmus (Chawla et al., 2017) angewendet, der in diesem Fall auf der Grundlage aller Datenpunkte angewandt wurde, die eine Auslastung von über 70% aufweisen. Aus diesen wurden per zufällig gewichteter Linearkombination mehrerer Datenpunkte neue Datenpunkte erzeugt, die damit eine synthetische Auslastung und synthetische Wetterverhältnisse vorweisen. So wurde die Anzahl an Datenpunkten mit einer Auslastung von über 70% in den Trainingsdaten so weit erhöht, dass sie 25% der Gesamtdaten beträgt. Dadurch konnte die Vorhersagequalität im Bereich hoher Auslastungen signifikant verbessert werden. Ein weiterer Ansatz, um die Prognosen im Bereich hoher Auslastungen zu verbessern, war die stärkere Gewichtung solcher Datenpunkte mit einer Auslastung von höher als 70%, wenn die entsprechenden Algorithmen eine solche Gewichtung ermöglichen.

Um das oben erwähnte Cold-Start Problem zu mindern, wurden im Pool des Ensembles zwei der vier reinen Regressionsmodelle (Long-Term Modelle) mit jeweils einem speziellen Feature ausgestattet. Eines dieser Modelle versucht, mit Methoden des Transfer-Lernens dem Cold-Start Problem entgegenzuwirken. Dafür wurde eine Gruppierung von Messstationen vorgenommen, bei denen ein ähnliches Verhalten erwartet wird. Dafür wurden beispielsweise die Parkplätze in Scharbeutz zu einer Gruppe zusammengefasst. Das Modell wurde dann auf den über die jeweilige Parkplatzkapazität zu einer prozentualen Auslastung skalierten Auslastungsdaten aller Parkplätze der Gruppe trainiert. So kann die ungefähre Dynamik der Parkplätze durch dieses Modell schnell erfasst werden. Sobald genügend Trainingsdaten dieses speziellen Parkplatzes vorliegen, wird erwartet, dass die Modelle, die ausschließlich auf diesen Daten trainieren, besser performen als das Modell, das auf den Daten der Gruppe trainiert. Sobald dies passiert, wird das Gruppen-Modell aus dem aktiven Ensemble verdrängt und nicht mehr in die Prognose eingebunden. Dieses Verfahren birgt allerdings die Gefahr, dass bei einer falschen Definition der Gruppen eine unpassende Dynamik durch das Gruppen-Modell erlernt wird. Wenn ein neu angeschlossener Parkplatz beispielsweise wenig besucht wird und in einer Gruppe mit Parkplätzen verordnet wird, die regelmäßig ausgelastet sind, würde für den neu angeschlossenen Parkplatz oft eine zu hohe Prognose gemacht werden. Allerdings hätte dies auch zur Folge, dass das Gruppen-Modell schlechter performt und so schneller von anderen Modellen des Ensembles verdrängt wird.

Das zweite speziell ausgestattete Long-Term Modell wurde mit einem aus der globalen Sensorik interpolierten Jahresverlauf der Auslastung in stündlicher Auslastung angereichert. Für neu angeschlossene Sensorstandorte, an denen globale Daten beschafft wurden, ist dieses Modell darauf ausgerichtet, die Differenz dieses Jahresverlaufs zu den gemessenen Daten zu bestimmen. Um dieses Verfahren sinnvoll

anwenden zu können, sollte allerdings schon ein kurzer Zeitraum an gemessenen Auslastungsdaten vorliegen, um den Skalierungsfaktor zwischen der globalen und lokalen Sensorik zu bestimmen. Abbildung 57 zeigt den Vergleich zwischen dem aus der globalen Sensorik interpolierten Jahresverlauf und den tatsächlich gemessenen Daten.

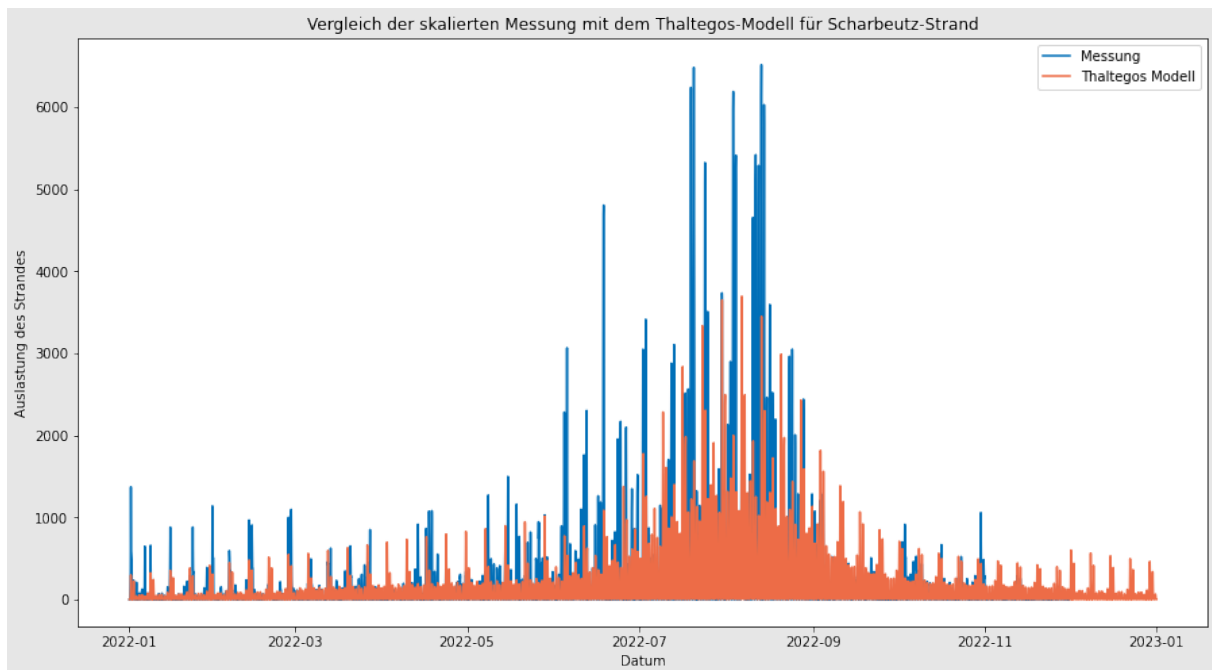


Abbildung 57: Vergleich zwischen Realdaten und Modellierung am Beispiel des Scharbeutzer Strandes

Der Kerngedanke dieses Modells ist es, den natürlichen Jahres-, Wochen- und Tagesgang über die globalen Daten aus dem Training des Modells herauszuhalten und nur noch den Wettereffekt zu lernen. Da die globalen Daten über jeweils drei Monate gemittelt sind, dürfte der vorliegende Wettereffekt vernachlässigbar sein, während dieser in den lokalen Daten zwingend auftaucht. Durch die Bildung der Differenz kann der Wettereffekt vermutlich besser isoliert und durch das Modell erlernt werden. In der Vorhersage addiert dieses Modell dann den prognostizierten Wert auf den Wert des entsprechenden Zeitpunkts aus dem Jahresgang der globalen Sensorik, um einen finalen Auslastungswert zu prognostizieren. Wieder ist der Gedanke, dass so bald genug Sensordaten vorhanden sind, andere Modelle dieses Modell aus dem Ensemble vertreiben, da sie langfristig wahrscheinlich besser performen.

Vor allem das Training über Gruppen ähnlicher POI hat dazu geführt, dass in simulierten Experimenten maßgeblich schneller eine zufriedenstellende Auslastungsprognose abgegeben werden konnte.

Prognose

Die von den individuellen Modellen im aktiven Ensemble abgegebenen Prognosen werden für die finale Vorhersage gemittelt. Dabei wird das Short-Term in den ersten 6 Stunden deutlich stärker gewichtet, da in diesem Zeitraum erwartet wird, dass das Short-Term Modell deutlich bessere Vorhersagen für diesen Zeitraum macht. Alle weiteren Zeitpunkte werden ausschließlich von den Regressionsmodellen vorhergesagt. Insgesamt werden vom Prognosemodul jede Stunde auf Grundlage der aktuellen Wetterdaten und -vorhersagen stündliche Vorhersagen für die Auslastung der nächsten 96 Zeitpunkte

gemacht. Um eine konstante Rückgabe an das Frontend zu gewährleisten, wurden im PoC diese 96 vorhergesagten Zeitpunkte so mit tatsächlichen Messungen des heutigen Tages verschnitten, dass eine aus Beobachtung und Prognose bestehende Zeitserie entsteht, die den kompletten heutigen Tag und die nächsten drei Tage beschreibt. An den Standorten der WLAN-Accesspoints wurde aufgrund der schwierigen Interpretierbarkeit der Daten keine Prognose durchgeführt. Diese werden im Recommender ausschließlich für die Live-Auslastungsmessung genutzt.

Propagieren von Auslastungen

Von den in LAB-TOUR SH installierten und weiteren Messtationen werden die zugehörigen Prognosen teilweise an umliegende POI weitergeleitet. Auch diese Funktionalität ist in den Anforderungen bereits ausführlich eingeleitet worden. In diesem PoC konnte aufgrund begrenzter Ressourcen und der Natur des Problems allerdings keine elaborierte, zufriedenstellende Logik entwickelt werden, nach der diese Propagation funktioniert. Um sich der Problemstellung trotzdem zu nähern, wurde das im Folgenden beschriebene Verfahren angewandt. In einem ersten Schritt wurde die vier Tage umfassende Zeitserie, die das Prognosemodul für jeden Messstandort erstellt, auf vier Werte reduziert. Diese Werte sind die maximale Auslastung eines jeden dieser vier Tage. In einem zweiten Schritt wurden diese vier Werte an alle POI im Umkreis von zwei Kilometern propagiert. Sollte ein POI eine Auslastungsprognose von mehreren Messstationen empfangen, wird der Mittelwert daraus gebildet. Es wurde nur der tagesmaximale Auslastungswert propagiert, um im Frontend nicht durch die Anzeige einer stündlich aufgelösten Zeitserie den Anschein zu erwecken, dass diese propagierte Prognose auch für den neuen POI sehr genau ist. Durch die Propagation des maximalen Wertes wird aus den vorliegenden Informationen das Worst-Case Szenario angenommen, sodass die Lenkungswirkung im Recommender weg von POI laufen kann, die irgendwann im Laufe des Tages eine hohe Auslastung erfahren. Die Propagation der Auslastung der WLAN-Accesspoints an umliegende POI ist noch schwieriger, da diese Auslastungen noch individueller und unspezifischer sind. Trotzdem wurde auch die gemessene Live-Auslastung dieser an alle POI in einem 2 km Umkreis propagiert. Das im PoC angewendete Verfahren zur Propagierung von Auslastungen birgt einige Schwächen, auf die im Abschnitt 5.1.6 noch einmal eingegangen wird.

5.1.5 Technische Details

Der folgende Abschnitt versucht, einen Überblick über die vorgeschlagene Architektur des Prognosemoduls zu vermitteln und dabei auf die technischen Besonderheiten dieses Ansatzes einzugehen. Da der Hauptfokus dieses Berichts aber auf der Exploration der Rohdaten und den logischen Ansätzen der eingesetzten Modelle liegt, wird der Detailgrad dieses Überblicks nicht einer technischen Dokumentation entsprechen.

Während der technischen Umsetzung der im Vorhinein beschriebenen Logiken stand die Skalierbarkeit und konstante Verfügbarkeit der Prognosen stets im Mittelpunkt der Bemühungen. Aus diesem Grund hat das prototypische Deployment des Prognosemoduls in Kubernetes stattgefunden.

Dafür wurde die entwickelte Software in Docker-Containern verpackt, deren Kommunikation untereinander über REST-Schnittstellen stattfindet. Diese Schnittstellen wurden mit FastAPI entwickelt.

Bei der Betrachtung der Gesamtarchitektur des Prognosemoduls (vgl. Abb. 58) lässt sich schnell eine Aufteilung in zwei Hauptbereiche feststellen. Einer dieser Bereiche umfasst das Training und Tuning der Modelle, der andere Bereich die Live-Vorhersage von Auslastungen. Diese beiden Komponenten wurden absichtlich voneinander entkoppelt, um der Anforderung zu entsprechen, dass die Prognose

immer in nahezu Echtzeit geliefert werden muss. Durch die Abkopplung von den rechenintensiven Aufgaben des Trainings und Tunings, steht immer genügend Kapazität zur Erstellung der Vorhersagen zur Verfügung. Außerdem können so beide Prozesse parallel laufen und Live-Vorhersagen gemacht werden, während neue Modelle trainiert werden. In dieser Zeit werden für die Vorhersagen weiterhin die „alten“ Modelle genutzt. Sobald das Training der „neuen“ Modelle abgeschlossen ist, können diese in der Live-Vorhersage verwendet werden.

Trotz dieser bewussten Trennung gibt es zwei zentrale Elemente, die beide Bereiche miteinander verbinden. Zum einen ist dies eine zentrale MongoDB. Hier werden Meta-Informationen zu den Ensemble- und Vorhersagemodellen gespeichert. Beispielsweise ist dort hinterlegt, welche Algorithmen aktuell im Ensemble für eine Messstation verwendet werden und wo diese gespeichert sind. Dies wird per Pfad angegeben, der das zweite zentrale Element referenziert, das die Komponenten „Training und Tuning“ und „Live-Vorhersage“ miteinander verbindet. Dabei handelt es sich um eine gemeinsame virtuelle Festplatte, die in alle Vorhersage-Pods eingebunden wird. Hier werden die Prognose-Algorithmen gespeichert und auch wieder geladen, wenn diese benötigt werden. Durch die Verwendung eines zentralen Speichers können die Pods universell verwendet werden. So können die Vorhersage-Pods jedes Modell laden und für jede verfügbare Messstation eine Vorhersage machen oder die zugrunde liegenden Modell tunen und trainieren. Gleichzeitig ermöglicht es, dass die Pods stateless konzipiert werden können, sodass die Anzahl der Vorhersage Pods uneingeschränkt skaliert werden kann.

Nach der Erörterung der zugrundeliegenden Infrastruktur folgt als nächster Schritt die Beschreibung der darauf aufsetzenden Software. Bei der Betrachtung des Pfades innerhalb der Gesamtarchitektur, der bei einer Anfrage zur Auslösung einer Vorhersage, des (Re-)Trainings oder (Re-)Tunings genommen wird, stellt der NGINX Loadbalancer den Startpunkt dar. Dieser leitet die Anfragen in die entsprechenden Silos für das Training und Tuning und der Vorhersage. In diesen Silos trifft die Anfrage auf einen Kubernetes-internen Loadbalancer, welcher nun die Last auf die einzelnen Pods verteilt, in denen das Ensemble-Modell läuft. Von hier werden dann die einzelnen Prognosen der individuellen Algorithmen in den Prediction-Pods getriggert und anschließend in den Ensemble-Pods wieder zusammengeführt. Diese beiden Ebenen der Erstellung der Prognose (vgl. Abb. 58) soll im Weiteren beschrieben werden.

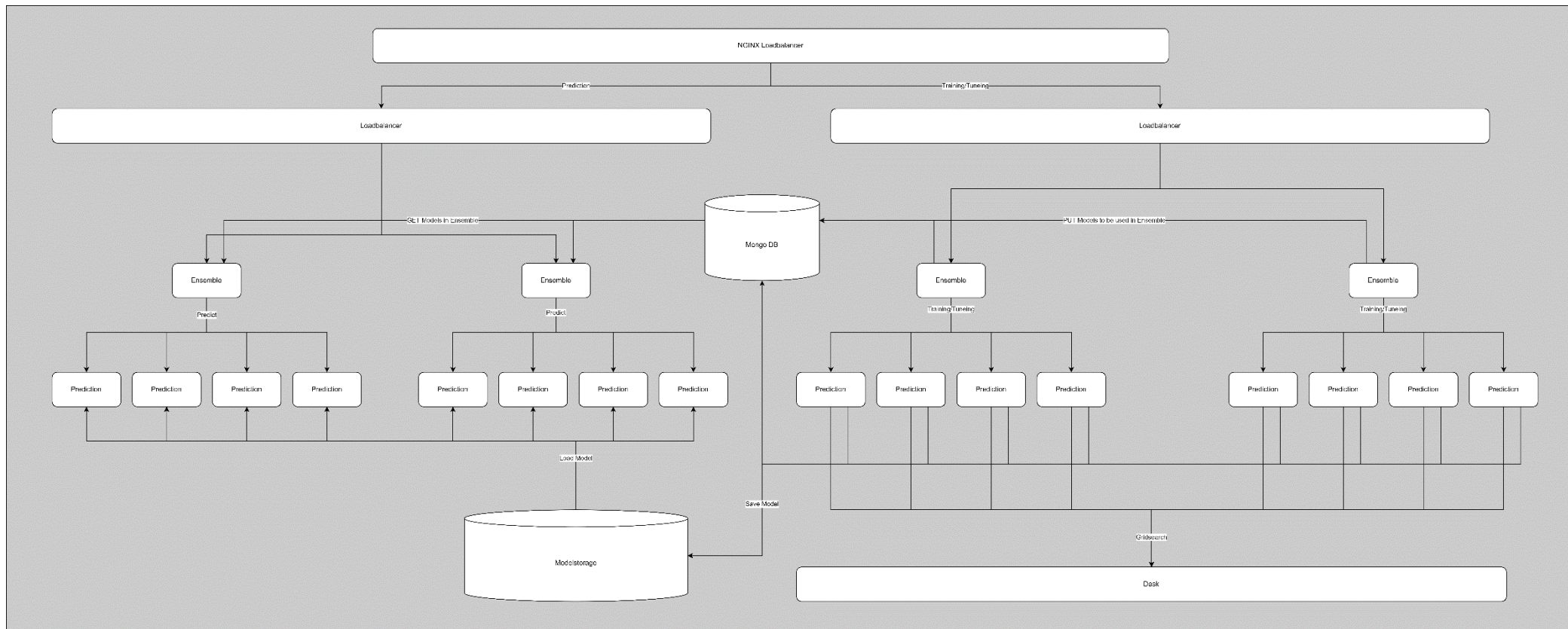


Abbildung 58: Architekturbild des Prognosemoduls

Zunächst soll die Software beschrieben werden, die in den Prediction-Pods läuft und die die einzelnen Algorithmen des Ensembles implementiert. Diese wurde mit Python entwickelt, damit externe Machine-Learning Bibliotheken eingebunden werden können, die in Python vielfach existieren. Dabei bietet die Software die Möglichkeit verschiedene Modelle zu entwickeln, um auf die Gegebenheiten einzelner Standorte einzugehen, wie bereits im Abschnitt 5.1.4 etabliert. Damit dies möglich ist, wurde eine Modell-Superklasse entwickelt, die über Callbacks und Parametrisierungen bei jeder Instanziierung die Möglichkeit bietet, ein anderes Feature-Engineering oder andere Algorithmen zu nutzen. Trotz dieser Flexibilität kann ein Modell, das aus dieser Superklasse erbt, direkt auf dem Prognose-Pod verwendet werden. Dabei ist jedes Modell für jeden Standort individuell getunt und somit für sich abgeschlossen. Die Parameter für die verschiedenen Modelle werden dafür in der MongoDB gespeichert, ebenso wie der Pfad zu der Datei, in welcher der spezifische, trainierte Algorithmus gespeichert wird. Diese Algorithmen werden auf der zentralen virtuellen Festplatte gespeichert.

Damit das Ensemble die einzelnen Prognose-Pods ansteuern kann, wurde eine Schnittstelle über FastAPI implementiert. Zur Steuerung der Modelle bietet diese API drei Routen an, eine zum Trainieren, zum Tunen und für die Vorhersage der einzelnen Modelle.

Der Ensemble-Pod, in dem die Software des Ensembles läuft, die ebenfalls in Python entwickelt wurde, vereint die Ergebnisse der einzelnen Prognose-Pods. Dadurch wirkt die Ansteuerung von Anfragen zu einem Messstandort über die Schnittstelle des Ensembles, als ob es nur ein zugrundeliegendes Modell geben würde. Die Steuerung des Ensembles ist dabei wieder über FastAPI realisiert und verfügt über die gleichen drei Routen wie die einzelnen Prognose-Pods zur Steuerung des Ensembles pro Messstandort.

Die erste hier erläuterte Route ist die Tune-Route. Bei einer Anfrage über diese Route werden alle im gesamten Ensemble zur Verfügung stehenden Modelle in den Prognose-Pods getunt. Dabei wird für einen angefragten Standort eine Optimierung der Modelle für diesen Standort angestrebt. Nach dem durchgeführten Tuning der Modelle folgt eine automatisierte Evaluation, die ebenfalls noch in den Prognose Pods durchgeführt wird. Diese spielen die Evaluation dann an den Ensemble-Pod, wo die Ergebnisse ausgewertet werden und entschieden wird, welche Modelle in das aktive Ensemble aufgenommen werden. Dieser Vorgang wird regelmäßig mit einigem Abstand für alle Standorte wiederholt, ist allerdings rechenintensiv.

Als zweite Route soll die Train-Route näher beschrieben werden. Durch diese werden die im aktiven Ensemble liegenden Modelle neu trainiert. Dabei berücksichtigen sie alle im Datenhub vorhandenen Messdaten, deren Umfang sich seit dem letzten Training vergrößert haben sollte. Dadurch sollte sich die Qualität der Vorhersagen wegen der größeren Datengrundlage im Training verbessern. Nach einem erfolgreichem (Re-)Training übermitteln die Prognose-Pods die ID der neu trainierten Modelle an den Ensemble-Pod, welche diese in der MongoDB persistiert. Aufgrund des geringeren Rechenbedarfs kann dieser Prozess in einem geringeren Abstand wiederholt werden als das (Re-)Tuning.

Die dritte angesprochene Route ist die Vorhersage-Route. Mit den trainierten Modellen des aktiven Ensembles für einen Standort kann über diese Route eine Vorhersage ausgelöst werden. Dafür werden die einzelnen Vorhersagen der Modelle im aktiven Ensemble aus deren Prognose-Pods gesammelt. Im Ensemble-Pod werden diese dann zu einer einzigen Zeitreihe verarbeitet, die von dem Ensemblemodul zurückgegeben wird. Dieser Prozess wird jede Stunde angeregt, sodass stets aktualisierte Vorhersagen gemacht werden, die die aktuellen Live-Auslastungen und Wetterverhältnisse berücksichtigen. Diese Vorhersagen werden dann auf einem FIWARE Context Broker zwischengespeichert. Dadurch ist eine Prognose mit sehr geringer Latenz verfügbar.

Wie bereits angesprochen, ist der Prozess des Tunings von Modellen sehr rechenintensiv. Zur Beschleunigung dieses Prozesses wurde ein Dask-Cluster verwendet. Über die dadurch ermöglichte Parallelisierung kann die Größe des Clusters reduziert werden, das die im Projekt entwickelte, spezifische Software betreibt. Das Dask-Cluster kann zudem in solchen Zeiten, in denen kein Tuning stattfindet, für andere Zwecke verwendet werden. Somit ist eine deutlich effizientere Nutzung der zur Verfügung stehenden Ressourcen möglich.

Zusammenfassend ermöglicht die beschriebene technische Umsetzung ein skalierbares und robustes Deployment des PoC. Für einen potentiellen produktiven Einsatz des Prognosemoduls bietet es sich an, sich an dieser erfolgreich explorierten Architektur zu orientieren.

5.1.6 Learnings und Potentiale

Während der Durchführungszeit von LAB-TOUR SH haben die beteiligten Verbundpartner* die Grundlagen für eine Auslastungsprognose an touristischen Anziehungspunkten gelegt. Allerdings hat das Forschungsprojekt auch einige Fragen aufgeworfen, deren Klärung vor einem produktiven Einsatz der Prognose stattfinden sollte. Durch die getane Arbeit lässt sich nun aber ein Status Quo zeichnen, anhand dessen der Weg des Prognosemoduls in die Schleswig-Holsteinische Tourismuspraxis geplant werden kann. Der folgende Abschnitt soll die wichtigsten Problematiken, in Bezug auf die Auslastungsprognose, zusammenfassen, die in diesem Projekt aufgetreten sind. Weiterhin werden teils bereits praktisch explorierte, teils theoretische Lösungsvorschläge für eben diese Problematiken gemacht.

Continuous Feed-Back-Loop

Für eine fundierte Evaluation der verschiedenen Mess- und Data Cleaning Methoden muss ein kontinuierlicher Vergleich mit tatsächlich beobachteten Auslastungswerten stattfinden. Um dabei Ressourcen effizient einzusetzen scheint es sinnvoll, Tourismusfachkräften vor Ort über eine Schnittstelle zur Integrationsplattform, die beispielsweise über eine App oder Website mit einer Nutzeroberfläche ausgestattet werden kann, die Möglichkeit zu geben, beobachtete Auslastungen zurückzumelden. Über einen solchen Feedback-Loop der Destinationen kann dann gezielt ein Vergleich der tatsächlichen Auslastung, mit der durch verschiedene Sensorik und Data Cleaning Methoden bestimmten Auslastung angestrebt werden. Über die reine Evaluation hinaus, könnten diese rückgemeldeten Auslastungen auch mit in das Training der Prognosemodelle eingehen, um dort schon während des Evaluationszeitraums die Modelle dahingehend zu beeinflussen, realistischere Auslastungen vorherzusagen.

Propagieren von Auslastungen

Das größte Hindernis, das LAB-TOUR SH für die Überwachung von Parkplatzauslastungen zur Bestimmung der touristischen Auslastung der Umgebung sichtbar gemacht hat, und für das kein klarer Lösungsweg skizziert werden kann, ist, dass unklar ist, wie die Auslastung eines Parkplatzes auf die Auslastung von POI im Umkreis verteilt werden soll. Im Abschnitt 4.2.2 wurde die Thematik bereits unter dem Schlagwort „Verknüpfung“ eingeführt. Es geht dabei um die Fragestellungen, an welche POI ein Parkplatz seine Auslastung propagiert und was die mathematische Funktion ist, die auf die Auslastung des Parkplatzes angewandt werden muss, um die Auslastung von Parkplätzen auf die Auslastung der POI übertragen zu können. Konkret geht es dabei um folgende Fragestellung: Wenn ein Parkplatz mit 100 Stellplätzen neben einem Strand liegt, der ca. 1.000 Menschen komfortabel Platz bieten könnte, und der Parkplatz komplett voll ist, welche Auslastung wird dann dem Strand zugeschrieben? Hierfür

lassen sich verschiedene Argumentationsstrategien aufbauen, von denen keine zu einem zufriedenstellenden Ergebnis führt. Es könnte argumentiert werden, dass in jedem Auto im Schnitt zwei Menschen saßen und alle an den Strand wollten, der damit zu 20% gefüllt ist. Damit würde sich die gemessene Auslastung des Strandes allerdings immer zwischen 0-20% bewegen und damit nie gemeldet werden, dass der Strand überfüllt ist. Weiterhin könnte die prozentuale Auslastung des Parkplatzes linear auf die des Strandes übertragen werden. Damit würde bei einem annähernd vollen Parkplatz auch der Strand schon als voll gemeldet werden, um Autofahrer*innen davon abzuhalten, in dieses Gebiet zu fahren. Es ist allerdings sehr wahrscheinlich, dass der Strand, selbst bei einem vollen Parkplatz, nicht übermäßig voll ist. Wenn also jemand die tatsächliche Auslastung des Strandes mit der angezeigten Auslastung im Recommender vergleicht, verliert dieser an Glaubwürdigkeit. Weiterhin wird die Chance verpasst, beispielsweise Radfahrern mitzuteilen, dass der Strand wenig gefüllt und für einen Ausflug geeignet ist.

In diesem PoC wurde für alle POI im Umkreis der Mittelwert aller Auslastungen aus WLAN-Accesspoints und Parkplätzen im Umkreis von 2 km als Auslastung angenommen. Dies hat sich allerdings als sehr unspezifisch erwiesen. Weiterhin ist denkbar, den Einfluss eines Parkplatzes oder einer Messstation mit steigender Entfernung abnehmen zu lassen und ein manuelles, auf Erfahrung basierendes Mapping einzuführen, welche Messstationen sich proportional zu welchen POI verhalten. Für eine solche Methode wäre es allerdings essentiell, dass möglichst flächendeckend, an vielen Parkplätzen die Auslastung gemessen wird. Außerdem würde so ein großer Teil der POI im Recommender keine Auslastung zugeschrieben bekommen und dementsprechend immer als neutral angegeben werden, unabhängig von der tatsächlichen aktuellen Auslastung.

Prognose

Auch für die tatsächliche Prognose lassen sich hier einige Punkte noch einmal besonders hervorheben. Das vorgeschlagene Prognosemodell liefert bereits gute Ergebnisse, die in der Lage sind, den Tages-, Wochen- und Jahresgang der Auslastung an unterschiedlichen POI abzubilden. Auch können unterschiedliche Dynamiken wie die Auslastung an Parkplätzen in Scharbeutz (ganzjährig immer wieder stark frequentiert) und die Auslastung am Strand von Scharbeutz (nur im Hochsommer stark frequentiert) automatisiert abgebildet werden. Das Auftreten von Auslastungs-Peaks ist jedoch stark vom Wetter und einer Eigendynamik abhängig und kann damit nicht immer präzise vom Prognosemodul weit im Voraus vorhergesagt werden. In der Evaluation der Daten hat sich gezeigt, dass sich die Auslastung des Strandes in Scharbeutz mit annähernd gleichen Rahmenbedingungen (Wetter, Uhrzeit, Wochentag) um viele tausend Menschen unterscheiden kann. Dies erschwert eine Vorhersage der Auslastung weit im Voraus. Zur visuellen Einschätzung der Prognosequalität wird in Abbildung 59 eine Gegenüberstellung der Prognosewerte und der Messwerte während des gesamten Messzeitraumes durchgeführt. Dabei liegt der Prognosehorizont in der oberen Grafik bei vier Tagen, während die untere Grafik einen Prognosehorizont drei Stunden aufweist.

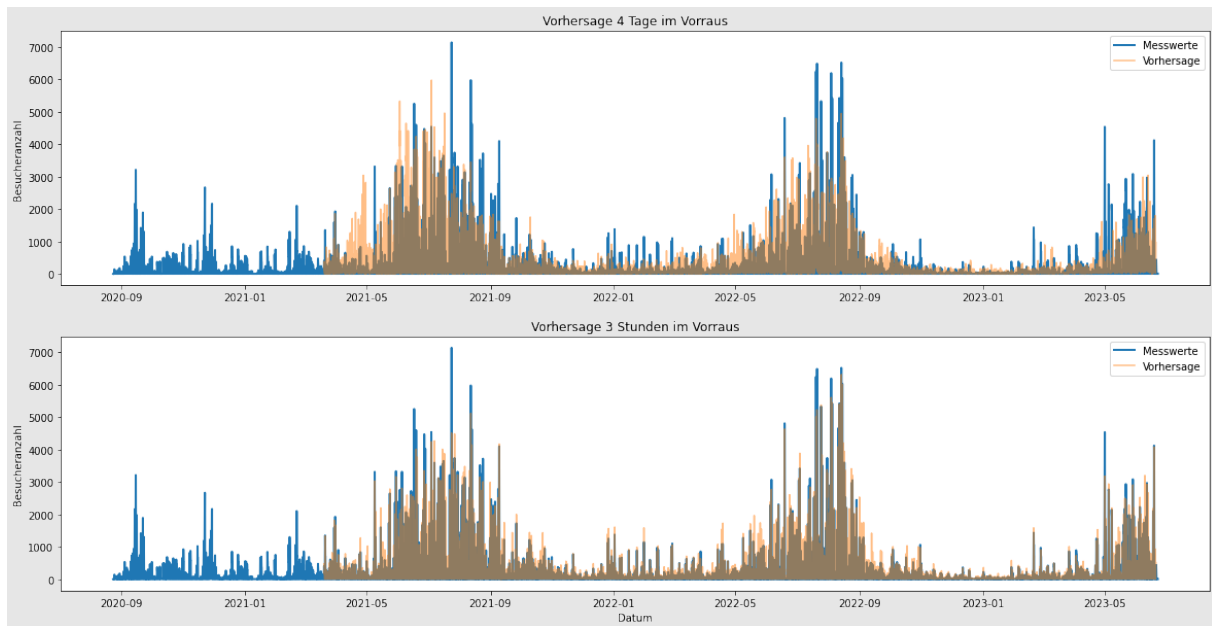


Abbildung 59: Prognose und Messwerte im Zeitverlauf

Um ein deutliches Bild der Prognosequalität zu schaffen, zeigt Abbildung 60 ebenfalls die Gegenüberstellung der Prognosewerte und Messwerte, jedoch für den Zeitraum Anfang Juni bis Anfang Oktober 2022.

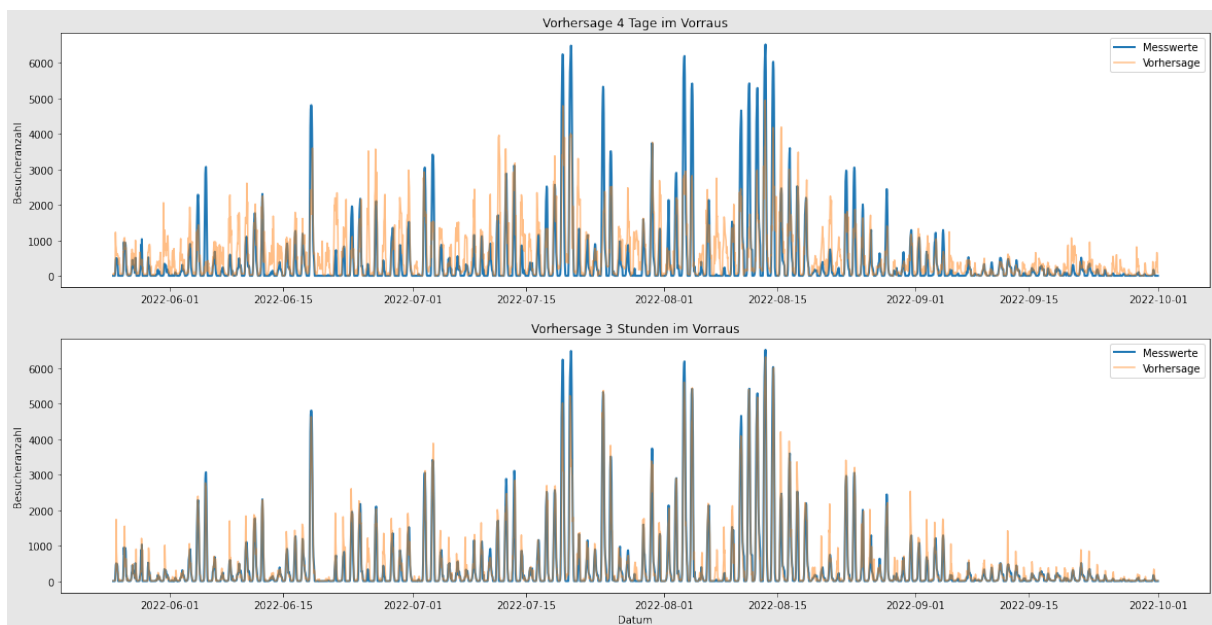


Abbildung 60: Prognose und Messwerte im Zeitverlauf, Juni bis Oktober 2022

Dies zeigt, dass das explorierte Modul vor allem bei kurzfristigeren Vorhersagen auch Peaks gut identifizieren kann und somit durch diese Prognosen auch Lenkungsmaßnahmen ausgelöst werden können, um das Besucheraufkommen zu entzerren. Allerdings bietet das vorgeschlagene Modul auch Raum für Verbesserungen.

Beispielsweise sollte der Pool an Algorithmen, auf den das Ensemble zugreifen kann weiter vergrößert werden und nach weiteren Verbesserungen für die potenziellen Ensemblemitglieder gesucht werden. Auch sollte evaluiert werden, wo der beste Trade-Off zwischen dem ressourcenaufwändigen Tuning der Modelle und der Qualität der Vorhersage liegt. Über die im PoC implementierte Trennung von Long- und Short-Term Modellen kann es weiterhin zu einem deutlichen Versatz der Prognosen für die Werte von sechs und sieben Stunden in die Zukunft kommen. An diesem Punkt hören die Short-Term Modelle auf in die Prognose einzugehen, die ab diesem Zeitpunkt nur noch von den Long-Term Modellen gemacht wird. Sollten Long- und Short-Term Modelle deutlich unterschiedliche Prognosen abgeben, würde sich dies hier bemerkbar machen. Die letzte hier genannte Verbesserung folgt dem wissenschaftlichen Anspruch, die Qualität des Prognosemoduls qualitativ einzuschätzen. Dies ist bislang nicht aussagekräftig möglich, da an vielen Standorten wenig Messdaten vorliegen und kein klares Evaluationsschema etabliert ist. In ein solches müsste klar mit eingehen, ob Tageswerte stärker als Nachtwerte gewichtet werden sollen, Sommerwerte mehr als Winterwerte, ob es einen Unterschied macht ob, zu viele oder zu wenige Besucher*innen vorhergesagt wurden etc.

Viele dieser Punkte werden zukünftig im bundesweiten Forschungsprojekt AIR aufgegriffen, das für eine touristische Auslastungsprognose das in LAB-TOUR SH entwickelte Modul weiter ausbauen und evaluieren wird. Weiterhin wird in diesem Rahmen beispielsweise auch eine Masterarbeit an der Fachhochschule Kiel erstellt, die sich dieser Thematik annimmt.

5.2 Alternativengenerierung

Für das Gesamtkonzept des LAB-TOUR SH Recommender ist die Vorhersage von Auslastungen an touristischen Attraktionspunkten essenziell, um die Zielsetzung zu erfüllen. Neben dieser Funktionalität existiert allerdings noch eine weitere, grundlegende Logik im Recommender, die unter anderem auch auf den Ergebnissen der Vorhersagen aufbaut. Gemeint ist damit die Auswahl von alternativen POI für einen bestimmten touristischen Attraktionspunkt. Anders als im Projektantrag beschrieben, wird dieser Funktionalität hier ein eigener Abschnitt gewidmet. Hintergrund dessen ist, dass die Logik hinter dieser Auswahl, wenngleich auf anderen Algorithmen basierend, eine ähnliche Komplexität wie die Auslastungsprognose einnimmt. Damit entspricht der Bericht der technischen Implementierung insofern, als das auch dort die Auslastungsprognose und die Alternativengenerierung logisch voneinander getrennt und in Submodule des Recommenders verpackt wurden („Divide and conquer“-Ansatz). Der Begriff „Alternativengenerierung“ wurde hier eingeführt um eine klare Abgrenzung dieses Submoduls zum (Gesamt-)Recommender, wie er im Projektantrag verstanden wird, zu schaffen. Der Recommender beschreibt dort die Gesamtheit der ausgespielten Information zu einem POI. Dies umfasst die Auslastungsprognose, alternative POI und statische Kontextinformationen zu dem ausgewählten POI. Der hier thematisierte Alternativengenerierer ist jedoch ausschließlich für die Auswahl der alternativen POI zuständig. Weiterhin gilt der Begriff „Alternativengenerierer“ auch der technischen Abgrenzung zu einem „Recommender“. Darunter werden in der Informatik weitestgehend Systeme verstanden, die auf Grundlage von Nutzerdaten arbeiten (bspw. kollaboratives Filtern), was der im Projektantrag erwähnte und prototypisch entwickelte Alternativengenerierer nicht tut. Über die Alternativengenerierung sollen für jeden POI passende, ebenfalls für den/die Nutzer*in relevante und zudem nahe gelegene POI ausgespielt werden. Über das Prinzip des Nudgings, also der gezielten Auswahl und Präsentation der ausgespielten alternativen POI, soll so ein Einfluss auf das Verhalten der/die Nutzer*in genommen werden. Entgegen der Grundidee bei Projektbeginn, Alternativen nur dann auszuspielen, wenn eine Überfüllung an einem POI vorhergesagt oder bereits eingetreten ist, werden nun dauerhaft alternative POI ausgespielt. Dafür muss allerdings gewährleistet sein, dass unter den so ausgespielten POI solche POI benachteiligt werden, die regelmäßig an ihre Kapazitätsgrenzen stoßen oder in Regionen liegen,

die regelmäßig an ihre Kapazitätsgrenzen stoßen. Werden solche POI von den Empfehlungen ausgeschlossen und dann aber fortwährend Alternativen angeboten, steigt die Wahrscheinlichkeit zumindest einen Teil der Nutzer*innen weg von Orten mit hohem Gästeaufkommen und hin zu Orten mit weniger Auslastung zu lenken. Im Idealfall lädt der Recommender so zum Stöbern ein und präsentiert attraktive POI, die allesamt Kapazität für mehr Besucher*innen haben.

5.2.1 Anforderungen an einen Recommender

Um ein angemessenes Set alternativer POI in Schleswig-Holstein effektiv auszuwählen und zu präsentieren, müssen bestimmte Voraussetzungen erfüllt sein. Im Wesentlichen zielen die Anforderungen darauf ab, dass die präsentierten alternativen Sets dazu beitragen, die Attraktivität von Schleswig-Holstein als touristisches Ziel zu steigern. Dies sollte jedoch nicht nur ein kurzfristiges Ziel sein. Vielmehr wird eine nachhaltige Verbesserung der Standortattraktivität angestrebt, die die natürliche Umgebung schützt und die Belange der Anwohner respektiert. Die gewählten alternativen POI sollten somit einen positiven, dauerhaften Einfluss auf die touristische Anziehungskraft Schleswig-Holsteins haben. In den nachfolgend aufgelisteten Punkten sind zusätzlich auch einige Grundvoraussetzungen an ein Recommender-System im Allgemeinen festgehalten. Die Anforderungen wurden nummeriert und verschlagwortet. Dabei gibt die Nummerierung nur sehr lose die subjektive Wichtigkeit der Anforderungen an und kann für unterschiedliche Nutzungsszenarien durchaus schwanken. Für spezifische Nutzungsszenarios kann die Liste an Anforderungen sich potenziell auch als unvollständig erweisen.

1. **Bereicherung.** Das Set an ausgespielten Empfehlungen für einen POI soll eine tatsächliche Bereicherung für die Nutzer*innen darstellen. Da das Kernprinzip des in diesem Projekt skizzierten Recommenders den Anspruch hat, so wenig invasiv wie möglich zu sein, liegt die tatsächliche Planung ihrer Ausflüge stets bei den Nutzer*innen. Aus dem zugrundeliegenden Prinzip des Nudging kann der Recommender also nur einen Effekt haben, wenn die ausgespielten Alternativen so bereichernd sind, dass Nutzer*innen sich selbst aktiv dafür entscheiden, ihren Ursprungs-POI durch einen der vorgeschlagenen POI zu ersetzen. Weiterhin kann sich eine Wirkung auch verzögert entfalten, wenn Nutzer*innen sich bei einer späteren Ausflugsplanung an einen der vorgeschlagenen POI erinnern.
2. **Nachhaltigkeit.** Die ausgespielten Alternativen sollen ein möglichst nachhaltiges Verhalten begünstigen. Sie sollen generell zu einer Umverteilung der Tourismusströme führen, sodass Crowding oder Congestion weniger wahrscheinlich auftritt und touristische Kapazitäten gleichmäßig genutzt werden. Durch die Umverteilung von Tourismusströmen, aber auch durch eigens dafür implementierte Ansätze (bspw. Vorschlagen von Parkplätzen für jeden POI) soll weiterhin der Parksuchverkehr an stark frequentierten Tourismusgebieten gemindert werden und so die Emission von gesundheits- und klimaschädlichen Treibhausgasen eingeschränkt werden. Um diese noch weiter einzuschränken, sollten alternative Ziele nahe am Ursprungspol oder - noch besser - sogar nahe am Ausgangspunkt des Nutzers oder der Nutzerin liegen und optimalerweise mit dem öffentlichen Nahverkehr zu erreichen sein. Je nach Datentiefe ist es auch denkbar, besonders nachhaltige POI im Recommender besonders kenntlich zu machen, beispielsweise durch ein kleines Icon im Frontend (bspw. ein grünes Blatt) Die entsprechende Definition eines „besonders nachhaltigen“ POI muss von der Tourismusforschung in Schleswig-Holstein im Kontext eines solchen Recommenders allerdings noch erarbeitet werden.

3. **Präsentation.** Generell ist es wichtig, die Alternativen anschaulich und attraktiv darzustellen. Dabei sollten einem/r Nutzer*in schnell die Kerninformationen einer Alternative zugänglich sein. Dazu gehört mindestens ein Titel, der entsprechende POI-Typ (Strand, Café, Museum, ...) und ein Bild. Das Bild sollte dabei einen ersten Eindruck des POI vermitteln. Je nach Möglichkeit wären auch Informationen wie die aktuelle Auslastung (bspw. über eine farblich codierte Auslastungsampel, wie sie weitläufig verwendet wird) oder die Entfernung der Alternative zum Ursprungs-POI Sinn machen. Alle diese Informationen sollten über ein nutzerfreundliches, modernes Design präsentiert werden, sodass ein/e User*in sich die Alternativen auch gerne anschaut. Hierbei gilt wieder, dass nur wenn die Präsentation der POI es schafft einem/r Nutzer*in Lust am Stöbern zu machen, der Recommender seine intendierte Wirkung entfalten kann.
4. **Wetter- und Jahreszeiteignung.** Das ausgespielte Set an Alternativen sollte sowohl an die Jahreszeit als auch an die aktuelle Wetterlage angepasst sein. Dabei sollte berücksichtigt sein, dass zum Beispiel die Ausspielung einer Badestelle im Winter wenig sinnvoll ist, ebenso wie bei schlechtem Wetter präferiert Indoor-Alternativen vorgeschlagen werden sollten.
5. **Erreichbarkeit.** Es muss gewährleistet sein, dass die vorgeschlagenen POI gut erreichbar sind. Optimalerweise könnte sogar direkt ein Vorschlag gemacht werden, wie der öffentliche Nahverkehr genutzt werden kann, um den POI zu erreichen. Zur Anforderung der Erreichbarkeit gehört weiterhin, dass Alternativen entweder möglichst nah an dem Ursprungs-POI liegen oder sogar näher am Ausgangspunkt des Nutzers oder der Nutzerin, falls dieser bekannt ist. Dies hat zum einen den Hintergrund einem/r Nutzer*in eine angenehme Anreise zu ermöglichen, zum anderen aber auch die Intention, die Anreise nachhaltiger zu gestalten.
6. **Öffnungszeiten.** Nachdem sich ein/e Nutzer*in einmal dazu entschieden, hat einen vom Recommender ausgespielten POI dem ursprünglichen POI vorzuziehen, ist es für eine langfristig wiederkehrende Nutzung des Recommenders enorm wichtig, dass der ausgewählte POI den Erwartungen weitestgehend entspricht. Während sich vom Recommender aber nicht kontrollieren lässt, ob ein/e Nutzer*in mit dem POI zufrieden ist, sollte zumindest gewährleistet sein, dass ein POI besucht werden kann. Dafür ist es unabdingbar, mindestens die Öffnungszeiten im Frontend anzuzeigen. Noch besser wäre es POI gar nicht erst als Alternative anzuzeigen, wenn sie beispielsweise saisonal oder an einem, für die Empfehlung ausgewählten, Tag nicht geöffnet sind. Zusätzlich könnten Warnhinweise ausgespielt werden, wenn ein POI zwar noch offen ist, aber bald schließt.
7. **Use Case spezifische Kriterien.** Je nach Use Case einer Alternativengenerierung kann es durchaus sein, dass spezifische, nur für diese Use Cases relevante Randbedingungen mitgedacht werden müssen. Beispielsweise muss als spezifisches Kriterium an der Nordseeküste von Schleswig-Holstein die Erreichbarkeit einiger POI im Kontext von Ebbe und Flut bei einer Empfehlung gewährleistet sein. So sollten bei Flut beispielsweise keine Wattwanderungen vorgeschlagen werden. Außerdem muss die Sicherheit der Nutzer*innen gewährleistet werden und so beispielsweise Warnhinweise bei entsprechenden, risikobehafteten Tätigkeiten mitgeliefert werden. Bei Radrouten zum Beispiel muss mitgedacht werden, dass User*innen diese sehr wahrscheinlich mit dem Fahrrad von ihrem Ausgangspunkt aus erreichen wollen und dass es hier verschiedene Ansprüche an die Schwierigkeit einer Radroute gibt. Solche Randbedingungen für verschiedene Use Cases treten bei genauerer Betrachtung regelmäßig auf und sollten mitgedacht sein, bevor der Alternativengenerierer um den jeweiligen POI-Typ erweitert wird.
8. **Barrierefreiheit.** Um möglichst vielen Menschen einen Mehrwert zu bieten, sollte der Recommender über ein intuitives Web-Frontend erreichbar sein. Das Frontend sollte leicht zu finden,

zu bedienen und zu interpretieren sein. Um den Zugang zum Recommender so leicht wie möglich zu machen, sollte dieser sowohl über gängige Suchmaschinen als auch von Destinationswebsites aus leicht zu finden sein. Durch gute Designüberlegungen und eine Funktionsweise die auch ohne User-Input arbeiten kann, können Nutzer*innen mit wenigen Klicks erste Informationen und alternative POI zugänglich gemacht werden.

9. **Rechtskonformität.** Ein Recommender muss alle nötigen Ansprüche an seine rechtlichen Verpflichtungen erfüllen. An dieser Stelle wird auf eine ausführliche Aufzählung aller potenziell risikobehafteten Rechtsthemen verzichtet, um nicht den Eindruck von Vollständigkeit entstehen zu lassen. Ein offensichtliches Thema sind jedoch Lizenzen unter denen Stammdaten und Bilder zu POI veröffentlicht werden. Weiterhin muss zum Beispiel gewährleistet sein, dass bei der Ausspielung von Badestellen nur solche Badestellen vorgeschlagen werden, an denen das Baden erlaubt ist. Vor einer Inbetriebnahme sollte geprüft werden, ob (am Beispiel der Badestellen) der Recommender nur Badestellen vorschlagen darf, die von der DLRG überwacht werden oder ein Warnhinweis vorhanden sein muss, dass Nichtschwimmer dort nicht schwimmen sollten.
10. **Diskriminierungsfreiheit.** Das Auftreten von potenzieller Diskriminierung innerhalb eines Recommenders sollte in der Entwicklung solcher Systeme stets mitgedacht werden. Diskriminierung kann dabei vor allem in drei Hauptdimensionen auftreten. Diese Dimensionen beziehen sich auf die Interaktion zwischen Recommender und Endnutzer (*Endnutzer*), auf die Abhängigkeit der jeweiligen POI-Eigentümer vom Recommender (*POI-Eigentümer*) und auf die Repräsentation verschiedener Regionen Schleswig-Holsteins im Recommender (*Regionen*). Um die vorhergegangenen Dimensionen zu verdeutlichen, folgt nun jeweils ein Beispiel aus jeder Kategorie. (*Endnutzer*): Es sollte beispielsweise gewährleistet sein, dass Menschen mit körperlichen Einschränkungen den Recommender nutzen können. So sollte zum Beispiel klar sein, ob ein Restaurant oder Museum rollstuhlgerecht ist um, bei einer entsprechenden Abfrage, eine rollstuhlgerechte Alternative ausspielen zu können. Mindestens jedoch muss, falls der Recommender dazu nicht in der Lage ist, diese Schwäche klar kommuniziert werden. (*POI-Eigentümer*): Zwei POI derselben Art sollten gleiche Behandlung erfahren. Es wäre nicht richtig, wenn der Recommender ein Café dem anderen Café vorzieht und durch die kontinuierliche Ausspielung des vorgezogenen Cafés eine Bevorteilung dieses erfolgt. Sollte eine Bevorzugung eines POI gegenüber eines anderen intendiert sein (bspw., weil mehr Daten zu einem POI vorliegen oder dieser nachhaltiger ist), muss klar kommuniziert werden warum ein POI bevorzugt wird, sodass POI-Eigentümer potenziell Einfluss darauf nehmen könnten. (*Regionen*): Unterschiedliche Regionen in Schleswig-Holstein sollten, in Relation zu ihrem touristischen Potential, gleichmäßig im Recommender vertreten sein. Um nicht den Anschein von Vollständigkeit zu erzeugen, verzichtet dieser Bericht an dieser Stelle bewusst auf eine umfänglichere Aufzählung. Stattdessen empfehlen die Autoren nachdrücklich eine kritische Auseinandersetzung mit allen ethisch relevanten Punkten, die die produktive Umsetzung eines Recommenders zu einem spezifischen Use Case mit sich bringt.
11. **Transparenz.** Diese Anforderung ist eng verbunden mit dem vorherigen Punkt. Um Nutzende des Recommenders und Eigentümer der im Recommender repräsentierten POI nicht zu irritieren, sollte gewährleistet sein, dass der Alternativengenerierer sich erklärt. Um das zu tun, kann beispielsweise auf der Ausspielungswebsite offen nachzulesen sein, wie der Alternativengenerierer seine Entscheidungen trifft. Außerdem braucht es eine Möglichkeit, mit der Korrekturen von fehlerhaften Informationen (falsche Öffnungszeiten, Rechtschreibfehler, etc.) an die Betreiber des Recommenders weitergeleitet werden können und dort dann auch zügig in die entsprechenden Wissensbasis eingepflegt werden. Die Mindestanforderung für diesen Punkt ist,

ein im Frontend sichtbar verlinktes Dokument, in dem die Funktionsweise in leichter Sprache erklärt wird und eine offen sichtbare E-Mail-Adresse, an die sich jeder mit einem Anliegen wenden kann.

12. **Genauigkeit der Wissensbasis.** Die dem Recommender zugrundeliegende Wissensbasis sollte fortlaufend aktualisiert und überprüft werden, um so zügig neue POI aufzunehmen, Änderungen an den Metadaten eines POI einzutragen und hinfällige POI zu löschen. Generell muss dabei berücksichtigt werden, dass es schwerer wiegt eine falsche Information zu einem POI auszuspielen, auf die sich Nutzer*innen verlassen, als gar keine Information zu einem POI auszuspielen. Dies sollte von der datengenerierenden Organisation stets bedacht werden, sodass genügend Kapazität für die Pflege der Wissensbasis bereitsteht.
13. **Quantität der Wissensbasis.** Der Recommender sollte über eine umfangreiche Wissensbasis verfügen und möglichst viele der POI in Schleswig-Holstein abbilden können, um landesweit geeignete Alternativen ausspielen zu können. Dabei spielt auch die Verteilung der POI eine essenzielle Rolle. Speziell ist darauf zu achten, dass POI nicht nur in touristisch ohnehin schon stark beanspruchten Regionen digitalisiert vorliegen, sondern gerade auch der bislang touristisch noch nicht stark vermarktete Punkte im Recommender repräsentiert werden. Für einen möglichst hohen Umfang der Wissensbasis ist außerdem die Anbindung möglichst vieler Datenquellen wichtig, um bestehende isolierte Datensilos aufzubrechen und in einem Daten-Hub vereinen zu können.
14. **Anpassung mit Anfragehorizont.** Je nach Anfragehorizont sollten sich die Zusammensetzung des ausgespielten Sets an Alternativen unterscheiden. Je kurzfristiger eine Anfrage für eine Alternative ist, desto näher sollte eine vorgeschlagene Alternative, sowohl räumlich als auch thematisch, an dem Ursprungs-POI liegen. Nutzende, die sich sehr wahrscheinlich schon auf den Ursprungs-POI vorbereitet haben und weniger offen für andere Aktivitäten sind, können/wollen kurzfristig nicht mehr umplanen. Wenn der Zeitpunkt, für den eine Alternative gesucht wird, mehrere Tage in der Zukunft liegt, ergibt eine größere räumliche und thematische Freiheit im Recommender Sinn, da Nutzende entsprechend Zeit haben, um sich auf eine neue Situation einzustellen und dementsprechend flexibler sind. Es müssen also in Abhängigkeit des Anfragehorizontes unterschiedliche räumliche und thematische Alternativen vorgeschlagen werden.
15. **Qualität der Wissensbasis.** Neben dem Umfang spielt auch die Tiefe der Wissensbasis eine wichtige Rolle für die potenziell umsetzbaren Funktionalitäten des Recommenders. Damit ist gemeint, welche Information zu einem, in der Wissensbasis vorhandenen POI vorliegen, und auch mit welcher Stetigkeit Informationen dieses Typs vorliegen. Zu den grundlegenden Attributen eines POI gehören nach dem Datenmodell *Point of Interest* (POI) (siehe Abschnitt 4.3.2) in der Smart Data Models Domäne unter anderem folgende Attribute, zwingend nötig für eine Aufnahme in die Wissensbasis sind. Ein *Titel*, eine *Beschreibung*, eine *Adresse*, eine (oder mehrere) *Kategorien* (Strand, Museum, Kirche, Café, ...) und eine *Geolocation*. Weitere Attribute aus dem Datenmodell, die hinreichend, aber nicht notwendig sind, sind eine Verlinkung zu einer *Website* und eine *Kontaktmöglichkeit*. Darüber hinaus bieten die deutlich feingranularen Datenmodelle der POI-Kategorien in der ODTA-Domäne die Möglichkeit weitere Informationen zu speichern. So kann darüber zum Beispiel in der Wissensbasis festgehalten werden, ob ein Strand über Toiletten und Volleyballnetze verfügt oder zu welchen Zeiten ein Museum geöffnet ist. Grundsätzlich erscheint es sinnvoll, eine möglichst hohe Tiefe der Wissensbasis anzustreben, um die Daten vielseitig für verschiedene Anwendungsfälle nutzen zu können. Allerdings steht die Tiefe der Wissensbasis in Konflikt mit den begrenzten Ressourcen für die

Datengenerierung. Um diesem Konflikt zu begegnen, wird empfohlen, in der Wissensbasis eine gewisse Datentiefe über die Grundattribute hinaus anzunehmen, wobei jedoch die Reichweite der Wissensbasis prioritär behandelt werden sollte. Sobald konkrete Use Cases identifiziert wurden, die eine höhere Datentiefe in einigen Bereichen der Wissensbasis voraussetzen, können diese dann gezielt nachgetragen werden. Eine Ausnahme von dieser Empfehlung bilden Use Cases, die eine Begehung voraussetzen. Diese sollten von Vorneherein möglichst umfangreich eingetragen werden – auch hier gilt aber, sollten die Grundinformationen auch anders abrufbar sein (Websites, Google Maps etc.), sollte dieser Ansatz bevorzugt werden.

5.2.2 Datengrundlage

Im Projekt LAB-TOUR SH war es allen Projektpartner*innen von Anfang an wichtig, neben theoretischen Aspekten auch praktische Erfahrungen in die Konzeption eines landesweiten Besuchermanagementsystems einfließen lassen zu können. Um dies zu gewährleisten, wurde von vorneherein angestrebt, eine prototypische Implementierung (PoC) eines Recommenders innerhalb der kurzen Projektlaufzeit zu implementieren. Dieser Aufgabe haben sich die Verbundpartner Fachhochschule Kiel, LHIND und ADDIX gewidmet und haben in Abstimmung mit der Fachhochschule Westküste einen solchen PoC entwickelt. Anhand der Entwicklung dieses Prototyps sollen die nächsten Abschnitt einen praktischen und detaillierten Einblick – angereichert mit theoretischen, verallgemeinerbaren Überlegungen – in die Entwicklung des LAB-TOUR SH Alternativengenerierers geben, um so auf alle praxisrelevanten Details eingehen zu können. Die Datengrundlage (Wissensbasis) des Alternativengenerierers in diesem PoC besteht aus 2.308 POI, die in ganz Schleswig-Holstein verortet sind (vgl. Abb. 61).

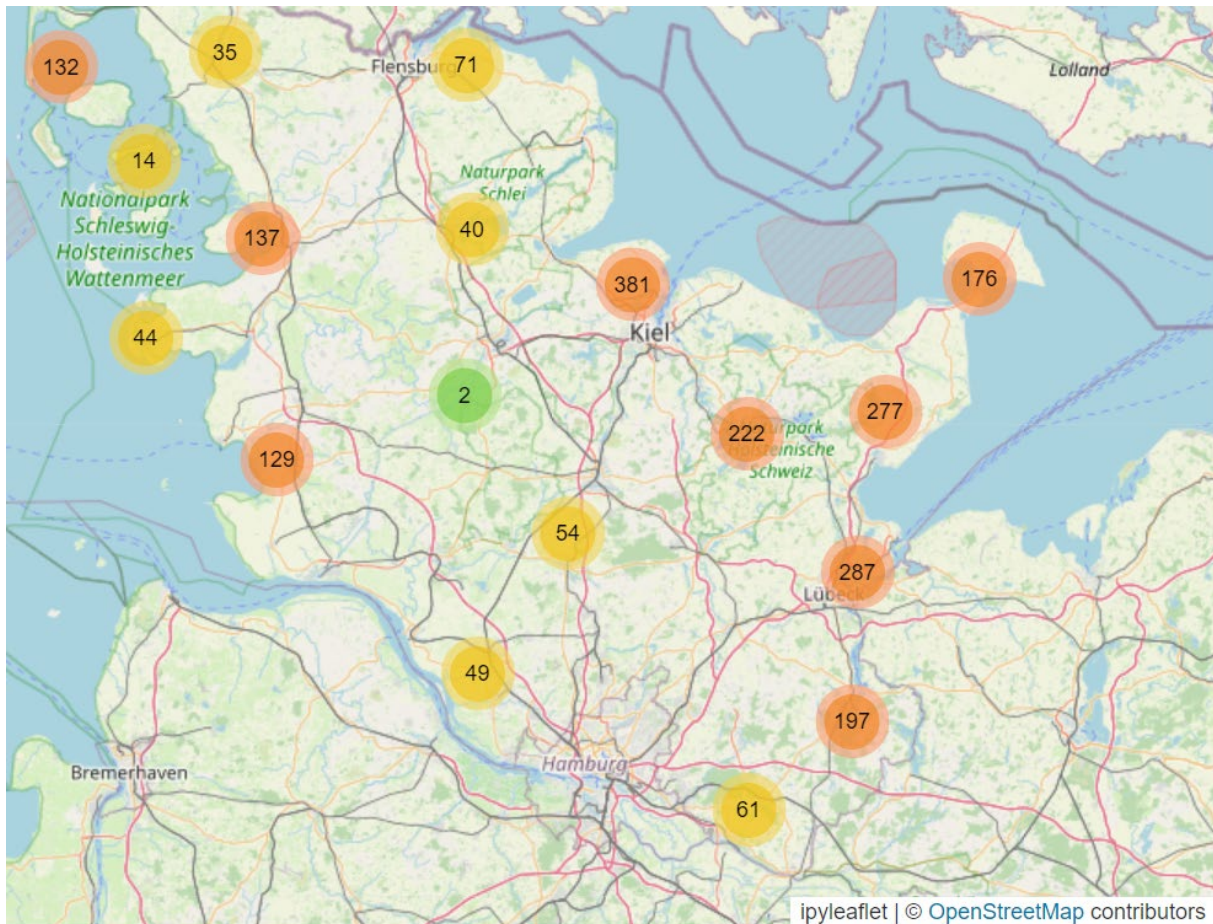


Abbildung 61: POI in Schleswig-Holstein

Entstanden ist diese Wissensbasis durch die Zusammenarbeit mit der Tourismusagentur Schleswig-Holstein (TA.SH), die dem Projekt alle in ihrer Datenbank vorhandenen POI zur Verfügung gestellt hat, die unter einer Open Data Lizenz stehen. Die Wissensbasis besteht dabei aus zwei Hauptteilen, für die unterschiedliche Schritte in der Datenbereinigung unternommen wurden. Etwa 1.000 POI der finalen Datengrundlage stammen aus dem Open Data Portal Schleswig-Holstein, in das die TA.SH alle POI aus ihrer Datenbank gespielt hat, die unter einer offenen Lizenz vorliegen und eine bestimmte Datentiefe erreichen. Diese Daten wurden manuell überprüft, um nachvollziehen zu können, wie aufwändig eine solche Prüfung in der Realität ist. Weitere 1.300 POI der finalen Datengrundlage wurden direkt aus der Datenbank der TA.SH exportiert. Auch sie liegen unter einer offenen Lizenz vor, erfüllen jedoch nicht die Datentiefe, die die TA.SH für einen Export ins Open Data Portal SH vorausgesetzt hat. Sie wurden in diesem PoC nicht aufwändig geprüft und dienen hauptsächlich der Erhöhung der POI-Dichte, um im PoC möglichst nah an eine für den Produktivbetrieb realistische Zahl an POI zu kommen und damit also in der Erprobung möglichst realistische Bedingungen zu schaffen.

Die Datengrundlage des PoC Recommenders teilt sich in 63 verschiedene POI-Typen auf. Unter POI wird "ein im Rahmen touristischer Reisen besuchenswerter geographischer Ort" (Reif, Schmücker, Staubert, 2023) verstanden. Dabei lassen sich Erlebnis-POI, die einen bestimmten Erlebniswunsch erfüllen (bspw. Sehenswürdigkeiten) und funktionale POI unterscheiden, die nicht als Hauptattraktionen dienen, sondern vielmehr dazu beitragen, die tatsächliche Reiseerfahrung zu ermöglichen. Obwohl sie von Reisenden frequentiert werden, dienen sie hauptsächlich als Werkzeuge oder Hilfsmittel zur Un-

terstützung der Reiseerfahrung (bspw. Parkplätze zur Erreichung einer Sehenswürdigkeit) (ausführlicher in Reif, Schmücker, Staubert, 2023). Im PoC sind drei POI der Gruppe funktionaler POI zuzuordnen (*Parkplatz*, *Wohnmobilstellplatz*, *Campingplatz*), die restlichen sechzig sind der Gruppe der „Erlebnis-POI“ zuzuordnen.⁸ Für einen Überblick über die Verteilung der POI in den POI-Typen wurden in Abbildung 62 die 15 häufigsten POI-Typen anhand der Anzahl der von ihnen repräsentierten POI aufgeschlüsselt.

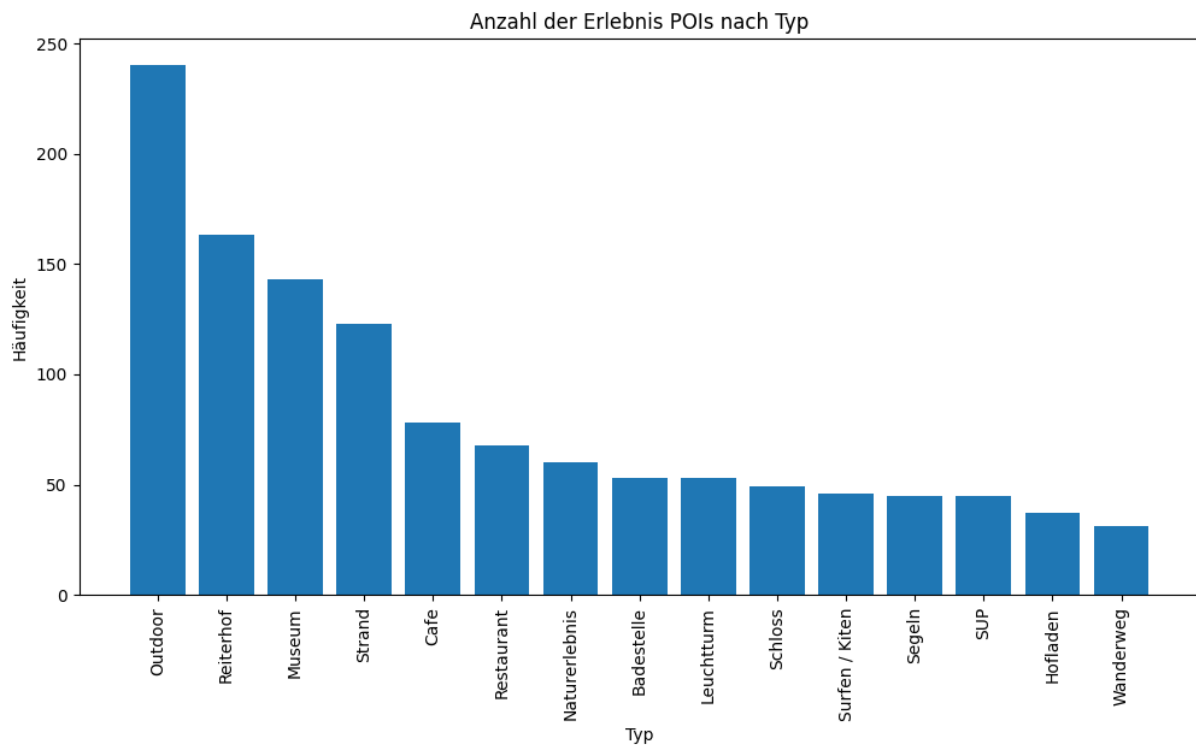


Abbildung 62: Anzahl von Erlebnis-POI nach Typ

Ausgewählt wurden diese POI-Typen anhand von Filterkriterien der TA.SH Datenbank, in der für jeden potenziellen Filter entschieden wurde, ob POI dieses Typs in die Wissensbasis eingehen sollen. Dabei wurde ein grundsätzlicher Fokus auf die Repräsentation von Erlebnis-POI gesetzt, die als Ausflugsziele gelten könnten. POI wie Supermärkte, Friseure, usw. wurden dementsprechend exkludiert. Die drei verbliebenden Kategorie-Typen *Parkplatz*, *Wohnmobilstellplatz* und *Campingplatz* wurden aufgenommen, um die konzeptionierten Ansätze zur Verringerung des Parksuchverkehrs testen zu können. Für manche POI-Typen liegt in der Datengrundlage nur eine einstellige Anzahl an POI vor. Dies liegt zum

⁸ Entsprechend der Häufigkeit ihres Vorkommens (beginnend mit dem Häufigsten) sind dies folgende POI-Typen: Outdoor, Reiterhof, Museum, Strand, Café, Restaurant, Naturerlebnis, Leuchtturm, Badestelle, Schloss, Surfen / Kiten, Segeln, SUP-Verleih, Hofladen, Wanderweg, eBike-Verleih, Therme, Kirche, Zoo, Freizeitpark, Fahrradverleih, Garten, Schwimmbad, Hafen & Marina, Galerie, Mode & Schmuckhandlung, Denkmal, Meierei & Käserei, Dekoration & Souvenir, Kanuverleih, Golfplatz, Theater, Räucherei, Science-Center, Schiff, Vogelkunde, Naturschutzgebiet / Naturpark, Aussichtspunkt, Herrenhaus, Freibad, Strandbar & Beach Club, Wellness Anwendungen, Fleischerei, Sauna, Tretbootverleih, Kneipe & Bar, Minigolf, Radweg, Buchhandlung, Spielplatz, Adventure Golf, diverse Gastronomie, Brennerei, Kletterpark, Erlebnishof, Imbiss, Bistro, Ruderbootverleih, diverse Kunst & Kultur und Imkerei

einen an dem geringen Umfang der Wissensbasis, zum anderen aber auch an den nicht optimal gewählten POI-Typen. Zur Vollständigkeit wurde zudem die Anzahl der funktionalen POI dieses Typs separat aufgeführt (vgl. Abb. 63).

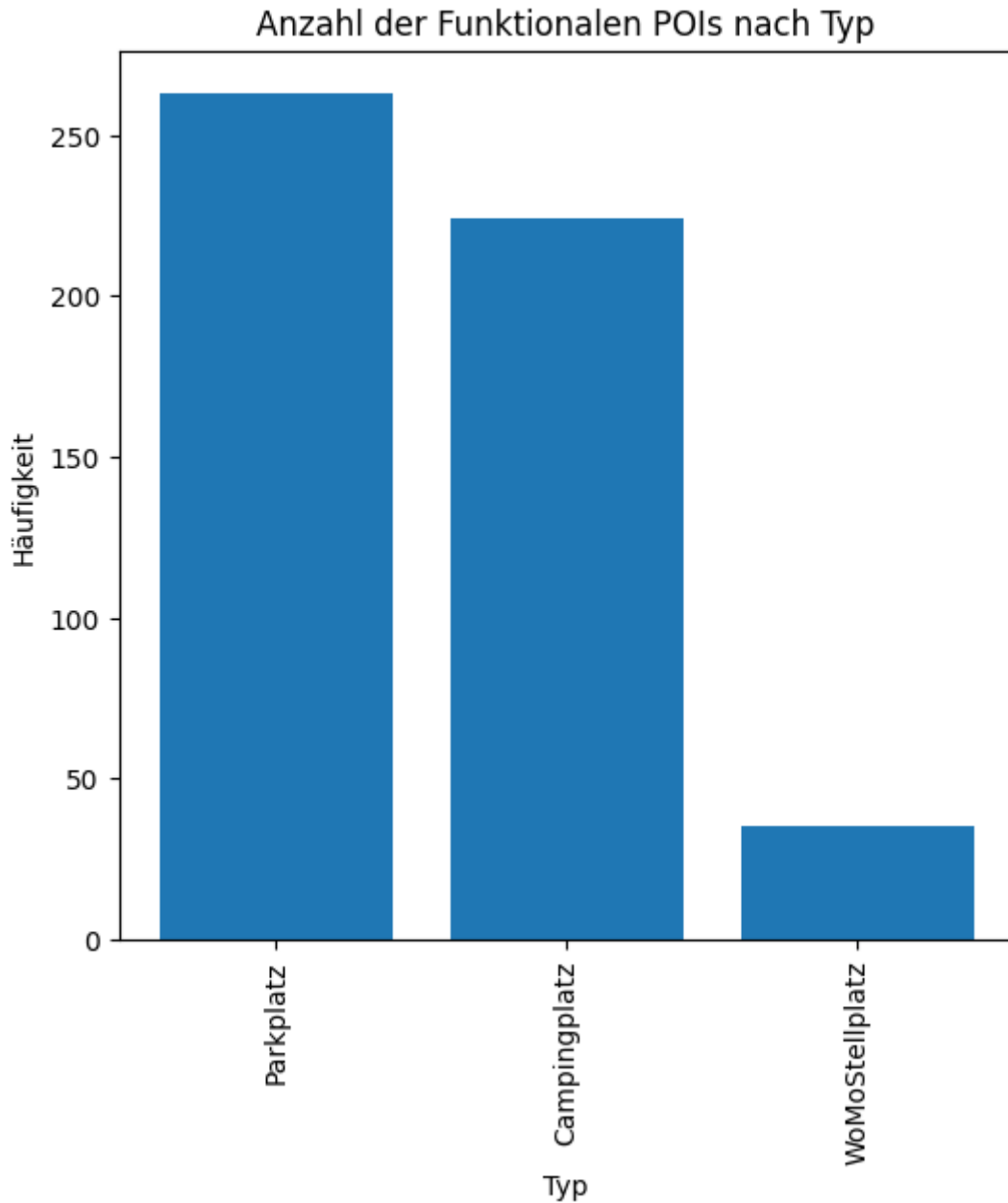


Abbildung 63: Anzahl von funktionalen POI nach Typ

5.2.3 Datenbereinigung

Ohne allzu sehr auf die technischen Details einzugehen, soll das folgende Abschnitt einen Überblick über die spezifischen Schritte der Datenbereinigung geben, die in LAB-TOUR SH für die Entwicklung des PoC durchgeführt wurde. Das Data Cleaning ist in vielen Data Science Projekten einer der wichtigsten Schritte der Datenwertschöpfungskette, in dem maßgeblich beeinflusst wird, wie gut das Endergebnis wird. Alle nachfolgend erläuterten Schritte zielen darauf ab, eine umfangreiche Datengrundlage

zu POI zu erstellen. Diese Datenbasis kann dann, gemäß den in dem Abschnitt 5.2.6 skizzierten Vorgaben, durch die dort beschriebenen Algorithmen verarbeitet werden.

Es ist gängige touristische Praxis, POI in ihrer digitalen Repräsentation mit Tags zu versehen, um sie so zu verschlagworten. Diese Tags folgen allerdings keiner festgelegten Nomenklatur und können von jeder datengenerierenden Organisation frei vergeben werden. Im konkreten Fall der Eintragung der POI in der TA.SH Datenbank waren mehrere Destinationen beteiligt, weshalb diverse Verschlagwortungssysteme dort nebeneinander existieren. Um diese Systeme an einem Beispiel aufzuzeigen, soll folgendes Muster in den Tags der POI genutzt werden. Eine der Tourismusorganisationen, die für die TA.SH ihre POI digitalisiert hat, hat dort den Tag *Schietwetter* genutzt, um anzuzeigen, dass ein POI auch bei schlechtem Wetter zu nutzen ist. Andere Organisationen haben dafür hingegen nur den Tag *Indoor* vergeben, der implizit auch angibt, dass ein POI auch bei schlechtem Wetter zu nutzen ist. Der Tag *Indoor* wurde jedoch nicht von der Organisation verwendet, die den Tag *Schietwetter* vergeben hat und andersherum. Schon anhand dieses Beispiels wird deutlich, dass allein aus einem Tag nicht eindeutig hervorgeht, ob ein POI für Indoor-Aktivitäten geeignet ist oder nicht. Zudem ist neben der bestehenden Heterogenität in der Verwendung von Tags auch bei einer konformen Nutzung der Tags deren Bedeutung nicht immer eindeutig. Abhängig vom Verständnis der jeweiligen tag-generierenden Person oder Organisation kann derselbe Tag in verschiedenen Kontexten verwendet worden sein. Der Tag *Radfahren* wird beispielsweise oftmals für das Taggen von Fahrradtouren, Fahrradgebieten und Fahrradverleihs genutzt – aber auch für diverse Cafés, Hotels und Herbergen, die sich wahrscheinlich gut in eine Radtour einbinden lassen oder sich als vom ADFC klassifizierte Beherbergungsbetriebe auszeichnen. Dementsprechend kann der Tag *Radfahren* nicht verlässlich genutzt werden, um alle Radrouten zu identifizieren. Zur Verdeutlichung der besprochenen Thematik folgt nun eine komplette Auflistung der in den Rohdaten vergebenen Tags zu drei ausgewählten POI.

- **Freizeitbad Brunsbüttel:** Wellness&Gesundheit, Freizeit- & Schwimmbäder, Therme, Produktlinien, Gesundheit & Wellness, NTS, individuelle Portale, Regionen, Bade&Strandurlaub, Dithmarschen, best of, Wellness&Entspannung, Nordsee Festland, Nordsee, Orte&Regionen, Themen, Touristische Landesdatenbank
- **Parkplatz L202 Frachtsee:** Outdoor, Natur erleben, Produktlinien, Radfahren und Wandern, Mobilität, individuelle Portale, HLMS, Parkplätze, Freizeit und Aktiv, Aktivität, Touristische Landesdatenbank, Mobil
- **Hotel Restaurant Dorfschänke:** Touristische Landesdatenbank, Hotels, Radfahren, Themen, Produktlinien, Bed & Bike, Teaser-Darstellungsoptionen, Radlerunterkunft, Übernachten, Service, Bett+Bike, Rendsburg, Quer / Klein, Hote, Module, Outdoor, TASH, Städte, Aktivität, individuelle Portale, Beherbergung, Landschaften, Sport, Binnenland

Durch bereits erwähnt und in diesen Beispielen klar sichtbare Ambiguität ist es schwergefallen, POI maschinell einem (oder mehreren) ganz konkreten Typen zuzuordnen. Während dies für einen Menschen oftmals schon aus dem Titel einfach zu folgern ist, ist dies für klassische Algorithmen deutlich schwieriger. Für einen ersten Versuch diese Zuordnung maschinell durchzuführen, wurde eine einfache, logische Baseline aufgesetzt. Zusätzlich wurde auch ein elaboriertes, auf Methoden des Machine Learnings und Textverständnis beruhendes Verfahren konzeptioniert und exploriert. Um zu der Baseline zu kommen, wurde das im Folgenden beschriebene Verfahren angewendet. In einem ersten Schritt wurden aus der Gesamtheit aller in der Datengrundlage vorhandenen Tags solche Tags ausgewählt, die klar auf einen Typ hinweisen, der zu den Erlebnis-POI gehört (*Therme, Freizeit & Gesundheit, Radfahren, ...*). Dann wurden rund 30 Regeln aufgestellt, die einen der vorher ausgewählten Tags löschen würden, wenn er gemeinsam mit einem oder mehreren vorher definierten, anderen Tags zusammen

bei einem POI auftauchen würde. Der Tag *Radfahren* wurde so zum Beispiel gelöscht, wenn einer der Tags *Beherbergung*, *Hotel*, *Hotels*, *Bed & Bike*, *Parkplätze*, *eBike-Verleih*, *Fahrradverleih*, *Café*, *Restaurant* ebenfalls bei einem POI vorkommen würde. So soll gewährleistet sein, dass diese POI die offensichtlich entweder Hotels, Parkplätze, Fahrradverleihs, Cafés oder Restaurants sind, nicht als *Radfahren* in die POI-Datengrundlage eingehen. Als *Radfahren* sollen dort nur Fahrradtouren und Radfahrgebiete hinterlegt sein. Diese Regeln aber so zu formulieren, dass sie alle in den POI-Daten und ihren Tags vorhandenen Muster abdecken, ist schwierig umzusetzen und sehr unwahrscheinlich vollständig. In einem dritten Schritt wurden dann alle solchen Tags, die nicht den im ersten Schritt ausgewählten Tags entsprechen, gelöscht. Daraus resultieren für die einzelnen POI drei Situationen. Entweder haben sie keinen Tag mehr, exakt einen, oder mehrere. POI ohne verbliebene Tags wurden aus dieser Datengrundlage gelöscht. Die Annahme liegt nahe, dass es sich bei Ihnen um POI handelt, die nicht in den Recommender aufgenommen werden sollten (da keine Erlebnis-POI), wie zum Beispiel Hotels und Supermärkte. Sollte genau ein Tag bei einem POI verblieben sein, wird dieser Tag zum Typ des POI. Für den Fall, dass mehrere Tags bei einem POI verbleiben (bspw. Museum mit angeschlossenem Café) wurde der POI vervielfacht, wobei jede Kopie einen der Tags als Typ zugewiesen bekommen hat. Dieses Verfahren führt zwar dazu, dass es denselben POI in der Datengrundlage mehrfach geben kann, wobei jede Version einen anderen Typ hat. Es führt allerdings auch dazu, dass klar ist wofür sich ein/e Nutzer*in genau interessiert, sodass entsprechende Alternativen ausgesprochen werden können. Für das Café des Museums können andere Cafés als Alternative vorgeschlagen werden und für die Ausstellung des Museums andere Museen und Kulturerlebnisse. Weiterhin wurden die Bezeichnungen der Typen geändert, sodass diese im Singular sind (Parkplatz statt Parkplätze) und möglichst eindeutig auf einen Ort gemünzt sind (Golfplatz statt Golfen).

Für die Empfehlung von Parkplätzen war es weiterhin im Interesse der Wissensbasis, die Kapazitäten der Parkplätze und Wohnmobilstellplätze mit in das Datenmodell aufzunehmen. Allerdings wurden diese in den Rohdaten oftmals nicht in einem gesonderten Feld hinterlegt, sondern als Fließtext in die Beschreibung der Parkplätze integriert. Um diese Informationen zu extrahieren, wurden reguläre Ausdrücke genutzt, die auf die Triggerworte „Kapazität“ und „Stellplätze“ reagieren.

Final wurde der Teil der Daten (ca. 1.000 POI), der aus dem Open Data Portal Schleswig-Holstein importiert wurde, händisch grob geprüft, um eine Einschätzung über die Qualität, der nach dem eingangs vorgestellten Baseline-Muster bereinigten Daten treffen zu können. Dabei hat sich herausgestellt, dass ca. 20% der Daten einer händischen Überarbeitung bedürfen. Die Gründe für den Überarbeitungsbedarf waren dabei sehr heterogen. Ein großer Teil der Bearbeitung wurde deshalb notwendig, weil eine fehlerhafte Zuordnung zum Typ POI vorlag. In Reaktion darauf mussten POI zum Beispiel gelöscht werden, wenn sie nicht als Erlebnis-POI gewertet werden konnten. Ein Beispiel dafür sind mehrere Physiotherapiepraxen, die unter dem Tag *Wellness Anwendungen* durch die vorherige Filterung in die Wissensbasis aufgenommen worden waren. Andere POI mussten einen anderen Tag zugewiesen bekommen. Dies war besonders eindeutig bei der Trennung von Badestellen im Inland und Stränden, die an dieser Stelle manuell eingeführt wurden. Aber auch bei dem beispielhaft ausgewählten POI „Hotel-Restaurant Rigoletto“, der in die Datenbasis als *eBike-Verleih* und *Outdoor* eingegangen ist, aber nicht als Restaurant. Ein weiterer großer Teil der Korrekturen bezog sich auf fehlerhafte Rohdaten. So mussten mehrere Testeinträge in den Daten gelöscht werden, Geolokationen und Parkplatzkapazität haben sich als fehlerhaft herausgestellt und Rechtschreibfehler in den Beschreibungen wurden korrigiert.

Bei dieser Art der händischen Evaluation der Datenqualität und der Baseline ist außerdem nicht klar, wie groß der Anteil der im Data Cleaning verlorenen POI ist. Die notwendigen Korrekturen wurden für die 1.000 POI aus dem Open Data Portal durchgeführt und die resultierende Datengrundlage dann mit weiteren POI-Daten aus der TA.SH Datenbank verschnitten. Diese wurden aufgrund ihrer geringeren

Datentiefe nicht in das Open Data Portal exportiert, eignen sich jedoch gut, um die POI-Dichte dieses PoC zu erhöhen. Dieser zweite Datensatz wurde nicht aufwendig bereinigt, um die dafür notwendige Zeit in die Konzeptionierung, Entwicklung und Evaluierung anderer Aspekte des Recommenders investieren zu können. Die PoC Datengrundlage ist dementsprechend als fehlerhaft zu betrachten, was in Kauf genommen wurde, um mehr POI im Recommender repräsentieren zu können und somit näher an realistischen Bedingungen zu arbeiten.

Nach Abschluss aller Data-Cleaning-Schritte umfasst die integrierte Wissensbasis in diesem PoC alle grundlegenden Informationen gemäß den Anforderungen. Darüber hinaus wurden zusätzliche Informationen hinzugefügt. Tabelle 20 bietet eine detaillierte Aufschlüsselung der Merkmale, die zur Beschreibung eines POI in diesem PoC verwendet werden. Des Weiteren wird angegeben, bei welchem Prozentsatz der vorhandenen POI Informationen zur "Dichte" verfügbar sind.

Feature	Beispiel	Dichte (in Prozent)
Unique ID	PoC-61-89797	100
Titel	Badestelle Elpersbüttel	100
Typ	Badestelle	100
Beschreibung	Die Badestelle Elpersbüttel ist ein ...	100
Geolokation	54.09 9.0755	100
Website	http://www.echt-dithmarschen.de/	76.02
Betreiber	Dithmarschen Tourismus e.V.	100
Ort	Meldorf	96.83
Postleitzahl	25704	98.80
Straße	Speicherkoog	88.31
Hausnummer	-----	42.21
Bildlinks	https://dam.destination.one/189535/5...	17.29

Tabelle 20: Beispiel von POI-Merkmalen anhand der Badestelle Elpersbüttel

Wie vorher beschrieben werden Parkplätze und Wohnmobilstellplätze außerdem auch durch das Attribut „Kapazität“ beschrieben. Dieses Attribut liegt für 38% der Parkplätze und Wohnmobilstellplätze vor.

5.2.4 Logik und Modelle

Dieser Abschnitt soll im Folgenden einen Einblick in die Logik geben, die hinter dem PoC-Alternativen-generierer steht. Dabei sollen sowohl theoretische als auch praktische Punkte aus der Entwicklung betont werden, um einen umfangreichen Überblick über die in LAB-TOUR SH geleistete Arbeit geben zu können. Grundsätzlich hat fast jedes Recommender-System zum Ziel zwei verschiedene Arten von Recommendations auszuspielen: "Mehr vom Gleichen" und "Innovationen". Es ist dabei oftmals deutlich schwieriger „Innovationen“ auszuspielen, gerade wenn keine Daten über das spezifische Verhalten

einzelner Nutzer*innen vorliegen. Diese beiden Arten von Empfehlungen zielen darauf ab, die Interessen und Intentionen einzelner Nutzer*innen zu antizipieren und Alternativen auszusprechen, die diesen entsprechen. Der „Mehr vom Gleichen“ Ansatz beschränkt sich hierbei darauf, sehr ähnliche Elemente zum Ausgangselement vorzuschlagen, bei denen somit gesichert ist, dass sie die Interessen eines/r Nutzer*in wahrscheinlich getroffen werden. Der innovative Ansatz versucht auch andersartige Elemente vorzuschlagen, wenn sie einem/r Nutzer*in gefallen könnten. Wenn zum Beispiel antizipiert wird, dass ein/e Nutzer*in an einem Strand wollte, um zu schwimmen, können auch Freibäder, Badestellen und Schwimmbäder vorgeschlagen werden, die dem Bedürfnis nach Aktivitäten an/im/auf dem Wasser nachkommen. Wollte ein/e Nutzer*in an den Strand, um dort spazieren zu gehen, können auch Wanderwege und Parks als Alternativen vorgeschlagen werden – nicht jedoch Schwimmbäder. Aus diesem Beispiel geht bereits hervor, warum dieser Ansatz schwieriger ist. Für den innovationsgetriebenen Ansatz ist die Kontextinformation einer Anfrage entscheidend. Diese können über diverse Möglichkeiten abgefragt werden. Mehr zu möglichen Personalisierungsansätzen wird deshalb in Abschnitt 5.2.6 dargestellt. Im LAB-TOUR SH Recommender werden beide dieser Ansätze verfolgt – wenn auch mit einigen Einschränkungen. Bei der Umsetzung von innovativen Empfehlungen wird hier versucht, die fehlenden Kontext-/Nutzerinformationen durch zwei Methoden auszugleichen. Einmal wird die Intention des/der Anwender*in grob approximiert, je nachdem für welchen POI-Typ er sich ursprünglich interessiert. Eine weitere Komponente des Versuchs, innovative Alternativen auszuspüren, beruht auf der zufälligen Auswahl von POI und auf der Möglichkeit für Destinationen, bestimmte POI auszuwählen, die in den Fokus des Alternativengenerierers aufgenommen werden. Dies können zum Beispiel Coldspots sein, bei denen ein großes Interesse besteht, sie zu vermarkten und an viele Nutzer*innen auszuspüren. In LAB-TOUR SH beruht der Recommender zum Hauptteil aus einem Knowledge Graphen, der alle für die Auswahl von alternativen POI notwendigen Informationen bereits vorhält, sodass diese nicht erst zur Laufzeit berechnet werden müssen. Der nächste Abschnitt wird diesen Knowledge Graphen ausführlicher beschreiben, bevor dann alle Einflüsse, die in die Alternativengenerierung eingehen, detailliert beschrieben werden.

Aufbau eines Knowledge-Graphen

Um alle für die Alternativengenerierung notwendigen Datenpunkte in einer gemeinsamen Datenstruktur zu vereinen, wurde in LAB-TOUR SH ein (*Meta-*) *Knowledge-Graph* aufgesetzt, der POI miteinander in Beziehung setzt. Dieser Graph besteht, um eine technische Grundlage für Graphstrukturen zu etablieren, aus sogenannten *Knoten* und *Kanten*. Dabei sind Knoten in einem Knowledge-Graph im Allgemeinen meist entweder Entitäten oder Konzepte – der LAB-TOUR SH Knowledge-Graph beschränkt sich allerdings in diesem PoC auf das Abbilden von Entitäten, in diesem Fall den individuellen POI, die hier die Knoten des Graphen bilden. Die Kanten eines Knowledge-Graphen beschreiben immer die Verknüpfung verschiedener Knotenpunkte. Oftmals werden damit die Verknüpfungen zwischen Entitäten und Konzepten beschrieben (ähnlich wie bei der Darstellung von Freundschaften in einem Social Media Netzwerk). Ein Beispiel im Kontext touristischer POI hierfür wären zwei Konzept-Knoten (*Indoor und Outdoor*) und eine Vielzahl an Entitäts-Knoten (*Wikingermuseum Haithabu, Schwimhalle Kiel, Schlei-Wanderweg, ...*). Die Kanten würden nun von den Entitäts-Knoten hin zu Konzept-Knoten laufen, die die jeweiligen Entitäten beschreiben (vgl. Abb. 64).

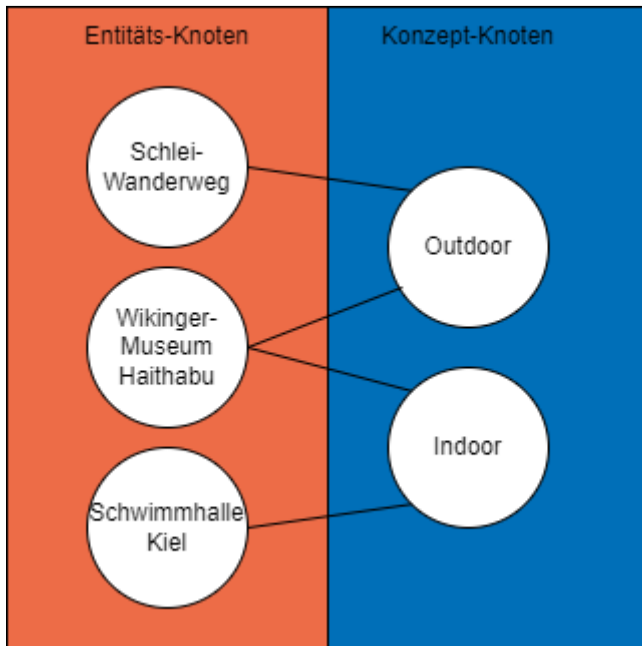


Abbildung 64: Beispiel für Konzept- und Entitätsknoten

Für eine Recommendation müsste aus einem solchen Graphen dann allerdings noch abstrahiert werden, welche POI füreinander als Alternative in Betracht gezogen werden können. Es interessiert also final vor allem, welche Beziehungen POI untereinander haben. Im LAB-TOUR SH Knowledge Graphen ist dieser Schritt direkt bei der Implementierung des Graphen durchgeführt worden, weshalb der Begriff *Meta-Knowledge-Graph* besser geeignet ist, um diesen Graphen zu beschreiben. Im Folgenden ist das Schema eines Graphen aufgezeigt, das sich an dem LAB-TOUR SH Meta-Knowledge-Graphen orientiert. Eine genauere Beschreibung aller im Graphen liegenden Kanten, also der Eigenschaften, die POI miteinander in Beziehung setzen und für die Auswahl von Alternativen genutzt werden, folgt im Anschluss.

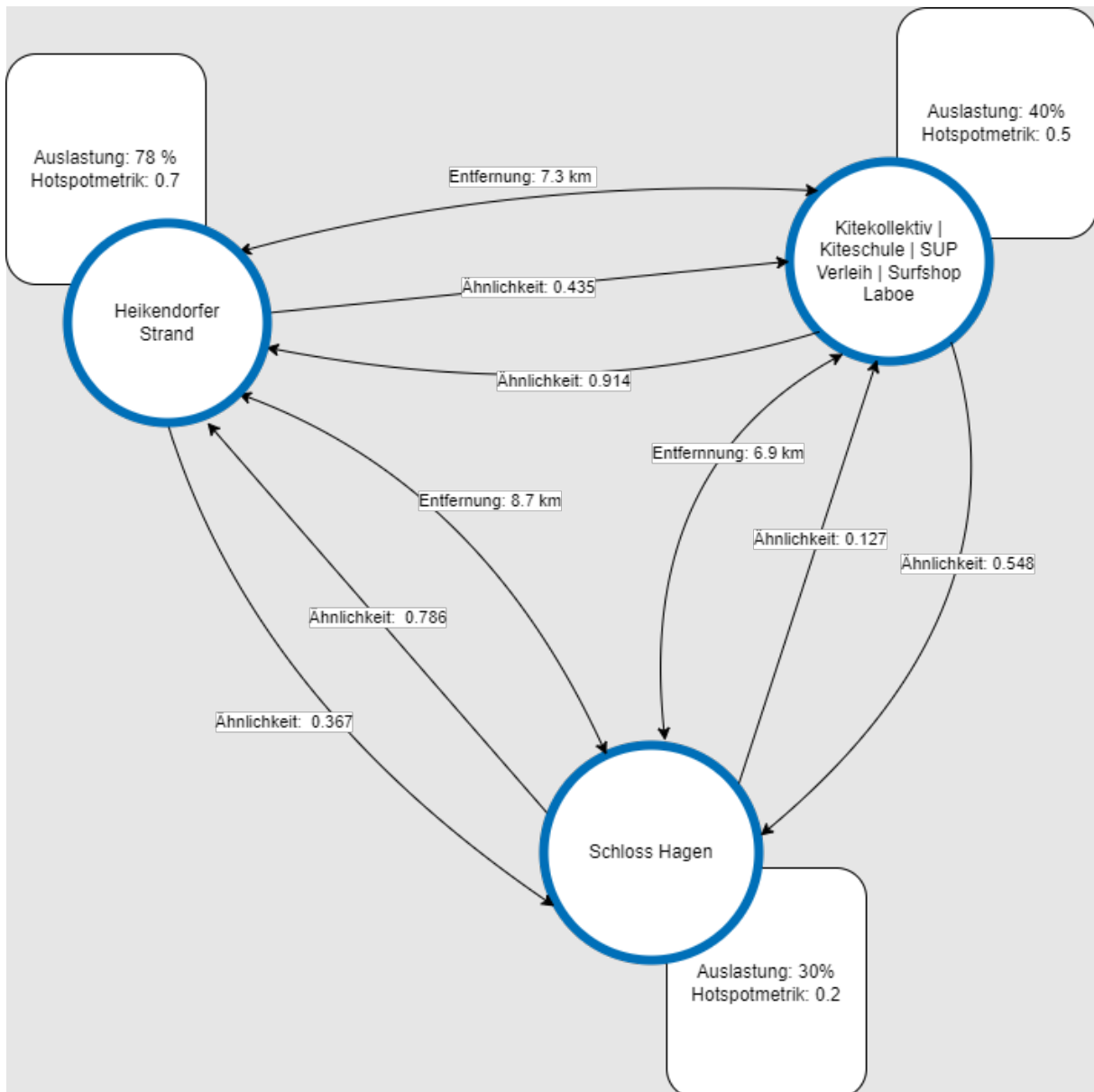


Abbildung 65: Meta-Knowledge-Graph in LAB-TOUR SH

Der vorliegende PoC Meta-Knowledge-Graph in Abbildung 65 wird fortlaufend aktualisiert und kann so immer die aktuellen Messungen und die aktuelle Prognose der Auslastung vorhalten, neue POI integrieren und die Hotspot-Metrik (siehe nachfolgender Abschnitt) aktualisieren. Die Frequenz dieser Aktualisierung sollte an die Frequenz der Sensordatenlieferung angelehnt sein, und war damit zuerst alle 15 min geplant. Mit der Einschränkung, dass die Datenlieferung alle 15 min nicht verlässlich gegeben ist, wurde die Aktualisierungsfrequenz in diesem PoC auf eine Stunde festgelegt.

Einflüsse auf die Alternativengenerierung

In die Ermittlung geeigneter Alternativen, die als Empfehlungen ausgespielt werden können, und in die Reihenfolge, in der sie sortiert werden, fließen mehrere Faktoren ein. Zudem ändert sich deren Relevanz, je nachdem für wie weit im Voraus ein Set an alternativen POI ermittelt werden soll. In diesem PoC gehen grundsätzlich folgende vier Faktoren in die Auswahl von Alternativen ein:

1. **Ähnlichkeit:** Die Ähnlichkeit des Ursprungs-POI zu einer potenziellen Alternative, in Bezug auf die Erfüllung der ursprünglichen Intention des/der Anwender*in.
2. **Entfernung:** Die Entfernung zwischen den beiden POI. Gemessen anhand der kürzesten, mit dem Auto befahrbaren Route, die diese beiden POI verbindet.
3. **Auslastung:** Die für diesen Tag prognostizierte oder gemessene maximale Auslastung des alternativen POI.
4. **Hotspot-Metrik:** Beinhaltet eine stark zusammengefasste Information über die vergangene Auslastung des alternativen POI. Diese bestimmt, ob ein POI ein sogenannter Hotspot oder Coldspot ist.

Auf diese Einflussfaktoren haben sich die Projektpartner*innen gemeinsam verständigt und weitere Faktoren wie das Wetter, die Jahreszeit und Öffnungszeiten vorerst nicht implementiert. Diese Simplifizierung dient der einfacheren Entwicklung der Kernalgorithmik dieses PoC-Alternativengenerierers, wobei die nicht implementierten Faktoren stets mitgedacht wurden. Zudem wurde sich bewusst dagegen entschieden, eine zeitliche Alternative für den Besuch des ursprünglich ausgewählten POI anzubieten. Zum einen würde dies einen User-Input über die ursprünglich intendierte Zeit des Besuches voraussetzen – und damit das Paradigma verletzen, für diesen PoC keinen User-Input zu verlangen. Andererseits liegt eine solche Empfehlung implizit schon vor, indem ein/e Nutzer*in sich über die prognostizierte Auslastung eines POI mit Auslastungs-Messung informieren und so selbst eine passende Besuchszeit auswählen kann. Um die Komplexität des PoC niedrig zu halten und Redundanzen zu vermeiden, werden also als Alternativen ausschließlich andere POI ausgespielt.

Um aus diesen Einflussparametern sinnvolle Recommendations zu generieren, wird ein Schema angewandt, das sich leicht unterscheidet, je nachdem für welchen Tag Recommendations gesucht werden. Dieses Schema wird im folgenden Abschnitt ausführlich behandelt – vorher werden die genannten Einflussfaktoren definiert.

Ähnlichkeiten von POI

Die Ähnlichkeit von POI fließt als entscheidender Faktor in die Auswahl von Alternativen für einen POI ein. Mit der Ähnlichkeit soll die Ursprungsintention des/der Anwender*in bestmöglich geschätzt werden. Daraus folgend sollen dann POI angeboten werden, die dieser Intention möglichst gerecht werden. Außerdem sollen POI angeboten werden, die einer Intention gerecht werden, die sehr häufig mit der Ursprungsintention einhergeht. Es handelt sich also um eine Ähnlichkeit der POI in Bezug auf die Intention oder die Erfüllung von Bedürfnissen der Nutzer*innen und ist daher stark subjektiv. Sollte es deshalb, anders als in diesem PoC, möglich sein, die Intention der Nutzer*innen besser einschätzen zu können wird dies stark empfohlen. Dies könnte durch User-Profile oder Eingaben im Frontend erreicht werden (mehr zu möglichen Personalisierungsansätzen in Abschnitt 5.2.6). Trotz der inhärenten Subjektivität lassen sich die Ähnlichkeiten von POI grundsätzlich in zwei Domänen herunterbrechen. Diese sind einerseits die Ähnlichkeit von POI-Typen (*Museum, Strand, ...*) untereinander und andererseits die Ähnlichkeit spezifischer POI (*Badestelle Plöner See, Wikingermuseum Haithabu*) untereinander. Die

Ähnlichkeit spezifischer POI besteht dabei immer auch mindestens zu einem Teil aus der Ähnlichkeit der beiden POI-Typen. Als Beispiel für die Formulierung der Ähnlichkeit von POI-Typen, könnte folgende Frage gestellt werden: Was erfüllt die Bedürfnisse einer Person, die ein Schwimmbad besuchen möchte, wahrscheinlich besser? Der Besuch einer Sauna, eines Strandes, eines Freibades, eines SUP-Verleihs oder eines Aussichtspunktes? Neben dieser Frage ist auch Reihenfolge und Quantifizierung der Ähnlichkeiten dieser Kategorie für die Alternativengenerierung von Belang. Die Ähnlichkeit der POI-Typen muss dementsprechend auf einer Intervallskala eingeordnet werden können. Um die Ähnlichkeiten spezifischer POI-Typen zu beschreiben, muss die Wissensbasis über eine ausreichende Datentiefe in den POI verfügen. Eine Beispielfrage, um Ähnlichkeiten dieses Typs zu verdeutlichen wäre: Was erfüllt die Bedürfnisse einer Person, die das Meerwasser-Wellenbad (mit angeschlossener Sauna) in Eckernförde besuchen wollte, wahrscheinlich besser? Der Besuch einer nahegelegenen Sauna mit angeschlossenem Fitnessstudio, einer anderen, etwas weiter gelegenen Sauna mit kleinem Schwimmbecken, der Besuch des Eckernförder Strandes mit Volleyballfeldern und DLRG-Bewachung oder der Besuch eines Cafés mit veganem Kuchenangebot. Wieder spielen auch hier die Reihenfolge und Quantifizierung eine essenzielle Rolle für die nachfolgenden Algorithmen. Aus dieser Verbalisierung wird bereits deutlich, dass die Ähnlichkeit des zweiten Typs deutlich komplizierter ist, da sie noch subjektiver ist. Im vorangegangenen Beispiel wäre die nahegelegene Sauna wahrscheinlich die beste Alternative für eine Person, die explizit wegen der Sauna in das Meerwasser-Wellenbad wollte, während der Strand oder die weiter entfernt liegende Sauna eine gute Alternative für eine Person wäre, die wegen des Schwimmens an sich in die Schwimmhalle wollte – aber auch hier wäre die Frage, ob es wichtig war, dass Wellen und Salzwasser im Ursprungs-POI vorhanden waren. Die dadurch entstehende Ähnlichkeit ist userspezifischer, anfälliger für Datenlücken und deutlich schwieriger zu implementieren. In der Literatur existieren zwar Ansätze, um die Ähnlichkeit zwischen POI zu bestimmen, allerdings basieren diese fast ausschließlich auf der Grundlage von Nutzerdaten (Zeng et al., 2017). In LAB-TOUR SH ist es allerdings gelungen, Ähnlichkeiten des ersten Typs mit sehr zufriedenstellenden Ergebnissen ohne Zuhilfenahme von Nutzerdaten zu erzeugen. Das dafür in diesem Forschungsprojekt entwickelte Verfahren soll nun in den Grundzügen vorgestellt werden.

Der Grundgedanke hinter dem hier verwendeten Ansatz ist es, POI-Typen als Vektoren darzustellen, wobei jede Dimension dieser Vektoren eine touristisch wichtige Eigenschaft beschreibt. Wenn es gelingt, diese Dimensionen richtig zu wählen, kann ein POI-Typ also sehr umfassend und facettenreich abgebildet werden. Eine gewählte Dimension ist dabei zum Beispiel die Eignung eines POI für Kinder, eine andere die Naturnähe des POI. Insgesamt wurden in diesem PoC 27 Dimensionen ausgewählt, um POI abzubilden (siehe Anhang). Die so generierten Vektoren werden dann darüber ineinander in Beziehung gesetzt, wie sie zueinander im Vektorraum stehen. Genauer gesagt wird dafür der Cosinus des Winkels zwischen zwei Vektoren genutzt. Diese sogenannte Cosinus-Similarity ist ein etabliertes Ähnlichkeitsmaß, das vor allem im Bereich des Natural Language Processings genutzt wird, in dem Wörter durch hochdimensionale Vektoren repräsentiert werden – also einem durchaus sehr ähnlichen Anwendungsfall. Dieses Verfahren sorgt dafür, dass die Ähnlichkeit zweier identischer Vektoren 1 ist und die Ähnlichkeit zwischen zwei orthogonal (senkrecht) stehenden Vektoren 0 ist. Daraus folgt, dass die Ähnlichkeit eines POI-Typs zu sich selbst immer 1 ist.

Die in diesem PoC ausgewählten Dimensionen beruhen zu einem großen Teil auf Daten des GfK DestinationMonitor Deutschland (GfK SE, 2023). Der Destinationsmonitor ist einer Verbraucherpanelbefragung, die im monatlichen Abstand das Reiseverhalten der deutschsprachigen Wohnbevölkerung ab einer Distanz von 50 km zum Zielort der Reise repräsentativ erhebt. Um Verzerrungen durch die Corona-Pandemie zu umgehen, konnten Daten für mehrtägige Urlaubs- und Freizeitreisen der

deutschsprachigen Bevölkerung nach Schleswig-Holstein im Jahr 2019 analysiert werden. Zur Verfügung stand eine Kreuztabelle der Aktivitätsarten bei Aktivitäten von privatreisenden Gästen, die während ihrer Reise nach Schleswig-Holstein eine wichtige Rolle gespielt haben. Dabei handelt es sich um 17 Aktivitäten wie „Radfahren“, „Wandern“, „Aktivitäten im/am/auf dem Wasser“ oder „Besuch von kulturellen/historischen Sehenswürdigkeiten“. So lassen sich Aussagen über die Verknüpfungen von Aktivitäten bei einer Schleswig-Holstein-Reise tätigen (Urlaubs- und Freizeitreisende, die shoppen gegangen sind haben auch ein Restaurant besucht usw.) (siehe Tab. 21).

	Besuch von kulturellen / historischen Sehenswürdigkeiten	Besuch von Museen / Ausstellungen	Besuch von Events / Veranstaltungen (z.B. Konzerte)	Besuch von Erlebnisrichtungen (z.B. Freizeitparks)	Zeit mit der Familie verbringen	Aufenthalt in der Natur	Spazieren gehen	Einkaufen / Shopping	Typische Speisen / Getränke genießen	Nachleben genießen (z.B. Disco Bar)	Wellnessangebote nutzen	Aktivitäten am / im / auf dem Wasser (inkl. Segeln, Surfen)	Wandern	Radfahren	Wintersportaktivitäten	Sonstiger Aktivport (z.B. Klettern, Mountainbike)	Sonstige Aktivitäten
Besuch von kulturellen / historischen Sehenswürdigkeiten	100	84	30	57	45	58	51	63	61	23	50	57	67	71	100	40	4
Besuch von Museen / Ausstellungen	18	100	8	14	11	12	11	14	16	3	11	14	11	16	100	9	0
Besuch von Events / Veranstaltungen (z.B. Konzerte, Sporte)	22	26	100	20	29	21	25	31	26	48	22	22	20	20	0	46	0
Besuch von Erlebnisrichtungen (z.B. Freizeitparks)	20	23	10	100	21	19	18	25	22	9	29	30	17	18	0	41	0
Zeit mit der Familie verbringen	48	56	43	65	100	46	48	53	48	57	45	56	45	49	0	48	8
Aufenthalt in der Natur	85	86	42	80	64	100	79	77	81	43	81	82	92	89	100	81	0
Spazieren gehen	55	55	37	55	49	58	100	69	59	51	48	51	69	48	100	50	3
Einkaufen / Shopping	24	26	16	27	19	20	25	100	32	40	30	22	19	22	0	24	0
Typische Speisen / Getränke genießen	47	58	28	48	35	42	42	63	100	70	46	44	51	48	100	36	3
Nachleben genießen (z.B. Disco Bar)	1	0	2	1	2	1	2	3	3	100	3	1	1	0	0	1	0
Wellnessangebote nutzen	5	5	3	8	4	5	4	8	6	10	100	5	6	4	0	9	0
Aktivitäten am / im / auf dem Wasser (inkl. Segeln, Surfen)	37	44	20	57	35	36	31	38	37	21	34	100	35	54	100	43	0
Wandern	32	24	14	24	20	30	31	23	32	16	28	25	100	28	100	19	0
Radfahren	23	24	9	17	15	20	15	19	20	3	12	27	19	100	100	14	0
Wintersportaktivitäten	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	1	0
Sonstiger Aktivport (z.B. Klettern, Mountainbike)	5	5	8	15	6	7	6	6	6	4	12	8	5	5	100	100	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100

Tabelle 21: Kombination von Aktivitäten bei Urlaubs- und Freizeitreisen ab 1 Übernachtung in Schleswig-Holstein im Jahr 2019.

Quelle: GfK SE, 2023

Diese Aktivitäten wurden als Eigenschaften in die Vektorrepräsentation der POI-Typen übernommen und von der Fachhochschule Westküste auf einer Skala von 0-10 für jeden POI-Typ bewertet. Alle weiteren Dimensionen der Vektoren wurden so ausgewählt, dass sie die Aktivitäten der Erhebung zu einem möglichst kompletten Set an touristisch relevanten Eigenschaften ergänzen. Dafür wurden Eigenschaften eingeführt wie „Wie gut kann man dort essen?“, „Kann man dort schwimmen?“, „Wie natürlich ist es dort?“ und ebenfalls auf einer Skala von 0-10 bewertet.

Ein weiterer wichtiger Schritt in der Bestimmung der Ähnlichkeiten von POI ist die Kreuzkorrelation der Aktivitäten aus der Marktforschung. Dadurch soll es möglich werden, im Alternativengenerierer besser auch „innovative“ Empfehlungen zu machen. Realisiert wird dies durch das Matrix-Vektor Produkt zwischen der Umfrage (als Matrix) und der Einordnung des Potentials des POI in Bezug auf jede der Aktivitäten in der Erhebung (als Vektor). Der dadurch entstehende Vektor hat die gleiche Länge wie der ursprüngliche Eigenschafts-Vektor der POI, schätzt aber nun viel mehr ein, wie „zufrieden“ jemand an diesem POI wäre, der sich ursprünglich für die entsprechende Aktivität interessiert hat. Damit wird nun berücksichtigt, dass jemand, der ursprünglich an der Aktivität „Besuch von kulturellen/historischen Sehenswürdigkeiten“ interessiert war, auch mit einem Wanderweg zufrieden wäre. Dort kann zwar primär keine Sehenswürdigkeiten besichtigt werden, aber 84% der Besucher*innen, die sich für den Besuch von Sehenswürdigkeiten interessieren, interessieren sich auch für den „Aufenthalt in der Natur“, zu dem der Wanderweg großes Potential hat.

In die Formulierung der Vektoren geht darüber hinaus auch eine Gewichtung mit ein. Diese Gewichtung ist abhängig vom Ausgangs-POI für eine Empfehlung und versucht damit abzubilden, welche touristischen Eigenschaften einem/r Nutzer*in wahrscheinlich wichtig sind. So suchen Nutzende beispielsweise eine Alternative für einen Wanderweg. Ein Wanderweg bietet von seinen Eigenschaften weder die Möglichkeit schwimmen zu gehen noch etwas zu essen, ist aber sehr naturnah. Entsprechend werden während der Gewichtung diese Gewichte (z. B. Schwimmen – 0; Essen – 0; Naturnähe – 9) auf die Eigenschafts-Vektoren aller POI-Typen multipliziert. Nach dieser Gewichtung hat sich also aus dem ursprünglich neutral beschriebenen POI eine Repräsentation ergeben, die auch die antizipierte Intention

des Nutzers mit abbildet. Durch diese Gewichtung, je nach Ausgangs-POI, wird die hier eingeführte Ähnlichkeit zu einer gerichteten Eigenschaft, d. h. die Ähnlichkeit eines Wanderwegs zu einem Strand, kann durchaus anders sein als die Ähnlichkeit eines Strandes zu einem Wanderweg.

Bei der Durchführung dieses Verfahrens hat sich die Beobachtungen ergeben, dass die berechneten Ähnlichkeiten alle recht nahe an 1 liegen und das manche POI-Typen nach dieser Berechnung nicht viele sehr ähnliche andere POI-Typen aufweisen. Um aber die gleichen Algorithmen für alle POI nutzen zu können und zur Erhöhung der Nachvollziehbarkeit dieser Algorithmen, wurden die Ähnlichkeiten pro Ausgangs-POI so skaliert, dass der ähnlichste, aber nicht gleiche, POI-Typ eine Ähnlichkeit von 0.99 hat und der unähnlichste eine Ähnlichkeit von 0. Die Ähnlichkeitsverhältnisse der POI dazwischen wurden beibehalten. Umgesetzt worden ist dies über eine einfache Min-Max Skalierung, in der das Maximum mit 0.99 multipliziert wurde. Dies führt dazu, dass manche POI-Typen sehr viele andere POI-Typen aufweisen, die eine Ähnlichkeit größer 0.9 haben. Andere POI-Typen weisen jedoch und nur den ähnlichsten POI in diesem Intervall auf, der dort per Definition liegen muss. Aus diesem Verfahren ergibt sich eine Matrix, die alle POI-Typen mit einer Ähnlichkeit zueinander beschreibt. Um die Ähnlichkeit der einzelnen, spezifischen POI in den Meta-Knowledge-Graphen einpflegen zu können, wurde jeweils der den beiden POI-Typen entsprechende Wert aus dieser Matrix genutzt.

Folgendes fiktives Beispiel verdeutlicht das Verfahren der Ähnlichkeitsberechnung: In diesem Beispiel wird es darum gehen, die Ähnlichkeit zwischen einem Strand und einem Wanderweg, bzw. einem Wanderweg und einem Strand zu bestimmen. Dafür müssen einige fiktive Annahmen getroffen werden, um die gleiche Ausgangslage zu schaffen, die auch im LAB-TOUR SH PoC geschaffen wurde. In den folgenden Tabellen (die oftmals Vektoren repräsentieren) wurden Farben verwendet, um die Nachvollziehbarkeit bestimmter Werte in den nachfolgenden Rechnungen zu erhöhen. Folgende Prämissen wurden in diesem vereinfachten Beispiel getroffen:

1. Maßgebliche Eigenschaften, die für die Klassifizierung von Tourist*innen relevant sind, sind das Bedürfnis nach „**Spazieren**“ und das Bedürfnis nach „**Schwimmen**“.
2. Es existiert eine repräsentative Erhebung, die festgestellt hat, wie das Interesse an „Schwimmen“ von Menschen ist, die an „Spazieren“ interessiert sind und wie das Interesse an „Spazieren“ von Menschen ist, die an „Schwimmen“ interessiert sind. Diese Umfrage hat ergeben, dass Menschen, die als Grundinteresse „Schwimmen“ angegeben haben, zu 50% auch gerne spazieren, während von den Menschen die als Grundinteresse „Spazieren“ angegeben haben, nur 20% auch an Schwimmen interessiert sind. Erklären lässt sich das in diesem Beispiel daran, dass insgesamt deutlich mehr Menschen gerne spazieren gehen, von denen manche nicht so sportlich veranlagt sind. Von den Leuten, die schwimmen hat ein größerer Teil generell Lust auf Bewegung und deshalb auch Interesse am Spazieren. Dieses Kreuzinteresse lässt sich (spaltenweise) festhalten, indem die prozentualen Werte auf eine Skala zwischen 0 und 1 gebracht wurden.

	Grundinteresse von Personen an der Aktivität	
	<i>Spazieren</i>	<i>Schwimmen</i>
Spazieren	1	0.5
Schwimmen	0.2	1

3. Es wurde von Experten festgestellt, wie sehr die POI-Typen „Strand“ und „Wanderweg“ die bereits etablierten touristischen Grundinteressen „Spazieren“ und „Schwimmen“ ermöglichen. Dies wurde auf einer Skala von 0 (keine Eignung) bis 10 (perfekte Eignung) umgesetzt.

	Grundinteresse (Eignung)	
	<i>Spazieren</i>	<i>Schwimmen</i>
Wanderweg	10	0
Strand	4	8

4. Weiterhin lässt sich ausmachen, dass die Interessen von Tourist*innen auch die Kategorien „**In der Sonne liegen**“ und „**Kinderfreundlichkeit**“ relevant sind. Für diese Interessen liegt allerdings keine Umfrage vor, die sie miteinander in Beziehung setzt. Nachfolgend werden ein Wanderweg und ein Strand in den Dimensionen dieser beiden Eigenschaften beschrieben:

	Weitere Interessen (Eignung)	
	<i>In der Sonne sein</i>	<i>Kindereignung</i>
Wanderweg	5	6
Strand	10	9

5. Mit den Eigenschaften „**Spazieren (Eignung)**“, „**Schwimmen (Eignung)**“, „**In der Sonne sein (Eignung)**“ und „**Kindereignung**“ lassen sich POI vollumfänglich beschreiben.

Mit der Aufstellung dieser Grundbedingungen lässt sich nun das im vorherigen Absatz eingeführte Verfahren anwenden, um die Ähnlichkeit eines Wanderwegs und eines Strandes bidirektional zu bestimmen. In einem ersten Schritt wird dafür die Grundinteressen-Korrelationsmatrix in einer Matrix-Vektor Multiplikation mit der Interesseneignung der jeweiligen POI-Vektoren durchgeführt. Dabei entstehen Vektoren, die beschreiben, wie geeignet ein POI für die Erfüllung der verschiedenen Grundinteressen ist. Dieses Matrix-Vektor-Produkt ist nachfolgend visualisiert:

	Grundinteresse (Erfüllung)	
	Spazieren	Schwimmen
Wanderweg	$1 \cdot 10 + 0.2 \cdot 0$	$0.5 \cdot 10 + 1 \cdot 0$
Strand	$1 \cdot 4 + 0.2 \cdot 8$	$0.5 \cdot 4 + 1 \cdot 8$

Die Terme wurden anschließend zusammengefasst und zur besseren Lesbarkeit so dargestellt:

	Grundinteresse (Erfüllung)	
	Spazieren	Schwimmen
Wanderweg	10	5
Strand	5.6	10

Hierbei lässt sich bereits anmerken, dass die Skala der Vektoren sich nicht über verschiedene POI-Typen hinweg vergleichen lässt. In der maximalen Ausprägung kann die Erfüllung eines POI in Bezug auf eines der Grundinteressen maximal der Summe aller seiner Eignung für eben diese Grundinteressen entsprechen. Die Erfüllung des POI-Typs Strand ist in diesem Beispiel also als „Punkte von 12“ zu verstehen, die des Wanderwegs als „Punkte von 10“. Im nächsten Schritt wird die Direktionalität in die Eigenschaft der „Ähnlichkeit“ eingebracht, indem eine Skalierung angelegt wird, die durch die Grundinteresseneignung des Ausgangs-POI bestimmt wird:

	Grundinteresse (Erfüllung, skaliert über Strand)	
	Spazieren	Schwimmen
Wanderweg	$4 \cdot 10$	$8 \cdot 5$
Strand	$4 \cdot 5.6$	$8 \cdot 10$

Wieder wurden die Terme zusammengefasst:

	Grundinteresse (Erfüllung, skaliert über Strand)	
	Spazieren	Schwimmen
Wanderweg	40	40
Strand	22.4	80

Wenn das gleiche Verfahren in die andere Richtung – ausgehend vom Wanderweg – angewendet wird, ergeben sich die folgenden Werte:

	Grundinteresse (Erfüllung, skaliert über Wanderweg)	
	<i>Spazieren</i>	<i>Schwimmen</i>
Wanderweg	100	0
Strand	56	0

Der Grundgedanke hinter dieser Skalierung ist, dass ein/e Nutzer*in, der/die ursprünglich Interesse an einem Strand gezeigt hat, auch an den Dingen interessiert ist, die an einem Strand erlebt und gemacht werden können. Andersherum spielt, ausgehend vom Wanderweg, zum Beispiel die ursprüngliche Schwimmeignung einer Alternative keine Rolle – der/die Nutzer*in war wahrscheinlich nicht daran interessiert schwimmen zu gehen.

Nach dieser Skalierung wird nun ein Vektor hinzugefügt, der die „weiteren Interessen“, in diesem Beispiel „In der Sonne liegen“ und „Kindereignung“ beschreibt. Um ungefähr den gleichen Einfluss jeder Dimension auf die später zu bestimmende Cosinus-Ähnlichkeit zu gewährleisten, wird auch der hinzugefügte Vektor skaliert. Diese Skalierung funktioniert so, dass der Betrag des angeflanschten Vektors dem Betrag des Originalvektors im Verhältnis der Dimensionen der beiden Vektoren entsprechen soll. In diesem Beispiel haben beide Vektoren eine Dimensionalität von zwei. Dementsprechend wird der anzufügende Vektor so skaliert, dass sein Betrag dem bisher erarbeiteten Vektor entspricht.

Ausgehend von der Skalierung über den Strand als ursprünglich ausgewählten POI ergibt sich

6. für den Strand ein Vektorbetrag von:

$$\sqrt{22.4^2 + 80^2} \approx 83.1$$

7. für den Wanderweg ein Vektorbetrag von:

$$\sqrt{40^2 + 40^2} \approx 56.6$$

Daraus resultieren die folgenden Skalierungsfaktoren:

8. für den Strand:

$$\sqrt{10^2 + 9^2} x = 83.1 \Rightarrow x \approx 6.2$$

9. für den Wanderweg:

$$\sqrt{5^2 + 6^2} x = 56.6 \Rightarrow x \approx 7.2$$

Damit ergeben sich final, ausgehend vom Strand als ursprünglich ausgewähltem POI, die folgenden Vektoren, die die POI repräsentieren:

	Grundinteresse		Weitere Interessen	
	skaliert über den Strand			
	<i>Spazieren</i>	<i>Schwimmen</i>	<i>In der Sonne sein</i>	<i>Kindereignung</i>
Wanderweg	40	40	5*7.2	6*7.2
Strand	22.4	80	10*6.2	9*6.2

Zusammengefasst also:

	Grundinteresse		Weitere Interessen	
	skaliert über den Strand			
	<i>Spazieren</i>	<i>Schwimmen</i>	<i>In der Sonne sein</i>	<i>Kindereignung</i>
Wanderweg	40	40	36	43.2
Strand	22.4	80	62	55.8

An diesem Punkt wird bereits klar, dass sich diese Vektoren nicht mehr ganz einfach nachvollziehen lassen. Das liegt zum einen daran, dass die Verrechnung der in der Umfrage erfragten Kreuzinteressen die Dimensionen weniger klar benennbar macht. Zum anderen liegt es aber auch daran, dass das Skalierungsverfahren darauf ausgelegt ist, einen ungefähr gleichen Einfluss jeder Dimension auf die Cosinus-Ähnlichkeit zu erreichen. Damit verzerren sich die absoluten Ausprägungen der Features etwas, an denen sich der Mensch allerdings oftmals zu orientieren versucht, wenn solche Vektoren vorkommen.

In einem letzten Schritt lässt sich zwischen diesen beiden Vektoren nun die Cosinus-Ähnlichkeit berechnen, welche definiert ist als das Skalarprodukt der Vektoren geteilt durch das Produkt ihrer Beträge. In der aufgestellten Formel bedeutet die Schreibweise $S \rightarrow W$, dass eine Ähnlichkeit mit dem Strand als Ausgangs-POI berechnet wurde.

$$\text{Ähnlichkeit}_{S \rightarrow W} = \frac{40 * 22.4 + 40 * 80 + 36 * 62 + 43.2 * 55.8}{\sqrt{40^2 + 40^2 + 36^2 + 43.2^2} \sqrt{22.4^2 + 80^2 + 62^2 + 55.8^2}} = 0.93$$

Sollte die Skalierung nicht über den Strand als Ausgangs-POI durchgeführt worden sein, sondern über den Wanderweg als Ausgangs-POI, würde sich folgende Cosinus-Ähnlichkeit ergeben.

$$\text{Ähnlichkeit}_{W \rightarrow S} = \frac{100 * 56 + 0 * 0 + 64 * 41 + 76.8 * 36.9}{\sqrt{100^2 + 0^2 + 64^2 + 76.8^2} \sqrt{56^2 + 0^2 + 41^2 + 36.9^2}} = 0.99$$

Dies verdeutlicht den unterschiedlichen Grad der Ähnlichkeit, die sich je nach Betrachtung der POI-Typen ergeben. Die Berechnungen zeigen, dass ein Strand sehr ähnlich zu einem Wanderweg ist (0.99), während ein Wanderweg nicht ganz so ähnlich zu einem Strand ist (0.93). Dies ist plausibel, da für Personen, die eine Alternative zu einem Wanderweg suchen, nur wichtig ist, ob sie spazieren gehen, die Sonne genießen und ihre Kinder mitnehmen können. Ob sie an der Alternative schwimmen können, spielt keine Rolle. Auf der anderen Seite ist dies für Personen, die vom Strand als Ausgangs-POI kommen, wichtig. Für sie sind alle vier betrachteten Features der POI relevant. Da an einem Wanderweg jedoch nicht geschwommen werden kann, ist die Ähnlichkeit entsprechend geringer. Die Cosinus-Ähnlichkeit hat im Vergleich zur inversen Euklidischen Distanz die Eigenschaft, dass sie die Ausprägung verschiedener Features nicht berücksichtigt, sondern nur deren Verhältnis untereinander.

	Grundinteresse		Weitere Interessen	
	<i>Spazieren</i>	<i>Schwimmen</i>	<i>In der Sonne sein</i>	<i>Kindereignung</i>
POI-1	100	50	80	20
POI-2	10	5	8	2

Die beiden fiktionalen POI sind nach der Cosinus-Ähnlichkeit also völlig identisch, obwohl POI-1 offensichtlich deutlich stärker in der Erfüllung mancher Eigenschaften ist und damit auch als „besser“ verstanden werden könnte. In diesem PoC wurden verschiedene Distanzmaße ausprobiert (Cosinus-Ähnlichkeit, inverse Euklidische Distanz, Pearson-Korrelation), die allesamt gute Ergebnisse produziert haben, allerdings aufgrund des Mangels eines geeigneten quantitativen Evaluationsverfahren nicht auf einer Ordinalskala bewertet werden konnten. Die Entscheidung für die Cosinus-Similarity als Distanzmaß in diesem PoC wurde hauptsächlich gefällt, da die Vektorrepräsentation der POI, wie eingangs erwähnt, nicht auf einer einheitlichen Skala liegen. Dementsprechend würde potenziell ein Fehler eingeführt werden, wenn die inverse Euklidische Distanz als Ähnlichkeitsmaß herangezogen würde.

An diesem Beispiel werden jedoch auch einige Nachteile dieses Verfahrens deutlich. Es gibt noch keine wissenschaftliche Grundlage, auf der Interessen in "Grundinteressen" und "Weitere Interessen" aufgeteilt werden können. Aus Gründen der Einfachheit stützte sich dieser PoC daher auf die vorhandene Marktforschung der GfK zu den Aktivitäten bei Urlaubs- und Freizeitreisen in Schleswig-Holstein. Es bleibt weiterhin wichtig, dass die berücksichtigten Dimensionen einen POI umfassend beschreiben können. Die Validierung dieser Methode ist eine schwierige Aufgabe, die aufgrund der stark subjektiven Einschätzung von Ähnlichkeit kaum quantitativ zu bewerkstelligen ist. Der Ansatz, der in diesem PoC zur Validierung gewählt wurde, besteht aus zwei Komponenten. Zum einen wurden die fünf ähnlichsten POI-Typen jedes POI-Typs per Durchsicht auf Plausibilität geprüft. Weiterhin wurden mit den aus dem Verfahren resultierenden Ähnlichkeitsvektoren der jeweiligen POI-Typen eine hierarchische Cluster-Analyse (Ward, Euklidische Distanz) durchgeführt. Die entstandenen fünf Cluster wurden auf inhaltliche Zusammengehörigkeit überprüft. Da beide Methoden gute Ergebnisse geliefert haben, wurde die vorgestellte Methode als plausibel durch qualitative Überprüfung angenommen. Tabelle 22 zeigt die fünf ähnlichsten POI-Typen für eine Auswahl von Ursprungs-POI-Typen.

	Alternative POI-Typen (sortiert nach Ähnlichkeit)				
Ausgangs-POI	1.	2.	3.	4.	5.
Badestelle	Strand	Freibad	SUP-Verleih	Tretbootverleih	Surfen / Kiten
Freizeitpark	Minigolf	Zoo	Therme	Erlebnishof	Science-Center
Strand	Badestelle	Strandbar	Freibad	SUP-Verleih	Tretbootverleih
Radweg	Fahrradverleih	Reiterhof	Kanuverleih	Wanderweg	Outdoor
Naturschutzgebiet	Naturerlebnispark	Vogelkunde	Aussichtspunkt	Outdoor	Wanderweg
Kirche	Diverse Kunst & Kultur	Galerie	Museum	Wellness-Anwendungen	Theater
...

Tabelle 22: Alternative POI ausgehend von Ursprungs-POI

Beide Validierungsmethoden haben allerdings auch bei anderen Ansätzen, wie zum Beispiel der Berechnung der Vektorähnlichkeiten auf Grundlage der inversen Euklidischen Distanz, oder unterschiedlichen Versionen der Vektoren Eigenschafts-Vektoren zufriedenstellende Ergebnisse geliefert. In der dargestellten Tabelle wurde darauf verzichtet den gleichen POI-Typ des Ausgangs-POI noch einmal aufzuführen, obwohl dieser per Definition immer der Ähnlichste ist.

Die inhärente Subjektivität dieser Bewertungsmethode kann in diesem Beispiel gut nachvollzogen werden, gerade anhand der Alternativen für den Radweg. Je nachdem, ob davon ausgegangen wird, dass jemand mit seinem eigenen Fahrrad anreisen will, ist ein Fahrradverleih eine gute oder schlechte Alternative. Außerdem ist der Reiterhof wohl sehr wahrscheinlich keine praktikable Alternative, obwohl beide Aktivitäten sportlich und in der Natur sind. Dies macht den Nachteil deutlich, dass die Eigenschaften, die einen POI repräsentieren mit dieser Methode so umfänglich gewählt werden müssen, dass sie alle Sonderfälle in den Daten beschreiben. Mit den in diesem PoC ausgewählten Dimensionen scheint die Trennung zwischen Fahrradfahren und Reiten nicht klar genug gelungen zu sein.

Entfernung und Auslastung

Als Maß für die Distanz zwischen zwei POI ist die Entfernung genutzt worden, die laut der Open-Street-Map Routenplanung zwischen den POI liegt. Dabei wurden auch die potenziell notwendigen Fußwege von einem POI zur nächstgelegenen Straße berücksichtigt. Gerade in Küsten- und Seeregionen unterscheidet sich die so berechnete Entfernung maßgeblich von der Luftlinie zwischen zwei POI. Im Durchschnitt sind die über Open-Street-Map abgefragten Distanzen ca. das eineinhalbfache der Luftlinie. Aus diesem Grund wurde dieser Wert als Default festgelegt, falls die Berechnung der Route über Open-Street-Map fehlschlagen sollte.

Um die Auslastung des alternativen POI in der Algorithmik des Alternativengenerierers angemessen darzustellen, werden die verfügbaren Daten vergangener Messungen mit den Prognosen für den aktuellen Tag kombiniert. Dadurch entsteht eine umfassende Repräsentation für den gesamten Tag. In diesem PoC liegen diese Datenpunkte in stündlicher Auflösung vor – also 24 Auslastungswerte. Als Auslastung wird in der Algorithmik nun der maximal erreichte Auslastungswert eines Tages verstanden, da keine spezifischere Information existiert, für welche Uhrzeit sich ein/e Nutzer*in interessiert. Dies bleibt zu unterscheiden von der Auslastung, die im Frontend für einen POI angezeigt wird. Diese ist immer live und bezieht sich auf die aktuelle Uhrzeit. In der Algorithmik hat es sich allerdings als sinnvoll herausgestellt nicht mit den Live-Werten zu arbeiten, sondern eben gerade mit den Tagesmaxima der Auslastung. Dies hat den Hintergrund, dass erwartet wird, dass Nutzer*innen den Recommender einige Stunden vor der tatsächlichen Aktivität im Rahmen der Ausflugsplanung nutzen werden. Dort sollen dann keine POI als Alternativen angezeigt werden, die zwar zum Zeitpunkt der Planung wenig ausgelastet sind, aber im Laufe des Tages noch sehr voll werden. Eine Auslastung ist minimal 0 und hat kein Ende nach oben, wobei eine Auslastung von 100% als 1 beschrieben wird. Zu diesem Zeitpunkt hat ein POI dann eine Auslastung erreicht, die der Kapazität entspricht, die von einer Tourismusorganisation festgelegt wurde. Das Datenmodell und die Algorithmik sehen aber durchaus vor, dass ein POI auch über 100% ausgelastet sein kann. In diesem Fall kommt es zu einer eindeutigen Überlastung. Sollte für einen POI keine Auslastung vorliegen, weil keine Messtation in der Umgebung existiert, wird diese als 30% angenommen. Damit hat sie in der Algorithmik eine neutrale Wirkung.

Hotspot-Metrik

Die "Hotspot-Metrik", die in LAB-TOUR SH entwickelt und eingeführt wurde, dient dazu, das Auslastungsverhalten einzelner POI über einen längeren Zeitraum zu beschreiben, basierend auf der zuvor beschriebenen Größe der tagesmaximalen Auslastung. Ziel ist es, POI als Hotspots, Coldspots oder durchschnittliche Orte zu klassifizieren. Im Alternativengenerierer von LAB-TOUR SH wurde die Hotspot-Metrik als Zahl zwischen 0 und 1 definiert. Ein POI gilt als Hotspot, wenn seine Hotspot-Metrik größer als 0,75 ist, und als Coldspot, wenn die Metrik kleiner als 0,25 ist. Eine Metrik von 0,5 wird als durchschnittlich angesehen. Die Nutzung einer solchen Metrik wird für den Betrieb eines touristischen Recommenders empfohlen, um eine Lenkungswirkung auf einer größeren Zeitskala zu erzielen als die bloße Betrachtung aktueller Auslastungen. Es ist wenig wünschenswert, derzeit wenig ausgelastete POI als Alternativen vorzuschlagen, wenn bekannt ist, dass sie regelmäßig zu Überfüllungen neigen. Dies könnte dazu führen, dass Menschen für diese POI begeistert werden, die sie vorher nicht kannten, und möglicherweise weiterhin diese Orte besuchen. Dadurch könnte Crowding sogar noch verstärkt werden. Daher sollte versucht werden, so wenig Aufmerksamkeit wie möglich auf solche POI zu lenken, die bereits als Hotspots bekannt sind. Es gibt jedoch eine Ausnahme von dieser Regel, wenn es keine anderen geeigneten POI in der Nähe gibt. In diesem Fall sollten auch bekannte Hotspots als Alternativen präsentiert werden, um den Benutzern eine konsistente User-Experience zu bieten. Auf der anderen Seite besteht ein großes Interesse daran, Coldspots so vielen Menschen wie möglich zu präsentieren. Die Einführung der Hotspot-Metrik ermöglicht es, dieses Verhalten im Alternativengenerierer-PoC abzubilden. Die Hotspot-Metrik lässt sich beispielsweise durch den im Folgenden skizzierten Weg bestimmen. Je nach Spezifika der im Recommender verwendeten Algorithmik ist es für die Hotspot-Metrik notwendig, lückenlos bei allen POI vorzuliegen. Um diesem Anspruch von vorneherein gerecht zu werden, kann in einem ersten Schritt die durchschnittliche Hotspot-Metrik (0.5) allen POI zugeschrieben werden. In einem zweiten Schritt können dann vergangene Auslastungsdaten eines jeden POI, falls verfügbar, herangezogen werden, um aus diesen statistisch eine Hotspot-Metrik abzuleiten. Dafür könnte zum Beispiel die maximale prozentuale Auslastung eines POI in den letzten 12 Monaten zwi-

schen 0 und 1 skaliert werden. Ein POI der im vergangenen Jahr über seine Belastungsgrenze gekommen ist und eine Auslastung größer als 100 Prozent erfahren hat, wäre demnach ein Hotspot mit der Metrik 1. Um das Verfahren weniger anfälliger für Ausreißer zu machen, könnte auch der Mittelwert der zehn Maxima der Auslastung im letzten Jahr für die Formulierung der Metrik herangezogen werden. Dieser Schritt wurde in LAB-TOUR SH aufgrund der sehr begrenzten Messzeiträume nicht durchgeführt. Eine genauere Analyse geeigneter statistischer Verfahren zur Bestimmung dieser Metrik ist allerdings ein neues, durch dieses Projekt geschaffenes, Forschungsziel der Tourismusforschung in Schleswig-Holstein. Im finalen Schritt zu Bestimmung der Hotspot-Metriken aller POI macht es durchaus Sinn, Tourismusdestinationen die Möglichkeit zu bieten, eine manuelle Überschreibung dieser durchführen zu können. Damit kann eine gezielte Lenkungswirkung in den Alternativengenerierer eingebracht werden, die zum Beispiel eine vielerorts gewünschte Lenkung von der Küste in Richtung Innenland möglich macht.

Auswahl von alternativen POI

Die zwei im Vorhinein ausgiebig diskutierten Eigenschaften Ähnlichkeit und Entfernung liegen in dem LAB-TOUR SH Meta-Knowledge-Graphen als Kanten zwischen allen POI vor, die nicht weiter als 25 km Luftlinie voneinander entfernt sind. Dies entspricht der maximalen Distanz, die im Recommender für eine Alternative zugelassen wird. Die Einschränkung dieser Kanten auf POI im Umkreis versucht die Größe des Knowledge Graphen und die Anzahl der notwendigen Berechnungen zu reduzieren. Weiterhin liegen die tagesmaximale Auslastung und die Hotspot-Metrik stets auf dem aktuellen Stand in den Eigenschaften der Knoten (POI) ebenfalls im Knowledge Graphen vor. Bei der Auswahl eines beliebigen POI innerhalb dieses Graphen können nun über die ausgehenden Kanten alle anderen POI identifiziert werden, die potenziell als Alternativen in Betracht kommen (alle im 25 km Umkreis). Außerdem liegen die Ähnlichkeit, Entfernung, Auslastung und Hotspot-Metrik direkt ebenfalls für alle diese POI vor.

Um aus diesen POI, entsprechend der im Abschnitt 5.2.1 eingeführten Zielsetzungen, ein Set an geeigneten Alternativen auszuwählen, wurde ein Algorithmus entwickelt, der zielführend, transparent und modular erweiterbar ist. Gerade die Komponente der Transparenz wurde als besonders wichtig erachtet, um die Funktionsweise des Recommenders stets genau verstehen zu können und so Fehler in Konzept und Programmierung leichter aufzudecken. Grundsätzlich funktioniert die Auswahl an Alternativen in diesem PoC so, dass mehrere, kleinere Sets an Alternativen zusammengefasst werden, die jeweils recht eindeutig tituliert werden können und durch klare Kriterien an die vier POI-Eigenschaften des Recommenders definiert werden können. Insgesamt entspricht die Anzahl an potenziellen alternativen POI, je nach lokaler POI-Dichte bis zu maximal acht POI. Diese Anzahl kann in einem Frontend gut dargestellt werden und einem/r Nutzer*in trotzdem genug Abwechslung anbieten.

Die folgende Tabelle 23 zeigt die vier Sub-Sets aus denen Alternativen für eine Empfehlung am heutigen Tage zusammengefasst werden. Jedes Sub-Set hat dabei seine eigene Definition anhand der vier POI-Eigenschaften. Im Folgenden werden die vier Sub-Sets auch schriftlich noch einmal genauer erläutert.

	Einflussfaktoren			
Anzahl von Alternativen	Ähnlichkeit	Entfernung	Auslastung	Hotspot-Metrik
0 - 3	= 1	< 10 km	< 60%	-----
	POI des exakt gleichen Typs. Auswahl der nächsten drei POI.			
0 - 2	> 0.9 und < 1	< 10 km	< 50%	< 0.75
	POI eines sehr ähnlichen Typs. Zufällige Auswahl.			
0 - 2	> 0.6 und < 0.9	< 10 km	< 50%	< 0.75
	POI eines ähnlichen Typs. Zufällige Auswahl.			
0 - 1	> 0.25	< 10 km	< 50%	< 0.25
	Coldspot. Zufällige Auswahl.			

Tabelle 23: Sub-Sets und Alternativen

Gemäß dem Konzept werden die Alternativen für den heutigen Tag aus vier kleineren Sub-Sets gebildet. Das erste Sub-Set besteht aus bis zu drei POI desselben Typs wie derjenige, für den eine Alternative gesucht wird. Wenn beispielsweise ein Strand der Ausgangs-POI ist, werden in diesem Sub-Set auch Strände präsentiert, vorausgesetzt sie liegen innerhalb eines Radius von 10 km. Falls es mehrere POI desselben Typs im Umkreis von 10 km gibt, werden immer die nächstgelegenen ausgewählt. Dieser Ansatz zielt darauf ab, POI im Sub-Set zu präsentieren, die in allen Dimensionen möglichst ähnlich zum Ausgangs-POI sind, einschließlich ihrer Lage. Allerdings werden nur solche POI in Betracht gezogen, deren maximale Tagesauslastung unter 60% der Belastungsgrenze liegt. POI, die eine höhere Auslastung aufweisen, sind grundsätzlich von einer Empfehlung ausgeschlossen. Damit soll verhindert werden Crowding künstlich zu bestärken, wenn laut Messungen oder Prognose ohnehin schon die Gefahr besteht, dass dies bereits von selbst auftreten könnte. Das zweite Set an Alternativen liefert bis zu zwei POI, die sehr ähnlich, aber nicht vom gleichen Typ sind. Dafür wird eine Ähnlichkeit von größer als 0.9, aber kleiner als 1 verlangt. Weiterhin wird ab dem zweiten Sub-Set eingeführt, dass es sich bei den vorgeschlagenen POI nicht um bekannte Hotspots handeln darf, selbst dann nicht, wenn diese am heutigen Tag weniger als 50% ausgelastet sind. Um dies zu gewährleisten wird das Kriterium eingeführt, dass die Hotspot-Metrik kleiner als 0.75 sein muss. Damit werden einem/r Nutzer*in POI vorgeschlagen, die hoffentlich nahe an seiner ursprünglichen Intention liegen, aber auch etwas Neues sind. Eine der Hauptintention dieses Sub-Sets im praktischen Sinne ist es, Ausflügler von den oft überlaufenen Stränden hin zu Badestellen und Freibädern weiter im Innenland zu lenken. Das dritte Sub-Set hat die Intention den Horizont eines/r Anwender*in zu erweitern und andere attraktive POI vorzustellen, die vielleicht etwas weiter weg von der ursprünglichen Intention des/der Anwender*in liegen. Das Ziel dieses Sub-Sets ist dabei nicht unbedingt, dass ein/e Nutzer*in noch heute seinen Ursprungs-POI durch

einen dieser POI austauscht, sondern viel mehr die Aufmerksamkeit des /der Anwender*in langfristig auf die Vielfalt der Schleswig-Holsteinischen Ausflugslandschaft zu lenken und Nutzer*innen zu neuen Aktivitäten zu inspirieren. Das letzte der Sub-Sets ist gedacht, um eine übergeordnete Lenkungswirkung zu etablieren. Hier wird ein Coldspot ausgespielt, falls ein solcher im Umkreis von 10km vorhanden ist. Dafür wird verlangt, dass die Hotspot-Metrik unter 0.25 liegt. Diese Coldspots müssen nicht sehr ähnlich zum Ausgangs-POI sein, da dieses Sub-Set eher als eine Art Promotion-Fläche konzipiert ist über die Tourismus-Destination ganz bewusst einen Steuerungseffekt einsetzen können. Dementsprechend müssen diese POI keinen großen Bezug zum Ausgangs-POI aufweisen.

Die Kriterien ändern sich bei der Auswahl von Alternativen, die für ein Zeitraum vorgeschlagen werden, der weiter in der Zukunft liegt. Die Auswahl der Alternativen für einen Tag in der Zukunft setzt sich beispielsweise zusammen aus den folgenden Sub-Sets (siehe Tab. 24):

	Einflussfaktoren			
Anzahl	Ähnlichkeit	Entfernung	Auslastung	Hotspot-Metrik
0 - 2	= 1	< 15km	< 60%	-----
	POI des exakt gleichen Typs. Auswahl der nächsten beiden POI.			
0 - 3	> 0.9 und < 1	< 15km	< 50%	< 0.75
	POI eines sehr ähnlichen Typs. Zufällige Auswahl.			
0 - 2	> 0.5 und < 0.9	< 15km	< 50%	< 0.75
	POI eines ähnlichen Typs. Zufällige Auswahl.			
0 - 1	> 0.25	< 15km	< 50%	< 0.25
	Coldspot. Zufällige Auswahl.			

Tabelle 24: Sub-Sets und Alternativen für 1 Tag in Zukunft

Die prominenteste Änderung ist hierbei die Erweiterung des Suchgebietes auf 15km. Weiterhin wird die Anzahl der pro Sub-Set gelieferten POI leicht angepasst. Es werden hier nur maximal zwei POI vom gleichen Typ vorgeschlagen, dafür aber drei eines sehr ähnlichen Typs. Diese beiden Anpassungen sollen die höhere Flexibilität eines/r Anwender*in bei einem erweiterten Recommendation-Horizont abbilden.

Für zwei Tage in der Zukunft werden folgende Sub-Sets genutzt (siehe Tab. 25):

Einflussfaktoren				
Anzahl	Ähnlichkeit	Entfernung	Auslastung	Hotspot-Metrik
0 - 2	= 1	< 20km	< 60%	-----
	POI des exakt gleichen Typs. Auswahl der nächsten beiden POI.			
0 - 2	> 0.9 und < 1	< 20km	< 60%	< 0.75
	POI eines sehr ähnlichen Typs. Zufällige Auswahl.			
0 - 1	> 0.6 und < 0.9	< 20km	< 50%	< 0.75
	POI eines ähnlichen Typs. Zufällige Auswahl.			
0 - 2	< 0.6	< 20km	< 50%	<= 0.5
	Alternative POI. Zufällige Auswahl.			
0 - 1	> 0.25	< 20km	< 50%	< 0.25
	Coldspot. Zufällige Auswahl.			

Tabelle 25: Sub-Sets und Alternativen für 2 Tage in Zukunft

Für einen zweitägigen Recommendation-Horizont wird das Suchgebiet auf 20km erhöht. Außerdem wird die Anzahl der Sub-Sets erhöht, indem ein neues Sub-Set eingeführt wird, das alternative POI vorschlägt, die nicht ähnlicher als 0.6 sind und maximal eine neutrale Hotspot-Metrik aufweisen. Ziel dieses Sub-Sets ist wieder die Erweiterung des Horizonts eines/r Anwender*in.

Um keine Redundanz zu erzeugen, wird an dieser Stelle darauf verzichtet, die Kriterien für einen dreitägigen Recommendation-Horizont aufzuzeigen. Die einzige Änderung im Vergleich zu einem zweitägigen Horizont ist dabei, dass der Suchradius auf 25km erhöht wird.

Neben der Auswahl von alternativen POI wird außerdem eine Auswahl von geeigneten Parkplätzen für einen POI durchgeführt. Dafür werden alle Parkplätze im Umkreis von 2km um den POI anhand nachfolgender Gewichtung bewertet. $Score = Entfernung/Kapazität$. Nach diesem Score werden dann alle Parkplätze aufsteigend sortiert und die besten acht ausgespielt. Die Auslastung geht bewusst nicht mit in die Sortierung mit ein. Das beruht auf dem Ansatz, dass ein potenziell voller Parkplatz im Frontend der Vollständigkeit halber auftauchen soll, dort allerdings als auch gut sichtbar als bereits voll markiert wird. Sollte dies nicht getan werden, könnten Nutzer*innen den bereits vollen Parkplatz trotzdem anfahren, wenn sie denken, er sei nicht im System erfasst.

Sortierung von alternativen POI

Die über das beschriebene Verfahren ausgewählte Alternativen werden nun in einem weiteren Schritt geordnet. Dem zugrunde liegt die Annahme, dass eine implizite Priorisierung von Elementen gegeben ist, wenn diese von links nach rechts in einer gewissen Reihenfolge angezeigt werden. Diese Priorisierung bewertet das erste Element (ganz links) als das wichtigste Element und alle anderen Elemente absteigend entsprechend ihrer Reihenfolge. Aus diesem Grund wurde entschieden, die Alternativen nach einer „Qualität“ zu sortieren. In diese Metrik gehen alle Faktoren ein, die auch in den Kriterien-Sets zu einer Auswahl der POI geführt haben. Folgende Formel beschreibt die zur Sortierung der POI genutzte Metrik.

$$\text{Qualität} = 2 \cdot \text{Ähnlichkeit} - \left(\frac{\text{Entfernung}}{z} \right) - \text{Auslastung} - \frac{1}{2} \cdot \text{Hot-Spot-Metrik}$$

Mit z als der maximal erlaubten Entfernung in diesem Kriterien-Set. Diese Transformation sorgt dafür, dass alle Faktoren in dieser Metrik zwischen 0 und 1 liegen. Die Ähnlichkeit wurde in einer eingeführten Gewichtung der Faktoren doppelt gewichtet, da davon ausgegangen wird, dass diese für eine/n Nutzer*in die wichtigste Eigenschaft einer Alternative ist. Die Hotspot-Metrik wurde über eine Gewichtung in ihrem Einfluss halbiert, da diese Information für eine/n Nutzer*in wahrscheinlich nicht sehr relevant ist. Diese Qualität wird vor der Ausspielung für alle ausgewählten Alternativen berechnet und diese dann anhand dessen absteigend sortiert.

Die vorherigen Seiten haben einen detaillierten Einblick in die interne Logik des Alternativengenerierers geboten. Der nun folgende Abschnitt soll weiterhin einen kurzen Überblick über die Details bieten, die in der technischen Umsetzung in der Logik in diesem PoC eine Rolle gespielt haben.

5.2.5 Technische Details

Dieser Abschnitt beschreibt die Funktionalität des Recommender-Moduls aus der Programmcodeperspektive. Es behandelt sowohl dessen Inbetriebnahme als auch das Verhalten in Produktion. Zuletzt werden Limitationen aufgeführt, die der zugrundeliegenden Funktionalität unterliegen.

Die Initialisierung des Recommender Programmcodes erfolgt wie folgt: Beim Start des Recommender-Moduls werden die POI-Daten beim Context Broker angefragt. Zu einem POI gehören statische Daten wie Art oder Ort, aber auch dynamische Daten wie die aktuelle Auslastung oder, im Falle eines Parkplatzes, die Kapazität dessen. Anschließend werden die POI-Daten zueinander in Relation gesetzt. Dabei wird die Luftlinien-Entfernung von jedem POI zu jedem anderen POI anhand von Längen- und Breitengrad eines POI berechnet. Wenn die Luftliniendistanz in Kilometern unter 25 liegt, wird die Luftliniendistanz durch die Fahrtroutendistanz mit einem Auto basierend auf Open-Source-Route-Machine Daten ersetzt. Die Fahrtroutendistanz wurde in einem vorherigen Schritt berechnet und liegt im Recommender-Modul statisch vor. Im darauffolgenden Schritt werden die POI-Daten als Knoten und die Relationen als Kanten in eine Neo4j Graph-Datenbank geladen sofern die Distanz einer Relation nicht die maximale Distanz, die ein Alternativvorschlag von einem Ausgangs-POI entfernt sein darf, übersteigt (25 km). Die Daten in der Neo4j Datenbank werden anschließend persistiert. Darauffolgend werden die vorhandenen Auslastungsprognosen von einem POI auf umliegende POI innerhalb eines 2 km Radius übertragen. Diese Übertragung geschieht mit Hilfe der Graph Datenbank, auf die mit der Abfrage-Sprache Cypher zugegriffen wird. Der übertragende Auslastungswert berechnet sich aus dem Durchschnitt der umliegenden POI-Auslastungen, für die es eine Prognose und folglich Sensor- oder WLAN-Nutzungsdaten gibt.

Während des Betriebs werden die POI, zu denen regelmäßig eine Auslastungsprognose erstellt und die Hotspot-Metrik berechnet wird, über eine Subscription zum Context Broker aktualisiert. Mit der Auslösung der Subscription auf dem Context Broker werden die Auslastungsdaten mit einem GET-Request in das Recommender-Modul geladen. Aus dem vom Prognose-Modul prognostizierten 24-Stundenintervalle für heute, morgen, übermorgen und überübermorgen wird jeweils der Maximalwert selektiert und auf die POI in der Neo4j Datenbank übertragen. Anschließend werden die berechneten Auslastungsspitzen auf die umliegenden POI übertragen. Im Abschluss werden die propagierten Auslastungswerte zurück zum Context Broker gespielt, damit das Frontend auf diese Informationen Zugriff hat. Abschließend werden für alle 4 Zeitintervalle (heute, morgen, übermorgen und überübermorgen) für jeden Knoten (POI) in der Neo4j-Datenbank die POI als auch Parkplatz Alternativen generiert und auf dem Context Broker gespielt.

Zusammenfassend werden während der Initialisierung folgende Schritte durchlaufen:

1. Laden der POI-Daten (Get-Request) aus dem Context Broker
2. Berechnung der Knoten (POI) und Kanten (Relationen zwischen den POI) für die Graph Datenbank Neo4j
3. Hinzufügen der Knoten und Kanten zur Graph Datenbank

Nach der Initialisierung erfolgen die Inbetriebnahme und regelmäßige Aktualisierung in folgenden Schritten:

1. Löschen der Auslastungsdaten in der Graph Datenbank
2. Laden der aktuellen POI-Auslastungsdaten vom Context Broker
3. Hinzufügen der Auslastungsdaten zur Graph Datenbank
4. Propagierung der Auslastungsdaten auf umliegende POI
5. Berechnung der Alternativen für verschiedenen Zeitintervalle und POI
6. Zuspielung der berechneten Alternativen an den Context Broker

Die Fahrroute-Berechnung, von einem zum anderen POI, umfasst zwei wichtige Klassen: Driving Route Processor und Batch Processor. Sie ermöglichen die Verarbeitung und Analyse von Fahrtrouten, einschließlich Distanz, Fahrtdauer und Luftlinienentfernung. Der Code arbeitet in Chargen, speichert den Fortschritt und verwendet externe Ressourcen wie die OSRM-API. Der Driving Route Processor berechnet einzelne Routen zwischen geographischen Koordinaten, während der Batch Processor die Batch-Verarbeitung von Routen ermöglicht. Es gibt auch ein Batch-Skript zur Automatisierung des Prozesses und eine Abhängigkeiten-Datei für die erforderlichen Python-Pakete.

Für ein besseres Verständnis der verwendeten Technologie-Stacks ist im Folgenden eine Auflistung der genutzten Technologien aufgelistet.

1. Programmiersprache: Python
2. Containervirtualisierung: Docker
3. REST (Representational state transfer)-API Framework: FastAPI
4. Graph-Datenbank: Neo4j
5. Python-Packages: os, ast, dataclasses, logging, math, time, dotenv, numpy, pandas, itertools, sys, py2neo, datettime, urllib, json

6. Eigens entwickelte Klassen: Neo4jManager, ContextBrokerManager

Konkret auf den Recommender Programmcode sollten folgende Punkte berücksichtigt werden: In einem Feldversuch sollte evaluiert werden, ob die Methodik, bei der die POI und Parkplätze nach ihrer Relevanz sortiert werden, überprüft werden. Des Weiteren ist zu berücksichtigen, dass bei der Berechnung der Beziehungen (Kanten) für die Graph Datenbank steigt die Menge der Kanten exponentiell zu der Menge an Knoten. Ein Graph mit zwei Knoten enthält genau eine Kante. Die analog semantische Beschreibung ist: Es gibt zwei POI, bei denen von denen vom einen POI zum anderen POI gereist werden kann. Letzteres beschreibt die Beziehung der POI untereinander. Ein Graph mit 3 Knoten hat dann genau 3 Kanten, einer mit fünf Knoten 10 Kanten und mit 7 Knoten 21 Kanten. Daraus ergibt sich folgende Formel:

$$\sum_{n=1}^n (n - 1)$$

Bei 2.300 POI ergibt das 2.649.350 Kanten, welche zur Graph-Datenbank hinzugefügt werden müssen. Die Laufzeitkomplexität steigt mit der Menge an Kanten. Wenn ein Knoten zu den 2.300 hinzugefügt wird, erhöht sich die zu verarbeitende Kantenmenge um 2.299 Kanten. Mit einer Beschränkung der Beziehungen durch eine obere Schranke, eine Kante zwischen zwei Knoten in dem Graphen existiert genau dann, wenn die Fahrtroutendistanz zwischen diesen unter 25 km liegt, kann die zu verarbeitende Menge an Kanten in der Graph-Datenbank signifikant reduziert werden und bleibt bis zu einem Bereich von ca. 3.000 Knoten gut beherrschbar. Ferner muss, durch die statische vorab Fahrtroutendistanzberechnung, diese beim Erweitern der POI-Daten um einen neuen weiteren POI für alle Kanten neu berechnet werden. Die Menge an dabei zu berechnenden Fahrtrouten kann reduziert werden, indem vorab die leichter zu berechnende Luftliniendistanz berechnet wird. Ist diese größer als 25 km, ist die Fahrtroutendistanz auch größer als 25 km und muss daher nicht mehr berechnet werden. Dieser POI steht dann in keiner Beziehung zum neu hinzugefügten POI, folglich existiert dann keine Kante in der Graph-Datenbank zwischen diesen POI. Hier wird die Annahme getroffen, dass die Fahrtroutendistanz immer länger oder gleich lang wie die Luftliniendistanz ist. Zudem könnte eine dynamische Fahrtroutendistanzberechnung auch noch Informationen wie Stau- und Wetterdaten in Zukunft berücksichtigen.

5.2.6 Learnings und Potentiale

Dieser Abschnitt soll analog zum gleichen Abschnitt im vorigen Abschnitt reflektieren, was für Learnings in Bezug auf den Alternativengenerierer im Projekt LAB-TOUR SH gemacht wurden, welche Potentiale entdeckt und welche Schwierigkeiten sich ergeben haben. Außerdem soll ein Ausblick gegeben werden, über die nächsten Schritte die stattfinden sollten, um die Erkenntnisse dieses Projekts in der Schleswig-Holsteinischen Tourismusforschung und -praxis zu verankern.

Datengrundlage

Einer der Punkte, der in diesem Projekt als maßgeblich wichtig für den Erfolg eines produktiv eingesetzten Recommenders ausgemacht wurde, bezieht sich auf die Wissensbasis aus POI, die einem solchen Recommender zur Verfügung steht. In einem Produktivsystem müsste es ermöglicht werden,

dass die Datengrundlage des Recommenders stetig wachsen kann und stets die aktuellen Daten verwendet werden. Hierfür wäre eine Live-Anbindung an die Datenbank der TA.SH sinnvoll. Dabei sollte dann ein Quality-Gateway eingeführt werden, sodass neue oder veränderte POI nach ihrer Anreicherung im Recommender noch einmal geprüft werden. Alternativ wäre weiterhin denkbar, das gesamte Preprocessing in Abstimmung mit der TA.SH auf deren Plattform auszulagern. In diesem Fall könnte das Quality-Gateway auch bei der TA.SH liegen, was durchaus empfehlenswert wäre, um klare Verantwortlichkeiten zu schaffen. Damit würde einhergehen, dass die TA.SH sich bei der Erstellung der POI näher an dem POI-Datenmodell von Smart Data Models und/oder der ODTA orientiert.

Die POI-Datengrundlage dieses PoC ist zum jetzigen Stand nicht ausgereift genug, um produktiv ausgespielt zu werden. Dies liegt zum einen an der limitierten Anzahl der zur Verfügung stehenden POI, zum anderen an der teilweise mangelhaften Genauigkeit dieser und letztlich auch daran, dass das Verfahren zur Extraktion des POI-Typs aus den sehr heterogen vergebenen Tags der POI nicht fehlerlos funktioniert. Ein anderes Verfahren, das sich zur Zuordnung der POI-Typen nicht auf die Tags konzentriert, sondern viel mehr auf den Titel und die Beschreibung eines POI, mit Zuhilfenahme der vergebenen Tags, wird in Abschnitt 5.2.3. vorgeschlagen. Um die Genauigkeit der Datengrundlage zu verbessern, ist eine komplette Überprüfung der bisher eingepflegten POI notwendig. Dabei sollte darauf geachtet werden, dass Geokoordinaten richtig gesetzt wurden, Dummy-Einträge gelöscht und Rechtschreibfehler korrigiert werden. Weiterhin sollten Parkplatzkapazitäten und Öffnungszeiten aus den Beschreibungen von POI entfernt und in die dafür vorgesehenen Felder der Datenmodelle übertragen werden. In einer solchen Sichtung der gesamten Datengrundlage könnte außerdem der POI-Typ ergänzt werden. Zur Vergrößerung der Datengrundlage im Allgemeinen empfiehlt es sich, Daten aus OpenStreetMap zu extrahieren und diese mit der Datenbasis zu verschneiden. In OpenStreetMap sind beispielsweise Parkplätze deutlich umfangreicher gepflegt als in der Datenbasis der TA.SH, aber auch weitere Strände, Museen und Restaurants. Ein Nachteil der Daten von OpenStreetMap ist jedoch, dass nicht notwendigerweise ansprechende Beschreibungen für die POI existieren. Diese können jedoch in einigen Fällen, zusammen mit Bildern der POI, über die in OpenStreetMap vorhandene Verlinkung zu WikiData extrahiert werden. Neben dieser Option der externen Datenbeschaffung könnte außerdem versucht werden, bereits vorhandene POI mit eingeschränkter Lizenz unter eine offene Lizenz zu bringen, damit sie in die Datengrundlage eines Recommenders einfließen können. Gerade der ländliche Raum ist in diesem PoC stark unterrepräsentiert. Die Vermutung liegt nahe, dass dies besonders so ist, weil die Aktivitäten Radfahren und Wandern in der Datengrundlage dieses PoC stark unterrepräsentiert sind. Potenziell bietet sich hier eine Kooperation mit kommerziellen Anbietern von Wander- und Radtoursplanung an, um diese Lücke zu schließen.

Zuordnung von POI-Typen und Ähnlichkeit

Wie bereits erwähnt, ist es gängige touristische Praxis, POI mit Tags zu versehen, die nicht einer einheitlichen Nomenklatur folgen, sondern individuell von der datengenerierenden Organisation erstellt werden. Da so allerdings eine sehr deutliche Heterogenität innerhalb der Tags erzeugt wird, ist es sehr schwer aus ihnen zu folgern, um was für einen Typ POI es sich genau handelt. Da diese Information aber nach dem vorliegenden Konzept in einem Recommender gegeben sein sollte, wurde in diesem PoC ein Verfahren exploriert, das versucht aus den Tags eines POI, trotz der Heterogenität, den Typ zu folgern. Diese Exploration hat die Schwierigkeiten dieses Ansatzes offengelegt und gezeigt, dass ein solches Verfahren nicht skalierbar gute Ergebnisse liefern kann. Aus diesem Grunde wurde ein weiteres Verfahren entworfen, das darauf beruht, Gruppen von POI zu Clustern zusammenzufassen. Dafür wurden die POI-Elemente der Beschreibung, des Titels und der Tags in einem ersten Schritt bereinigt. Aus der Beschreibung wurden dabei, wie im Natural Language Processing gängig, Füllwörter entfernt

und aus den Tags solche Tags entfernt, die keinen Bezug zu einer touristischen interessanten Eigenschaft haben, beispielsweise „Themen“ oder „individuelle Portale“. Dann wurden jeweils für den Titel, die Beschreibung und die Tags, über ein vortrainiertes Sprachmodell in Vektoren (Document Embeddings) übersetzt. Aus den resultierenden drei Vektoren wurde dann der Mittelwert gebildet, so dass final ein Vektor einen POI beschreibt. Anders als im vorherig beschriebenen Verfahren sind diese Dimension allerdings losgelöst von touristischen Eigenschaften und für einen Betrachter nicht nachvollziehbar. Auf diese Vektoren wurde dann ein Clustering-Algorithmus angewandt, der die Bildung einer fest definierten Anzahl an Clustern zum Ziel hat. In einem ersten Versuch wurden dafür 25 Cluster vorgeschrieben – da dies der Größenordnung der Anzahl an sinnvoll unterscheidbaren Aktivitätsgruppen ungefähr entsprechen sollte. Über Methoden des Topic-Modelling wurden dann die Kern-Wörter, die ein jedes Cluster ausmachen, ausgegeben, um so die Cluster für einen Menschen nachvollziehbar zu machen. Aufgrund fehlender Zeit konnte keine genauere Prüfung dieser Cluster innerhalb des Projektzeitraums stattfinden, allerdings scheint die Methode vielversprechend. Die entstehenden Cluster waren händisch verschlagwortbar, so dass auch an dem Frontend eine Information ausgespielt werden kann, zu welcher Aktivität, oder zu welchem Typ sich ein POI zuordnen lässt. Abbildung 66 zeigt exemplarisch die Kernbegriffe so erstellter Cluster.



Abbildung 66: Cluster von Kernbegriffen für Aktivitäten oder POI

Auf Grundlage dieser Methode ist ein weiteres Verfahren denkbar, dass mit Hilfe von Dokument Embeddings einer weiteren, in diesem Forschungsprojekt identifizierten Problematik entgegenzutreten könnte. Gemeint ist dabei die Ähnlichkeit von POI untereinander, die bis jetzt auf den händisch generierten Eigenschafts-Vektoren beruht. Denkbar wäre hier beispielsweise die Ähnlichkeit der verschiedenen Cluster zu bestimmen, indem die Centroid-Vektoren eines jeden Clusters über die Cosinus-Similarity miteinander verglichen werden. Die Ähnlichkeit zweier POI könnte dann, ähnlich wie in diesem PoC, der Ähnlichkeit der jeweiligen POI-Typen (POI-Clustern) entsprechen. Um individuelle Eigenschaften des POI zu berücksichtigen, könnte sogar der Vektor eines POI direkt mit allen anderen POI-Vektoren über die Cosinus-Similarity verglichen werden. Dies würde allerdings einen Verlust an Nachvollziehbarkeit für das Verfahren bedeuten und eine Abhängigkeit von der Art der Beschreibungen eines POI einführen. Außerdem wäre eine so eingeführte Ähnlichkeit nicht nur auf den touristischen Kontext bezogen, sondern auf alle Kontexte, die in den Trainingsdaten des genutzten Sprachmodells vorkommen. Dieses Verfahren sollte für dein Einsatz in einem produktiven Recommender weiter untersucht werden. Falls die Untersuchung ergeben sollte, dass das Verfahren nicht gut performt, können die Cluster, wie in diesem PoC händisch in verschiedenen touristischen Dimensionen eingeordnet werden und auf Grundlage dessen verglichen werden.

Personalisierung

Als weiterer Kernaspekt wurde in diesem Forschungsprojekt der Faktor „Personalisierung“ als Chance für den Recommender ausgemacht, um sein volles Potential auszuschöpfen. Grundsätzlich sind im Bereich der Personalisierung aus Sicht der Autoren drei Möglichkeiten gegeben. Eine davon wäre die Einführung eines User-Systems, in dem ein/e User*in sich über einen Login an der Recommender-Plattform anmeldet. Dadurch könnten sämtliche Interaktionen, die diese/r User*in mit den Empfehlungen hat, gespeichert und verarbeitet werden. Weiterhin könnten bei der Anmeldung persönliche Daten wie Alter, Geschlecht und Interessen abgefragt und in die Empfehlung von Alternativen mit eingebunden werden. Die zweite Möglichkeit wäre die Etablierung eines Session-basierten Users. Dafür könnte jeweils während der Nutzung des Recommenders ein Profil der/des User*in angelegt werden, das nach der einmaligen Nutzung wieder verworfen wird. Ein solches User-Profil kann in der einfachsten Form durch ganz konkreten User-Input über das Frontend angelegt werden. Dafür könnten beispielsweise Fragen zum heutigen Interesse gestellt werden, für die mehrere Antwortmöglichkeiten existieren würden. Mit Hilfe dieser Antwortmöglichkeiten könnten User*innen, entsprechend der in Schleswig-Holstein definierten touristischen Zielgruppen, aufgeteilt werden und versucht werden, passende Alternativen auszuspielen. Weiterhin könnte erfragt werden, ob ein/e User*in mit Kindern oder einem Hund unterwegs ist und auch darauf eingegangen werden. Zusätzlich könnte auch Live getrackt werden, welche angebotenen Alternativen eine/n User*in interessiert haben (ausgemacht durch das Anklicken der jeweiligen Alternative) – und diese Information mit zur Auswertung herangezogen werden. Der dritte denkbare Ansatz beruht auf einem/r Interaktions-basierten User*in. Dieser würde jedes Mal neu erstellt werden, wenn ein/e User*in nach Alternativen für einen bestimmten POI sucht, und dann wieder verworfen werden. Dafür würde bei der Auswahl eines POI erst einmal ein generisches Set an Alternativen ausgespielt werden. der/die User*in bekommt allerdings auch die Möglichkeit auszuwählen, was seine Motive für den Besuch dieses POI sind. Dafür könnte eine vordefinierte Liste an möglichen Motiven ausgespielt werden. Im Folgenden würde dann ein neues spezifisches Set an Alternativen gebildet, das diesen Motiven möglichst gut entspricht. Diese „Aktivitätsfilterung“ könnte dann in die Alternativen die Intention eines/r Anwender*in klarer einbeziehen. Dafür müsste zu jedem POI eine Zuordnung an mögliche Aktivitäten vorliegen. Für einen Strand könnten dies beispielsweise die Aktivitäten „Spazieren“, „In der Sonne sein“, usw. sein. Diese Aktivitäten könnte nun im Frontend sichtbar gemacht und einem/r Nutzer*in die Möglichkeit gegeben werden, aus diesen auszuwählen, welche Aktivitäten oder Eigenschaften ihm/ihr besonders wichtig sind. Über die Jaccard-Similarity der ausgewählten Intentionen unter den POI könnte somit eine weitere Filterbedingung geschaffen werden, die in der Bildung von Sub-Sets berücksichtigt werden kann und dort steuern kann, dass POI die Erfüllung möglichst vieler derselben Intentionen erlauben. Dieser Ansatz könnte zu einem Session-basierten Useransatz ausgedehnt werden, indem die ausgewählten Motive gespeichert werden und zum neuen Default auch bei anderen angeklickten POI werden.

Beinahe unabhängig davon welcher Ansatz hier zur Personalisierung gewählt wird, macht es Sinn die Umfrage, auf der die Ähnlichkeit der POI im LAB-TOUR SH Alternativengenerierer-PoC zum Großteil beruht, daran anzupassen und neu durchzuführen. Sollte im Personalisierungskonzept versucht werden über Zielgruppen zu arbeiten, sollte dies entsprechend in der Umfrage berücksichtigt werden und die Umfrage mit diesen Zielgruppen gesondert durchgeführt werden. Eine große Schwäche des von uns eingeführten Verfahrens ist, dass die genutzte Umfrage sich auf weitere Interessen innerhalb eines Kurzurlaubes bezieht. Es geht dort also nicht um die direkte Austauschbarkeit, sondern vielmehr um die Frage, was ein Befragter gerne noch in seinem Urlaub machen möchte. Sollte die Möglichkeit bestehen, Einfluss auf das Design der Umfrage zu nehmen, könnte diese zeitlich enger gefasst und so mehr auf die Austauschbarkeit innerhalb eines Tages ausgelegt werden. Im vorhergegangenen Abschnitt wurde die Möglichkeit eingeführt, Cluster von POI zu bilden, die im Recommender als eine Art

POI-Typ fungieren könnten. Je nach Granularität könnte dies beispielsweise alle POI zusammenfassen, an denen schwimmen gegangen werden kann (Strand, Badestelle, Freibad, Schwimmbad, ...). Es könnte sinnvoll sein, die Umfrage so durchzuführen, dass die Fragestellung zweistufig formuliert wird, als „Der Besuch welcher Gruppe von POI liegt heute in Ihrem Interesse?“ und dann folgend für jede der in der Eingangsfrage genannten Cluster „Wenn der Besuch dieser POI nicht möglich wäre, welche Gruppe von POI würden sie dann gerne besuchen?“. Über die resultierenden Antworten könnten gut generische Alternativen ausgespielt werden, die sich entweder im gleichen oder einem, nach dieser Umfrage, nahen Cluster an POI befinden. So kann auch abgefragt werden, wie flexibel ein/e Nutzer*in der Auswahl der Alternativen ist.

Eine weitere Möglichkeit zur Personalisierung wäre es, Anwender*innen den Entscheidungsspielraum darüber zu geben, in welchem Umfeld für sie nach Alternativen gesucht werden soll. Wenn es außerdem gelingen würde die jetzige Lokation des /der Anwender*in, beispielsweise über die IP-Adresse zu approximieren, könnte auch die Entfernung vom aktuellen Standort mit in die Auswahl einfließen und im Sinne der Nachhaltigkeit Nutzer*innen möglichst nahe gelegene Empfehlungen ausspielen.

Weitere Filterbedingungen

In den eingeführten Anforderungen an ein Set an Alternativen wurde verlangt, dass die Öffnungszeiten, das Wetter und die Jahreszeit in einer Recommendation berücksichtigt würde. Aufgrund der vorliegenden Datengrundlage und der begrenzten Entwicklungskapazitäten konnten diese im LAB-TOUR SH PoC allerdings nicht berücksichtigt werden. Die entsprechende Funktionalität wurde in der Entwicklung jedoch bereits mitgedacht und kann ohne allzu großen Aufwand in die bisherige Algorithmik integriert werden. Da im Knowledge-Graphen alle Informationen alle 15min aktualisiert werden, kann dort auch stets der aktuelle Öffnungszustand eines POI hinterlegt werden. Ebenso die aktuelle Wettereignung – die auf dem ebenfalls aktuellen Wetter beruht. Bei einer Anfrage für Alternativen könnte dann in die Filterkriterien aufgenommen werden, dass ein POI geöffnet und für das aktuelle Wetter geeignet sein muss, um vorgeschlagen werden zu können. Zusätzlich kann ein Frontend die detaillierten Öffnungszeiten anzeigen und auch Warnhinweise ausspielen, falls ein POI noch geöffnet ist, aber schon bald schließt. Ein ebensolches Verfahren kann auch für die Wettereignung eingeführt werden, wobei noch definiert werden muss, wie diese für jeden POI klar bestimmt werden kann. Denkbar ist vor allem die Temperatur und Niederschlagsmenge zu berücksichtigen, potenziell aber auch eine Wassertemperatur und die aktuellen Windverhältnisse. Im besten Falle wird dabei außerdem nicht nur das aktuelle Wetter berücksichtigt, sondern die kurzfristige Vorhersage für die nächsten Stunden, sodass keine Strandbesuche empfohlen werden, wenn ein Sommergewitter vor der Tür steht. Dieses Verfahren sollte für jeden POI-Typen eine Funktion etablieren, die ausgehend von den Wetterdaten als Input einen Wert zwischen 0 (keine Eignung) und 1 (perfekt geeignet) erzeugt. Indoor-POI-Typen können dabei stets mit einer perfekten Eignung versehen werden. In den Filterkriterien kann dann ein Threshold gesetzt werden, der beispielsweise eine Wettereignung größer 0.5 verlangt. Wenn das Verfahren zur Bestimmung der Wettereignung umfassend genug ausgelegt ist, berücksichtigt dies gleichzeitig auch die Jahreszeiten und würde z.B. bei Schneefall keine Fahrradtouren vorschlagen. Über die Öffnungszeiten könnte ein anderer Teil der Saisonalität berücksichtigt und beispielsweise alle Badestellen im Winter künstlich geschlossen werden, sodass diese dann nicht mehr vorgeschlagen werden.

Sonstige Learnings

Der Algorithmus, der die Auswahl von Alternativen anhand der Eigenschaften der POI steuert, kann ebenfalls verbessert werden. Das in diesem PoC explorierte und vorgestellte Verfahren der Auswahl

von Alternativen anhand von Kriterien-Sets ist dabei nicht das einzig denkbare Verfahren. In der Konzeptionierungsphase des PoC wurde ausgiebig ein Verfahren diskutiert, das auf Grundlage der verfügbaren Einflussparameter ein Ranking aller umliegenden POI ausrechnet. Aus diesem Ranking folgend könnten dann die zehn besten POI ausgespielt werden. Das Ranking selbst könnte dabei auf einer Formel beruhen, die ganz ähnlich der Formel ist, die in diesem POI für die Sortierung der ausgespielten Alternativen verantwortlich ist. Es wurde sich gegen in diesem PoC gegen ein solches Verfahren entschieden, um die Nichtlinearität der Einflussparameter nicht zu stark zu vereinfachen und um Abwechslung in die angebotenen Alternativen zu bringen, sodass diese bei jedem Aufruf leicht anders sein können. Nutzer*innen können so durch ein „Neu-Laden“ des Recommenders stetig neue und spannende Empfehlungen gefunden werden. Mit einigen Modifikationen könnte eine Auswahl der Alternativen auf Grundlage eines vorher berechneten Rankings aber durchaus ebenso denkbar sein, wie ein Ansatz, der auf Kriterien-Sets beruht.

In dem innerhalb des Projektes umgesetzten PoC wurde eine große Bandbreite an POI mit aufgenommen, ohne allzu sehr auf die Spezifika einzugehen, die für eine sinnvolle Einbindung dieser POI in einen Recommender notwendig wären. Beispielsweise wurden Radrouten mit aufgenommen, ohne zu berücksichtigen, dass für eine ausgereifte Empfehlung von Radrouten sowohl der gewünschte Startpunkt des/der Anwender*in und eine gewünschte Länge bekannt sein sollten. Weiterhin sollten die Routen in Gänze ausgespielt werden, während der PoC nur einzelne Gebiete ausspielt. Solche Spezifika lassen sich für viele Gruppen von POI finden, wobei sich viele durch eine Einbindung von Öffnungszeiten beheben lassen. Trotzdem sollte in der produktiven Einführung eines Recommenders abgewägt werden, welche POI-Typen in den Recommender aufgenommen werden und diese dann mit allen notwendigen Funktionen implementiert werden. So können dann iterativ neue Gruppen von POI ergänzt werden.

Eine Komponente, die in unserem PoC des Recommenders aufgrund einer mangelnden Datengrundlage keine Berücksichtigung gefunden hat, ist die der Events. In einem Recommender könnte es dennoch Sinn machen, einzelne Events wie Konzerte als Alternativen anzubieten. Allerdings muss dafür von der Tourismuspraxis vorher noch klar definiert werden, welche Events ausgespielt werden sollten und ob dies abhängig ist von dem Ursprungs-POI, den ein/e User*in ausgewählt hat.

Um diesen Absatz zu schließen, soll ein letztes Learning erwähnt werden, das sich auf den Radius bezieht, innerhalb dessen nach Alternativen gesucht wird. Dieser ist im PoC abhängig vom Vorhersagehorizont und beträgt für null, respektive ein, zwei oder drei Tage in die Zukunft 10, 15, 20 und 25km. In der Datengrundlage des PoC existieren allerdings Regionen, gerade im ländlichen Raum, innerhalb derer im Umkreis von 10km keine weiteren POI verzeichnet sind. Die Anzahl dieser Regionen sollte natürlicherweise zurückgehen, wenn die Datengrundlage des Recommenders größer wird. Um trotzdem präventiv dagegen vorzugehen, könnte einem/r Nutzer*in nach jeder Anfrage im Frontend die Möglichkeit gegeben werden, den Suchradius um 5km zu vergrößern und die Anfrage mit dem erhöhten Radius erneut durchzuführen. So ist sichergestellt, dass ein/e Nutzer*in für jeden POI die Möglichkeit hat, im Frontend alternative POI zu finden.

5.3 Zwischenfazit

In AP 4 „Recommender“ des Projektes LAB-TOUR SH konnten über die Projektlaufzeit diverse Erkenntnisse gewonnen werden, die grundlegend für die tatsächliche produktive Nutzung eines touristischen Recommenders in Schleswig-Holstein sind. Dabei wurde ein Status Quo etabliert, anhand dessen für diverse Komponenten eines Recommenders überprüft wurde, welche Ansätze erfolgversprechend sind und wie viel weiterer Aufwand in diesen Komponenten fließen müsste, um sie zu einer produktiven Ausspielung fortzuführen. Neben der theoretischen Betrachtung des Recommenders wurden di-

verse analytische Experimente gemacht und über den Projektrahmen hinaus ein PoC des Recommenders, inklusive eines durchdachten Web-Frontends entwickelt. Anhand dieses Frontends lassen sich die Anforderungen der schleswig-holsteinischen Tourismuspraxis nun konkret diskutieren, sodass auf der Grundlage der Arbeit von LAB-TOUR SH ein deutlich klareres Bild eines digitalen Besuchermanagementsystems gezeichnet werden kann. In dem PoC, der über die letzten Abschnittskomponentenweise vorgestellt wurde, sind die folgenden Elemente eines digitalen Besuchermanagementsystems prototypisch integriert.

1. Automatisierte, kontinuierliche Abfrage von Auslastungsdaten aus dem Datenhub von ADDIX, Luft-hansa Industry Solutions und Eco Counter mit einer Frequenz von sechzig Minuten. Dadurch wird sowohl die historische Auslastung aller Sensorstandorte als auch deren Live-Auslastung in die Integrationsplattform aufgenommen und dort für den Recommender vorgehalten.
2. Anreicherung der Auslastungsdaten mit ortsabhängigen Kontextinformationen wie Wetterdaten, Ferien- und Feiertagen sowie zeitlichen Spezifikationen, wie Wochentag und Uhrzeit. Außerdem wird eine Transformation der gemessenen Ins- und Outs zu einer prozentualen Auslastung durchgeführt und diverse Korrekturen auf dieser durchgeführt, um auf Messungenauigkeiten zu reagieren.
3. Durchführung einer qualitativ hochwertigen, automatisierten Auslastungsprognose mit einer Frequenz von sechzig Minuten über die Nutzung der historischen Auslastungen sowie der Live-Auslastung unter Zuhilfenahme von ortsbezogenen, aktuellen Wettervorhersagen. Das Prognosemodul wurde dafür so konzeptioniert, dass es robust gegen kurzfristige Sensorausfälle und -verzögerungen ist, sich auf verschiedene Dynamiken einstellen kann, für neu angeschlossenen Sensorstandorten schnell Prognosen liefert und automatisiert und regelmäßig die gesammelten Daten nutzt, um seine Vorhersagequalität zu verbessern.
4. Aufsetzen eines Meta-Knowledge-Graphen, in dem die unter Open-Data-Lizenz vorliegenden, teilweise bereinigten POI der TA.SH mit allen verfügbaren Metadaten sowie Informationen über ihre Ähnlichkeit und Entfernung untereinander vorliegen.
5. Implementierung eines vielseitigen Algorithmus, um alternative Attraktionen für alle POI vorzuschlagen, abhängig von der Live-Auslastung der POI und unter Berücksichtigung von Entfernung, Ähnlichkeit und historischer Auslastung der POI. Über diesen Algorithmus wurde eine Lenkungswirkung in den Recommender integriert, die weg von akut oder historisch überlasteten POI hin zu bisher weniger besuchten Attraktionen führen soll. Außerdem wird über die gezielte Sortierung und Empfehlung von verfügbaren Parkplätzen die Parkplatzsuche vereinfacht.
6. Puffern der POI-Metadaten, Prognosen und Recommendations auf einem FIWARE Context Broker, sodass diese mit sehr geringer Latenz verfügbar und immer auf dem aktuellen Stand sind.
7. Implementierung einer Schnittstelle für den FIWARE Context Broker, sodass alle verfügbaren Daten in dokumentierten Standards nach außen verfügbar sind.
8. Entwicklung eines ansprechenden Frontends, das die auf dem Context Broker liegenden Daten sichtbar macht und das Nutzungspotential des Recommenders demonstriert. Anhand dieses Frontends kann nun klar diskutiert werden, welche weiteren Funktionalitäten gebraucht werden, um ein solches digitales Besuchermanagementsystem in der Schleswig-Holsteinischen Tourismuspraxis zu etablieren.
9. Dokumentation der wichtigsten, während der Implementierung getroffenen Entscheidungen, eine ausführliche Formulierung von Anforderungen und theoretischen Hintergründen zum Recommender Modul und eine Erläuterung der Ergebnisse in diesem Bericht.

Bei der Umsetzung dieser Komponenten wurde außerdem auf die in den jeweiligen Abschnitten ausführlich beschriebenen Anforderungen der Prognose und des Alternativengenerierers eingegangen.

Zudem konnte gezeigt werden, wie die in der Einleitung dieses Abschnitts genannten Ziele „Erhöhung der Ausflugsattraktivität“, „Reduktion von Crowding“ und „Reduktion von Parksucherverkehr“ über das vorgestellte Konzept erreicht werden könnten. Auf die in das Projekt aufgenommenen touristischen Use Cases konnte dank der flexiblen Logik des Recommenders individuell eingegangen werden.

In der Entwicklung eines touristischen Recommenders mit einem Fokus auf Nachhaltigkeit macht es zusätzlich Sinn, die Erreichbarkeit von dargestellten POI stärker mitzudenken. Der Recommender bietet dafür die Möglichkeit, interessierte Ausflügler sehr früh in ihrem Entscheidungsprozess adressieren zu können. Sollte es gelingen, den regionalen Nahverkehr als attraktive Möglichkeit zur Bewältigung des Fahrtweges zu einem POI zu platzieren, könnte dies durchaus Menschen dazu bewegen, auf die Anfahrt mit dem PKW zu verzichten und stattdessen den öffentlichen Nahverkehr zu nutzen. Besonders durch das Deutschlandticket scheint diese Option realistisch und sollte in einer produktiven Umsetzung mitgedacht werden.

6. Auspielung

Ziel bei der Auspielung sollte es sein, dass die Informationen zum Besuchsaufkommen und zu den Besucherströmen an möglichst vielen Stellen von Gästen abgerufen werden können. Eine solche integrierte Kommunikation kann nur dann gelingen, wenn die erhobenen Daten konsistent und möglichst flächendeckend an Gäste ausgespielt werden, damit diese an möglichst allen relevanten digitalen Touchpoints vorliegen. Um dem gerecht zu werden, wurde eine Konzeptionslogik entwickelt, mit der diese Informationskanäle identifiziert und bewertet werden können. Dabei wurden im Arbeitspaket 6 die folgenden Fragen beantwortet:

1. Identifizierung von Sekundärdaten zur Relevanzabschätzung: Welche Informationskanäle sind überhaupt relevant?
2. Identifizierung geeigneter Informationskanäle: Welche von den als relevant erachteten Informationskanälen sind geeignet für die Auspielung eines Recommenders und warum?
3. Exemplarische Pilotierung: Wie kann es aussehen, wenn ein Recommender ausgespielt wird?
4. Weitere Pilotierung: In welche weiteren Auspielkanäle könnte ein Recommender nach Projektende noch angebunden werden?

6.1 Konzeptionslogik

Bevor ein Frontend für einen Recommender entwickelt wird ist es sinnvoll abzuschätzen, welche Relevanz verschiedene Auspielkanäle haben könnten, welche Endgeräte häufig genutzt werden und welche Art von Informationen Reisende aufrufen, damit der Recommender so konzipiert werden kann, dass das Frontend sich bestmöglich verbreiten und damit eine entsprechende Sichtbarkeit erlangt. Hierfür wurde eine Konzeptionslogik entwickelt entlang derer wichtige grundlegende Parameter für die Entwicklung des Frontends festgelegt werden können. Diese Vorgehensweise ist übertragbar auf andere Projekte und wird nachfolgend erläutert. Die Konzeptionslogik, mit der abgeleitet werden soll, wie ein Recommender gestaltet werden und sich verbreiten kann, besteht in diesem Fall aus den folgenden zwei Analyseschritten mit ihren jeweiligen Unterarbeitschritten:

1. Relevanzabschätzung der Nutzung:

- Es sollten aktuelle Studien zur Nutzung von **Endgeräten** erfolgen, um die Priorisierung der Darstellung (optimiert für Smartphone, Website etc.) festlegen zu können.

- Daneben ist es für die Gestaltung wichtig zu wissen, **welche Informationen** von Nutzenden in der Regel im Urlaub gesucht werden (Wetter, Routenführung etc.)
- Schließlich sollte recherchiert werden, welche **Arten von Content** aufgerufen werden, um den Recommender ausgestalten zu können (Verhältnis von Video, Text, Bildern usw.)

2. Die Analyse der **Informationskanäle** kann auf zwei Ebenen erfolgen:

- Analyse relevanter Websites und Anwendungen, wo ein Recommender implementiert werden könnte: Dies können Touren- oder Routenplaner, Kartendienste oder andere Destinationswebsites sowie Use Case spezifische Anwendungen sein. Hieraus kann die Relevanz und Bandbreite einer Implementierung hergeleitet und danach eine Priorisierung vorgenommen werden, zu welchen Anbietern für die Verbreitung der Recommenders Kontakt aufgenommen werden könnte.
- Analyse des Funktionsumfangs von ähnlichen und komplementären Anwendungen und Websites: Wenn klar ist, welche Angebote es bereits auf dem Markt gibt, die von Reisenden genutzt werden, dann kann von diesen abgeleitet werden, welcher Funktionsumfang sich auch auf einen Recommender übertragen lässt. Dies sind in der Regel dann diejenigen Dienste und Informationen, die Reisende vor Ort nutzen. Hierzu zählen zum Beispiel Wetterangaben, Navigation und Routenplanung sowie weitere spezifischere Anwendungen. Wenn der Recommender entlang dieser bereits bei Nutzenden bekannten Funktionen entwickelt wird, dann ist er intuitiver bedienbar und kann von den Reisenden leichter verstanden werden.

6.2 Relevanzabschätzung verschiedener Informationskanäle

In AP 5.1 werden zunächst die Quellen identifiziert, auf deren Basis die spätere Relevanzabschätzung erfolgt. In AP 5.1 geht es demnach zunächst darum, mögliche Informationsquellen von Touristen und Touristinnen zu identifizieren. Dabei wird die Nachfrageseite untersucht und nach (1) Art der Anwendung (Skill, Website etc.), (2) Art des Contents (Video, Audio usw.), (3) Art der Information (Wetter, Veranstaltungen, Routenführung usw.) und (4) Art der Endgeräte (Smartphone, Laptop, Tablet usw.) differenziert.

Dieser Logik folgend sollen die anschließenden Fragen beantwortet werden:

1. Für welches Endgerät werden Informationen ausgespielt?
2. Wo sollten Informationen ausgespielt werden?
3. Auf welchen Anwendungen sollten Informationen ausgespielt werden?

Der Fokus von AP 5.1 liegt also darauf, wo welche Endgeräte und mit welchen Anwendungen für welchen Zweck im Urlaub genutzt werden.

Endgeräte

In der Postbank Digitalstudie (Postbank, 2022) wurde unter anderem die Internetnutzung nach Endgeräten untersucht (vgl. Abb. 67). Datenbasis der Studie ist eine Stichprobe der Bundesbürger mit einer Größe von n=3.050. Die Befragten sollten angeben, welche Endgeräte sie für den Internetzugriff nutzen. Mit 84% gaben die meisten Befragten an, dass sie das Smartphone für den Internetzugriff nutzen, gefolgt vom Laptop/Notebook mit 69% und dem Desktop mit 52%. Über 40% nutzen zudem jeweils Tablets und Smart-TVs. Für den spezifischen Kontext der Nutzung von Endgeräten im Urlaub können hier bedingt Ableitungen getroffen werden.

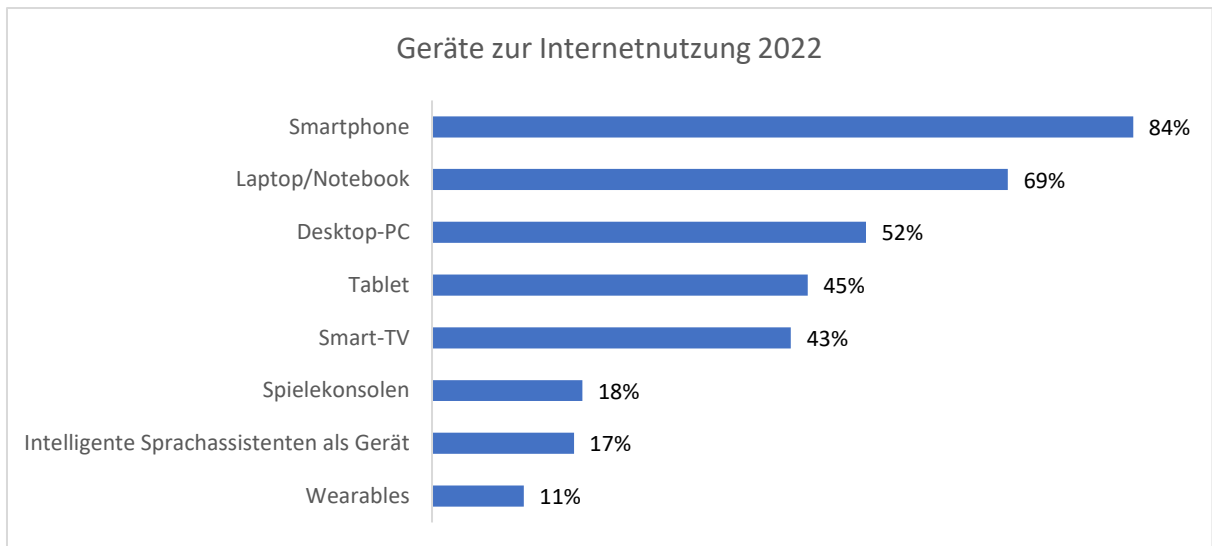


Abbildung 67: Geräte zur Internetnutzung 2022

Quelle: Eigene Darstellung basierend auf Postbank - eine Niederlassung der Deutsche Bank AG, 2022

Die Initiative D21 e.V. (Initiative D21 e. V., 2022) hat ebenfalls eine Studie zur Nutzung digitaler Geräte durchgeführt (siehe Anhang). Die Teilnehmenden sollten angeben, welche der aufgeführten Geräte sie beruflich oder privat nutzen. Datenbasis für den D21-Digital-Index ist eine Stichprobe von n=18.243 mit Personen ab 14 Jahren. Auch diese Studie belegt, dass die meisten Nutzer*innen per Smartphone auf das Internet zugreifen. Der prozentuale Anteil nimmt dabei ab, je älter die Nutzergruppe ist. Während bei den sogenannten Babyboomer*innen 87% das Smartphone nutzen, sind es bei der Nachkriegsgeneration 71% und bei der Generation bis 1945 nur noch 42%. Gefolgt wird das Smartphone vom Laptop mit 66%, dem Desktop-PC mit 47%, Smart-TV mit 46% und Tablet mit 41%.

In einer von DataReportal (2022) veröffentlichten Studie wurde nicht die Nutzung, sondern der Besitz von Endgeräten untersucht (vgl. Abb. 68). Datenbasis ist eine Untersuchung von GWI in Q3 2021. Teilnehmende sind Internetnutzer*innen zwischen 16 und 64 Jahren. In Deutschland besitzen mittlerweile 95,6% eine Art von Mobiltelefon, wobei es sich bei 95,1% um Smartphones handelt. Laptops und Desktop Computer werden von 82,8% besessen und Tablets von 52,9%.

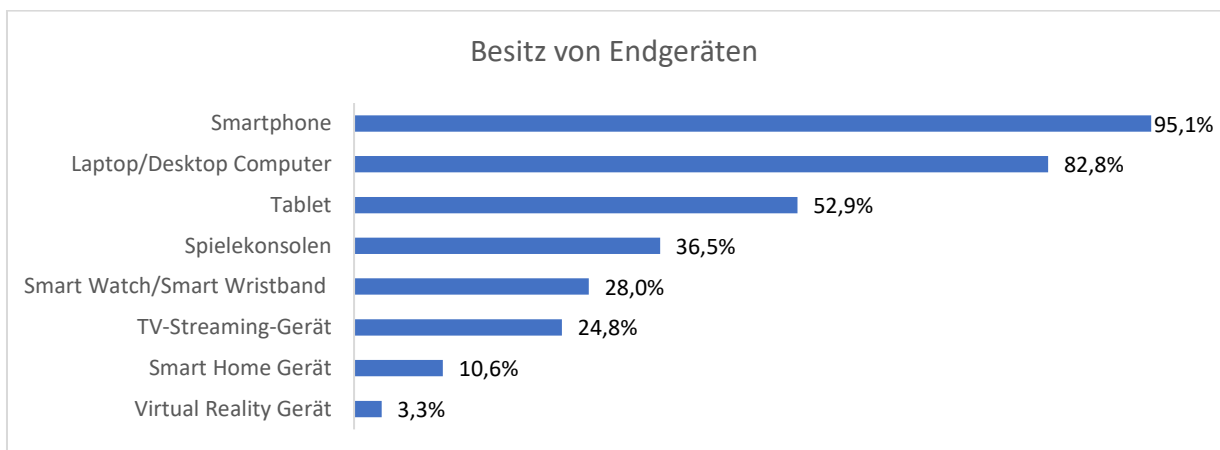


Abbildung 68: Besitz von Endgeräten

Quelle: Eigene Darstellung basierend auf DataReportal, 2022

In der Studie von DataReportal (DataReportal, 2022) wurde ebenfalls der Zugriff auf das Internet nach Endgeräten untersucht (siehe Anhang). Auch diese Studie ergab, dass die meisten Nutzer*innen per Smartphone auf das Internet zugreifen (87%), gefolgt vom Laptop oder Desktop (84,5%) sowie dem Tablet (45,6%) und vernetztem Fernsehen (37,4%).

In einer Studie von StatCounter (2022) wurde der Marktanteil von Endgeräten untersucht. Die Statistik in Abbildung 69 basiert auf aggregierten Daten, die für eine Stichprobe mit mehr als 5 Milliarden Seitenaufrufen pro Monat aus dem StatCounter-Netzwerk für den Zeitraum Januar 2022 bis Dezember 2022 gesammelt wurden. Den größten Anteil haben Laptops und Desktop Computer (50,13%), gefolgt von Mobiltelefonen (47,15%). Tablets haben einen geringen (2,66%) und Konsolen mit 0,06% so gut wie keinen Anteil.

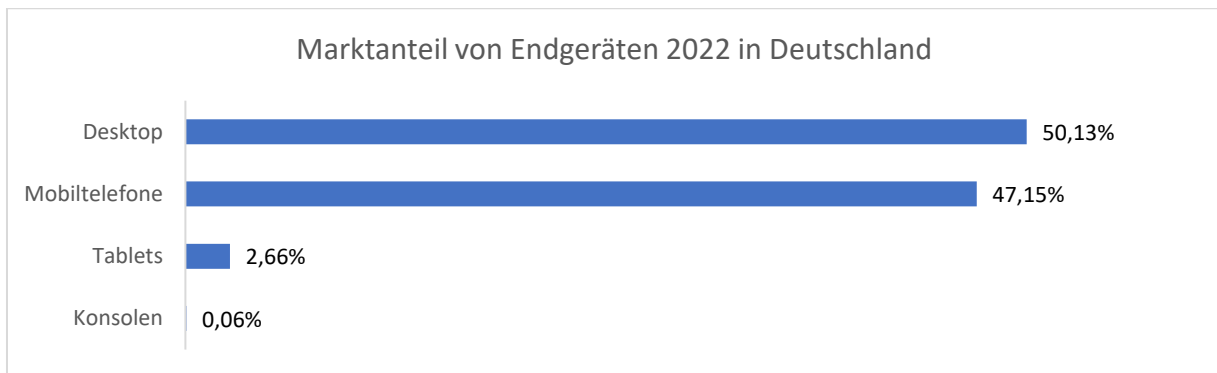


Abbildung 69: Marktanteil von Endgeräten 2022 in Deutschland

Quelle: Eigene Darstellung basierend auf StatCounter, 2022

In der ARD/ZDF-Onlinestudie (ARD/ZDF-Forschungskommission, 2021) wurde die Tagesreichweite von Endgeräten nach Alter untersucht (vgl. Abb. 70). Datenbasis ist eine Stichprobe der deutschsprachigen Wohnbevölkerung ab 14 Jahren von n=2.001. Die Ergebnisse der Studie zeigen, dass insgesamt die Tagesreichweite von Fernsehern/Smart TVs am größten ist, gefolgt von UKW/DAB+-Radiogeräten und Smartphones. Eine wesentlich geringere Tagesreichweite haben Laptops, stationäre Computer, Tablets und WLAN-Radios. Wird nach Altersklassen differenziert, fallen die Ergebnisse anders aus. Bei den 14-29-Jährigen hat das Smartphone mit Abstand die größte Tagesreichweite. Bei der Altersgruppe 30-49-Jahre hat ebenfalls das Smartphone die größte Tagesreichweite, wenn auch nur geringfügig.

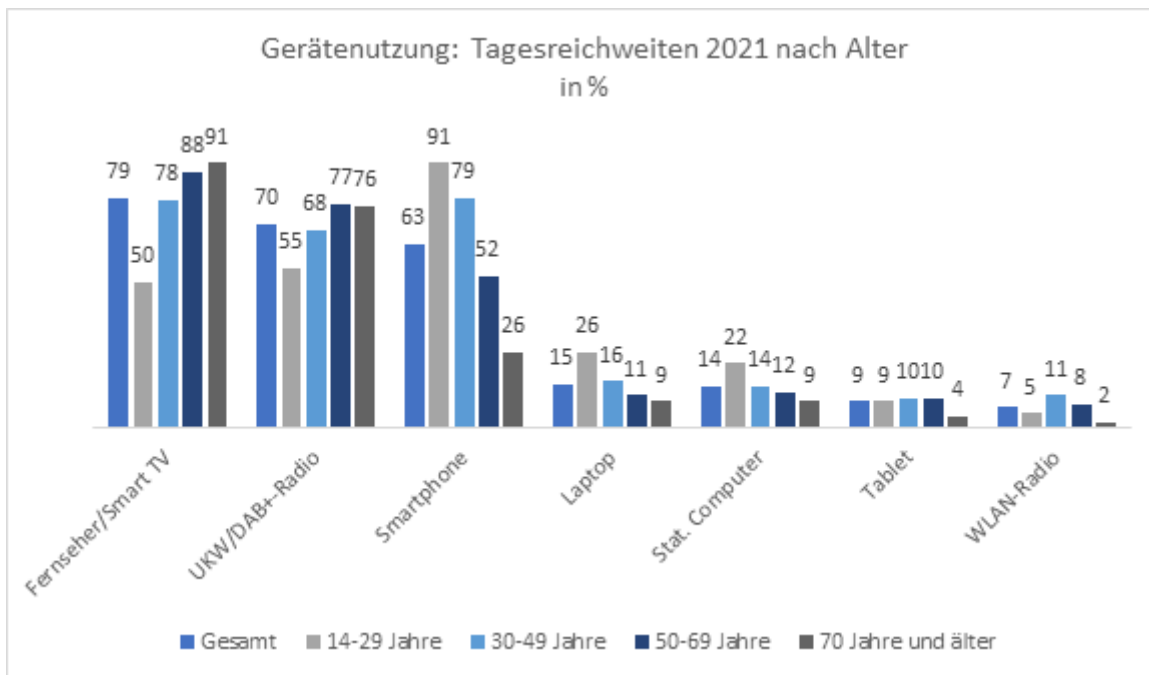


Abbildung 70: Gerätenutzung: Tagesreichweiten 2021 nach Alter

Quelle: Eigene Darstellung basierend auf ARD/ZDF-Forschungskommission, 2021

Bitkom Research (Bitkom Research, 2022) hat die Nutzung von Geräten auf Urlaubsreisen untersucht. Dazu wurden 935 Urlaubsreisende aus Deutschland ab 16 Jahren gefragt, welche der gelisteten Geräte sie zumindest hin und wieder auf Urlaubsreisen verwenden. Abbildung 71 verdeutlicht, dass 79% der Befragten angaben, ihr Smartphone während des Urlaubs zu nutzen. Am zweithäufigsten wird das Tablet genutzt, von 41% der Befragten, gefolgt vom Laptop mit 34%.

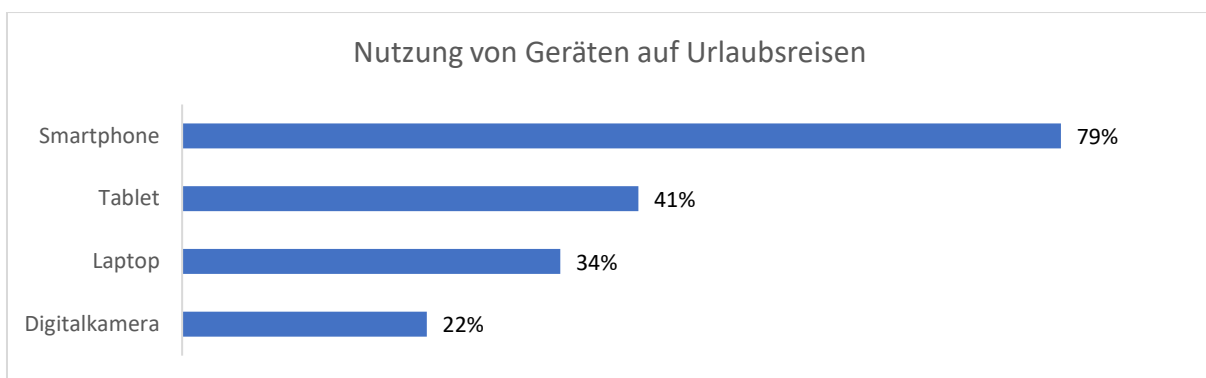


Abbildung 71: Nutzung von Geräten auf Urlaubsreisen

Quelle: Eigene Darstellung basierend auf Bitkom Research, 2022

Implikationen für die Relevanzabschätzung

Deutlich wird bei der Postbank Digitalstudie, dass die Verbreitung des Smartphones fast flächendeckend ist (mit 84%) und damit angenommen werden kann, dass die Nutzung auch im Urlaub erfolgt. Gestützt wird diese Einschätzung durch die Studie von DataReportal, dass auch der Besitz eines Smart-

phones mittlerweile bei 95,1% liegt sowie die Bitkom Research Studie (79%). Gleiches gilt für die Ausgabe auf Laptop, der in verschiedenen Studien eine hohe Verbreitung zwischen 69% (Postbank Digitalstudie) und 82,8% (DataReportal) besitzt (wobei die Bitkom Research Studie mit einer Angabe von 34% hier etwas abweichende Ergebnisse liefert). Dass sich die Nutzung auf Desktop (50,13%) und Mobile (47,15%) gleichermaßen aufteilt verdeutlicht die StatCounter Erhebung, wobei im Kontext des Urlaubs anzunehmen ist, dass Mobile aufgrund der Flexibilität der Nutzung überwiegt. Diese Relevanz kann bedingt auch dem Tablet zugesprochen werden (41% bei D21 und Bitkom Research, 45% bei der Postbank Digitalstudie und 52,9% bei DataReportal). Gleichzeitig ist die Verbreitung von Wearables deutlich geringer ausgeprägt (mit 11% bei der Postbank Digitalstudie und 28% bei DataReportal), sodass eine Auspielung auf diese Art von Endgeräten vermutlich (noch) vernachlässigbar ist, oder zumindest nicht dieselbe Relevanz besitzt. Auch eine Auspielung des Recommenders in Form einer Sprachausgabe scheint nicht zwingend zu sein, da Nutzungsgewohnheiten allgemein noch nicht in einem Maße (lediglich 17% bei der Postbank Digitalstudie) verbreitet sind, die eine Überführung in ein solches Format rechtfertigen würden. Gleiches gilt für Virtual Reality Geräte (3,3% bei DataReportal). Beachtenswert ist die Nutzung von Smart-TV (43% bei der Postbank Digitalstudie sowie 46% bei D21 und 24,8% bei DataReportal (wobei es hier um den Besitz und nicht die Nutzung geht)), da diese in Hotels oftmals verfügbar sind. Die Relevanz wird durch die ARD/ZDF Online-Studie noch einmal hervorgehoben, in der eine Tagesreichweite von insgesamt 79% konstatiert wird – und hier entgegen der Smartphone-Nutzung die Reichweite mit zunehmendem Alter ansteigt. Wenn Reisende diese auch im Urlaub verstärkt, nutzen sollten wäre es hier denkbar, einen Startbildschirm zu etablieren, wo aktuelle Besuchsaufkommen und deren Ströme abgebildet werden. Die Studie D21 stützt diese Einschätzung und zeigt, dass die Verbreitung des Smartphones insbesondere bei den Babyboomer*innen sehr hoch ist.

Die Ergebnisse zeigen, dass die Ausgabe auf dem Smartphone unabdingbar ist. Ähnliches gilt für den Laptop – wenn auch eingeschränkt und es bleibt abzuwarten, wie sich die Nutzung hier weiterentwickelt und wie diese im Urlaubskontext aussieht, wozu keine aktuellen Daten vorliegen. Ausgabekanäle wie Wearables, Sprachassistenten oder Virtual Reality können bis dato vernachlässigt werden. Die Ausgabe auf Smart-TV und – da ein ähnliches Format – auf Infostelen, sollten jedoch mit berücksichtigt werden.

Nutzung von Medienangeboten im Allgemeinen

DataReportal (DataReportal, 2022) hat zudem die tägliche Nutzungsdauer von Medien betrachtet (vgl. Abb. 72). Datenbasis ist die in Q3 2021 durchgeführte Studie von GlobalWebIndex (GWI). Im Jahr 2021 sank die tägliche Zeit, die im Internet verbracht wurde, im Vergleich zum Vorjahr, liegt jedoch immer noch bei 5h 22min pro Tag. Fernsehen schauen oder das in Anspruch nehmen von Streamingdiensten hat mit 3h 20min die zweithöchste tägliche Nutzungsdauer. Es ist jedoch auch hier ein minimaler Rückgang zu verzeichnen. Ebenso wurde im Schnitt eine Minute weniger Radio gehört, womit die tägliche Nutzungsdauer nun bei 1h 30min liegt. Gestiegen ist jedoch sowohl die Zeit, die mit der Nutzung von Social Media verbracht wurde, auf 1h 29min, als auch die des Lesens von Presse Medien, digital oder physisch, auf 1h 19min und die des Hörens von Musik über Streaming Dienste, auf 1h 07min. Einen deutlichen Anstieg von über 25% verzeichneten Podcasts. Hier beträgt die tägliche Nutzungsdauer nun 34min. Die tägliche Nutzungsdauer von Spielekonsolen bleibt mit 47min unverändert.

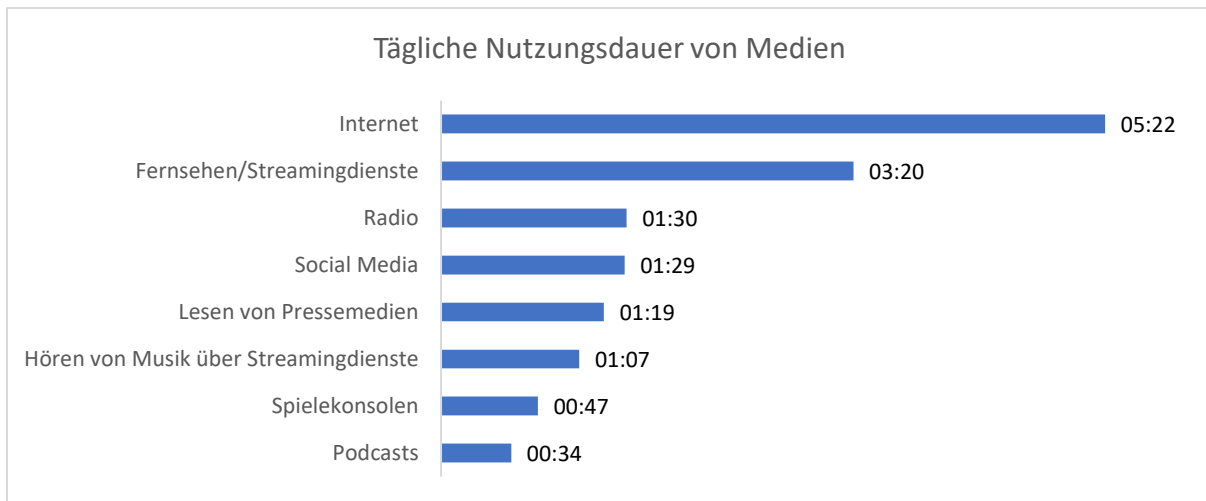


Abbildung 72: Tägliche Nutzungsdauer von Medien

Quelle: Eigene Darstellung basierend auf DataReportal, 2022

In der auf Daten von GWI basierenden Studie (DataReportal, 2022) wurden ebenfalls die Hauptgründe für die Internetnutzung in Deutschland untersucht (siehe Anhang). Zu den Hauptgründen für die Internetnutzung zählen vor allem Informations- und Recherchezwecke. Am häufigsten wird das Internet genutzt, um Informationen zu finden, Neuigkeiten und aktuelle Events zu verfolgen sowie Marken zu recherchieren. Des Weiteren wird das Internet zum Entertainment und zur Freizeitgestaltung genutzt sowie für soziale Aspekte wie das in Kontakt bleiben mit Freunden und Familie. Gut die Hälfte der Befragten nutzt das Internet zudem für die Suche nach Orten und Reisen sowie nach neuen Ideen und Inspiration.

Im Rahmen der Global Consumer Survey 2022 (Statista Consumer Insights, 2023) hat Statista die Smartphone-App-Nutzung in Deutschland untersucht (siehe Anhang). Die Daten wurden in vier Wellen zwischen Januar 2022 und Dezember 2022 erhoben. Teilnehmende waren 5.432 Deutsche zwischen 18 und 64 Jahren. Befragt wurden Personen, die ein Smartphone nutzen. Die Studie zeigt, dass die meisten Deutschen Messenger Apps (60%) und Wetter-Apps (60%) nutzen, dicht gefolgt von E-Mail-Diensten (59%), Internetbrowsern (58%) und Social-Media-Apps (56%). Navigations- und Verkehrsanbindungs-Apps werden von 40% der Befragten genutzt und Nachrichtenangebote von 31%.

Implikationen für die Relevanzabschätzung

Die Art der Nutzung stützt die zuvor identifizierten relevanten Endgeräte Smartphone, Laptop und Smart-TV, da in der DataReportal Studie die Internetnutzung mit 5:22 Stunden überwiegt und daran Fernsehen bzw. Streamingdienste mit 3:20 Stunden anschließen. Die Gründe für die Nutzung zeigen, dass der Recommender eine entsprechende Relevanz erfährt, da ebenfalls laut der Studie von DataReportal die Nutzenden sich über das Internet vornehmlich informieren und Recherche betreiben. Da Neuigkeiten entsprechend wichtig sind, geht die Verfügbarmachung von Live-Daten zum Besuchsaufkommen hiermit konform. Das Global Consumer Survey von 2022 zeigt zudem, dass für den Recommender Wetter-Angaben einen wichtigen Grund liefern könnten, um diesen anzurufen, da 60% aktiv nach diesen Informationen suchen. Mit immerhin noch 40% werden Navigations- und Verkehrsanbin-

dungs-Apps genutzt, wo der Recommender ebenfalls Aufschluss geben kann – zumindest in der Vorplanung, wenn es darum geht, wohin überhaupt navigiert werden soll und wie die Verbindung dorthin ist. Es kann somit festgehalten werden, dass Wetterinformationen in einen Recommender integriert werden sollten und es auch Weiterleitungen oder Anbindungen an Navigations-Apps geben sollte und wenn möglich bereits die Verkehrsverbindungsmöglichkeiten angezeigt werden.

Art des Contents

In der ARD/ZDF-Onlinestudie (ARD/ZDF-Forschungskommission, 2021) wurde die Tagesreichweite von Video, Audio und Text im Internet untersucht (vgl. Abb. 73). Die Tagesreichweite von Videos im Internet ist am größten. Audio besitzt die zweitgrößte Tagesreichweite und Text die geringste. Die Tagesreichweiten von Video und Audio Content verzeichneten seit 2018 einen kontinuierlichen Anstieg. Die Nutzung von Text im Internet stagniert hingegen weitestgehend.

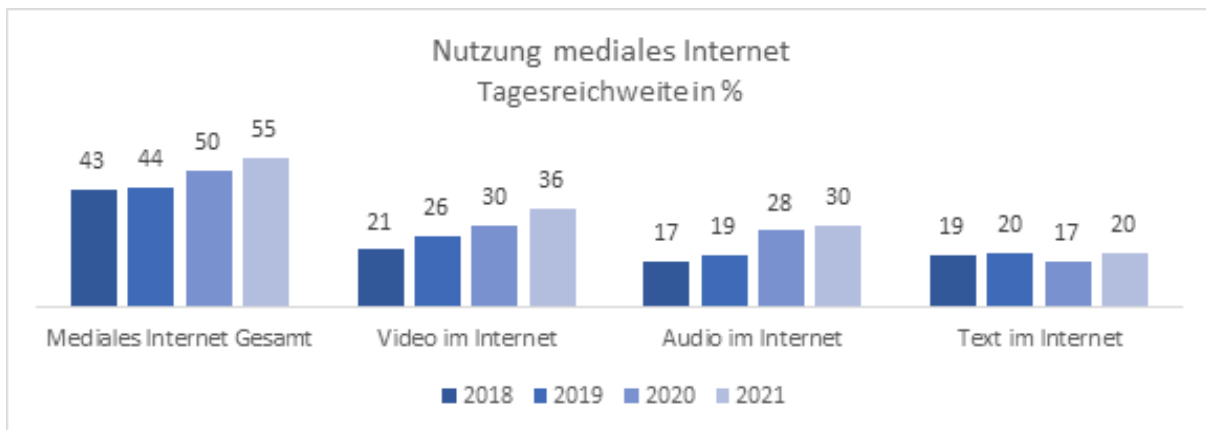


Abbildung 73: Nutzung mediales Internet

Quelle: Eigene Darstellung basierend auf ARD/ZDF-Forschungskommission, 2021

Die Nutzung und Tagesreichweite der unterschiedlichen Medien hängt ebenfalls von dem Alter der Nutzer*innen ab. Abbildung 74 zeigt, dass das mediale Internet insgesamt am meisten von 14-29-Jährigen genutzt wird und dieser Anteil mit steigender Altersgruppe abnimmt. Bei den 14-29-Jährigen und 30-49-Jährigen hat zudem Video Content die größte, Audio Content die zweitgrößte und Text die geringste Tagesreichweite. Bei den 50-69-Jährigen sind keine signifikanten Unterschiede zu erkennen. Die Altersgruppe 70 Jahre und älter nutzt hingegen am meisten Text.

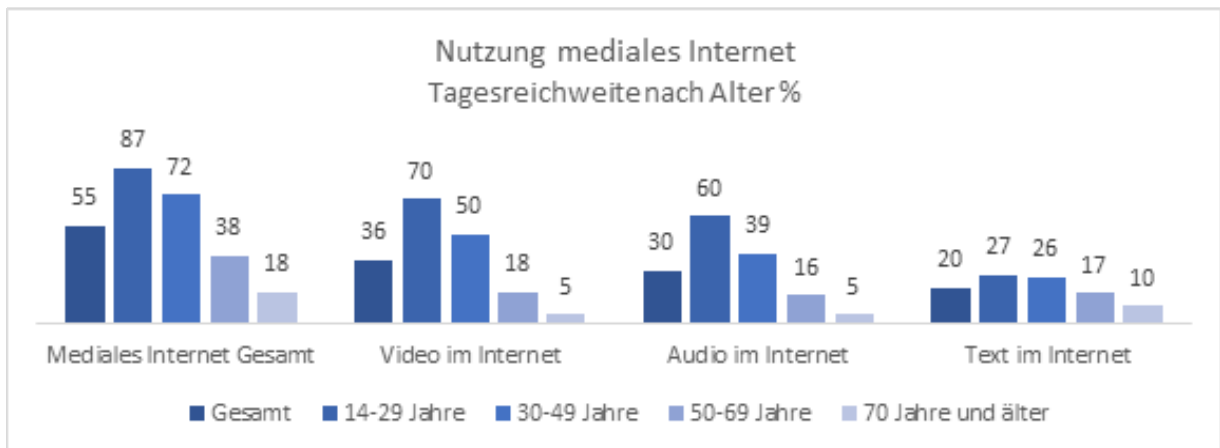


Abbildung 74: Nutzung mediales Internet nach Alter

Quelle: Eigene Darstellung basierend auf ARD/ZDF-Forschungskommission, 2021

In der ARD/ZDF-Onlinestudie (ARD/ZDF-Forschungskommission, 2021) wurde ebenfalls die Nutzungsdauer in Minuten pro Tag für das mediale Internet untersucht (vgl. Abb. 75). Die Nutzungsdauer für das gesamte mediale Internet ist 2020 und 2021 deutlich angestiegen. Die Nutzungsdauer von Videocontent im Internet ist am größten, für Audio am zweitgrößten und für Text am geringsten. Sowohl die Nutzungsdauer von Video- als auch Audiocontent stieg in den Jahren 2018 bis 2021 kontinuierlich an, während die Nutzungsdauer von Text im Internet bis zu einem leichten Anstieg in 2021 sank.

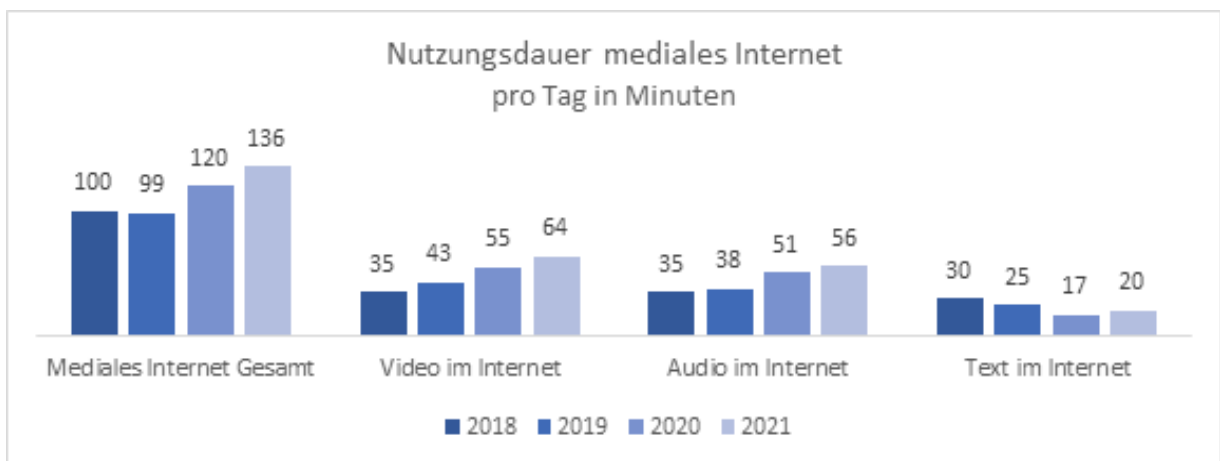


Abbildung 75: Nutzungsdauer mediales Internet

Quelle: Eigene Darstellung basierend auf ARD/ZDF-Forschungskommission, 2021

Auch bei der Nutzungsdauer der Medien sind je nach Altersgruppe starke Unterschiede zu verzeichnen, wie in Abbildung 76 deutlich wird. Die Nutzungsdauer nimmt mit steigender Altersgruppe ab. Die Altersgruppen 14-29 Jahre und 30-49 Jahre nutzt pro Tag am längsten Video, gefolgt von Audio und am kürzesten Text. Bei den Altersgruppen 50-69 Jahre und 70 Jahre und älter sind die Unterschiede zwischen der Nutzungsdauer der Medien geringer, jedoch nutzen die 50-69-Jährigen pro Tag am meisten Audiocontent und die 70+ am meisten Text im Internet.

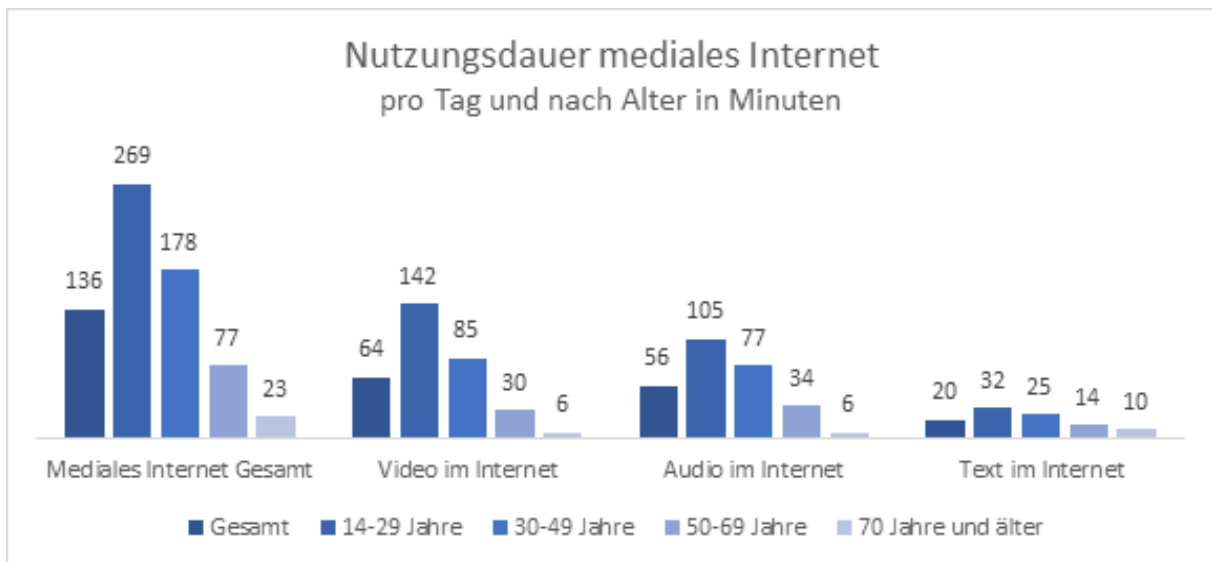


Abbildung 76: Nutzungsdauer mediales Internet nach Alter

Quelle: Eigene Darstellung basierend auf ARD/ZDF-Forschungskommission, 2021

Implikationen für die Relevanzabschätzung

Die Ergebnisse der ARD/ZDF-Onlinestudie legen nahe, dass die Vermittlung von Besuchsaufkommen und Besucherströmen multimedial und insbesondere via Video (36% Tagesreichweite in 2021 im Vergleich zu 20% von reinem Text) erfolgen sollte. Selbst wenn das Verhalten je nach Alter der Nutzenden stark schwankt (bei den 14–29-jährigen 70% Video und bei den 50–69-Jährigen lediglich 18%), ist eine klare Präferenz zum Format Video zu erkennen, wo insgesamt die Nutzungsdauer auch am höchsten ist mit 64 Minuten pro Tag (Daten ebenfalls auf Basis der ARD/ZDF-Onlinestudie). Die Übertragung von Videos ist datenschutzrechtlich allerdings schwierig. Eine Möglichkeit wäre es, wenn es Webcams gibt, die bereits installiert sind und die in den Recommender übernommen werden, damit Gäste ein Live-Bild vom Standort erhalten – was zum Beispiel an Standorten wie Stränden oftmals möglich ist.

Nutzung von Medien- und Informationsangeboten für den Urlaub

Bitkom Research (Bitkom Research, 2020) hat in einer Studie 1.003 Personen ab 16 Jahren gefragt, wo sie sich in der Regel informieren, wenn sie eine Reise planen (vgl. Abb. 77). Insgesamt gaben 68% an, dass sie sich online informieren. Dabei stellen Reise- und Vergleichsportale die wichtigsten Informationskanäle dar, dicht gefolgt von den online Informationsmöglichkeiten der jeweiligen Dienstleister. Einen geringen Anteil haben ebenfalls Reiseblogs, soziale Netzwerke und Podcasts.

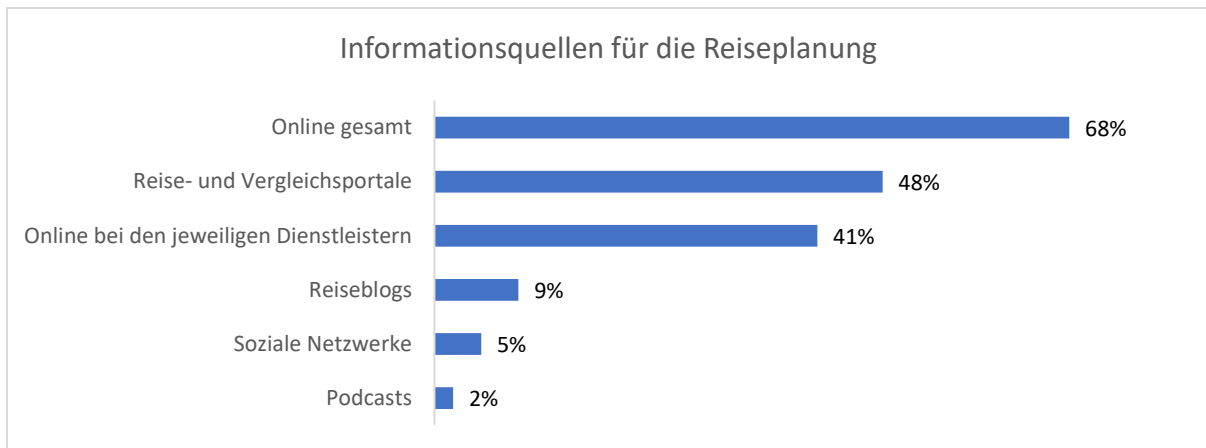


Abbildung 77: Informationsquellen für die Reiseplanung

Quelle: Eigene Darstellung basierend auf Bitkom Research, 2020

Eine weitere Untersuchung von Bitkom Research (Bitkom e.V., 2020b) ergab zudem, dass die deutliche Mehrheit ihr Smartphone mit in den Urlaub nimmt (siehe Anhang). Datenbasis ist ebenfalls eine Stichprobe von 1.003 Personen in Deutschland ab 16 Jahren, darunter 844 Internetnutzer*innen. Die Teilnehmenden wurden gefragt, welche der gelisteten Apps sie bereits im Urlaub genutzt haben. Die Studie ergab, dass im Urlaub mit Abstand am häufigsten Wetter-Apps (61%) genutzt werden, gefolgt von Übersetzungs-Apps (40%) und Navigations-Apps (34%). Weniger als ein Viertel nutzt im Urlaub ebenfalls Reiseführer-Apps (20%) und ein geringer Anteil Buchungs-Apps für Unterkünfte (12%) oder Transportmittel (10%).

Zwei Jahre später führte Bitkom Research erneut eine Studie mit 760 Urlaubsreisenden, die ihr Smartphone oder Tablet im Urlaub nutzen, durch (Bitkom Research, 2022). Die Teilnehmenden wurden gefragt, welche der gelisteten Apps sie zumindest hin und wieder auch auf ihren Urlaubsreisen nutzen. Abbildung 78 zeigt, dass in den Ergebnissen nun eindeutig der Einfluss der COVID-19 Pandemie erkennbar ist. Die meisten Befragten nutzen nun Corona-Apps. Wetter-Apps folgen auf Platz zwei. Am drittmeisten werden Navigations-Apps genutzt, gefolgt von Übersetzungs-Apps. Navigations-Apps haben demnach im Vergleich zur vorherigen Studie an Bedeutung gewonnen. Von den gelisteten Möglichkeiten nutzen die wenigsten Reiseführer-Apps.

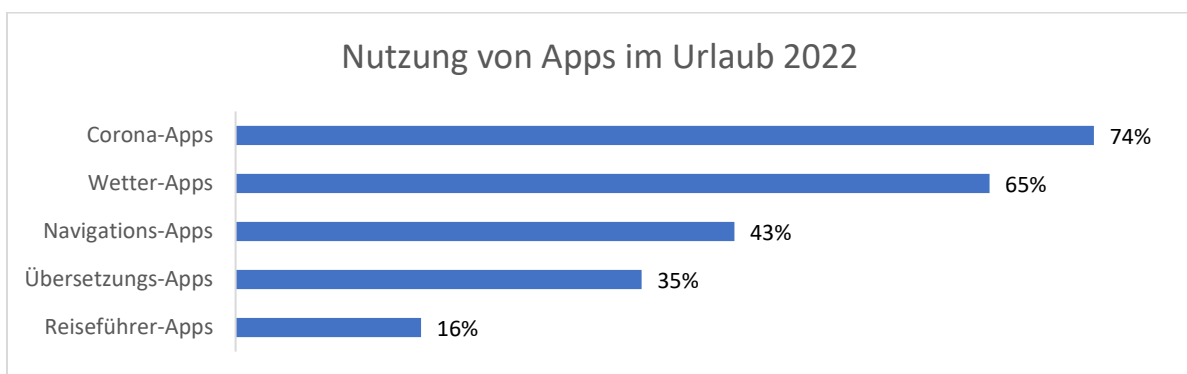


Abbildung 78: Nutzung von Apps im Urlaub 2022

Quelle: Eigene Darstellung basierend auf Bitkom Research, 2022

Bitkom Research (Bitkom e.V., 2020a) hat zudem konkret die Nutzung von Medien für die Information über Staus untersucht. In der Studie wurden 793 Auto- und Motorradfahrer*innen gefragt, wie sie sich über Stau informieren. Mit 91% nutzt die deutliche Mehrheit das Radio als Informationsquelle. Die am zweithäufigsten genutzte Quelle ist das Navigationsgerät (61%), gefolgt von Staumelder-Apps (20%) und Navigations-Apps (16%). Ein geringer Anteil nutzt ebenfalls die online Informationsangebote von Automobilclubs (6%) und Verkehrs-Websites (5%) sowie SMS-Services (4%).

Bitkom Research (Bitkom Research, 2019) hat ebenfalls die Potenziale weiterer digitaler Technologien auf Reisen untersucht. 1.004 Verbraucher ab 16 Jahren wurden gefragt, welche der genannten Services sie nutzen würden. 60% der Befragten gaben an, dass sie Augmented Reality mit Hilfe von Smartphone-Apps nutzen würden. Gut die Hälfte der Befragten gab zudem an, dass sie Virtual Reality nutzen würde, um virtuelle Reisen an existierende oder fiktive Orte zu unternehmen oder um sich nach dem Urlaub mit einer VR-Brille 360-Grad-Urlaubsvideos anzuschauen.

Auch der Bundesverband Digitale Wirtschaft (BVDW) e.V. (Bundesverband Digitale Wirtschaft (BVDW) e.V., 2018) hat sich mit der Smartphone Nutzung im Urlaub befasst. Im Auftrag des BVDW hat Kantar TNS 855 Personen gefragt, wozu sie ihr Smartphone bei einer Urlaubsreise nutzen. Die Studienergebnisse in Abbildung 79 zeigen, dass die meisten Personen ihr Smartphone im Urlaub nutzen, um zu telefonieren, E-Mails zu lesen/versenden oder für Messaging. Demnach wird das Smartphone zumeist für soziale Aktivitäten genutzt. 37% der Befragten nutzen es jedoch auch zur Navigation, 19%, um Reiseelemente für den laufenden Urlaub zu buchen und 8% nutzen Bewertungsportale wie Tripadvisor oder Yelp.

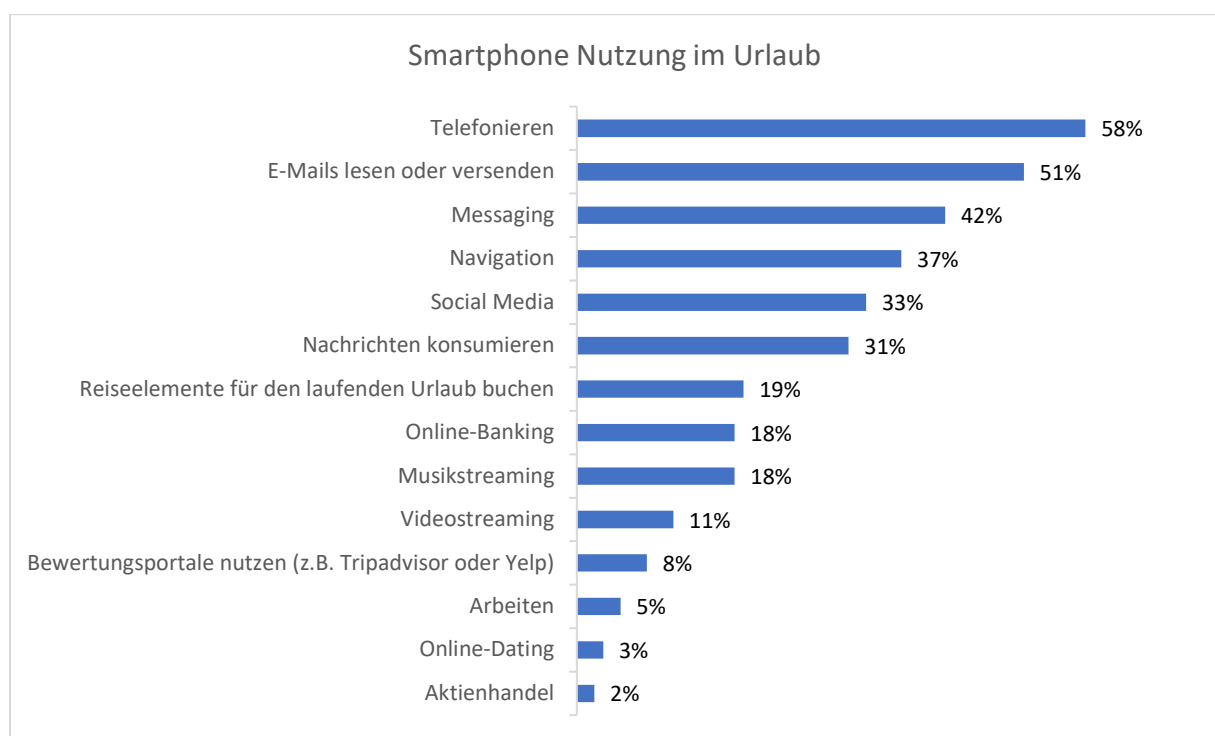


Abbildung 79: Smartphone Nutzung im Urlaub.

Quelle: Eigene Darstellung basierend auf Bundesverband Digitale Wirtschaft (BVDW) e.V., 2018

Implikationen für die Relevanzabschätzung

Für die Reiseentscheidung zeigt die Bitkom Research Studie aus 2020, dass Interessierte neben großen Reise- und Vergleichsportalen, wo eine Ausspielung aktuell unrealistisch erscheint, insbesondere die jeweiligen Dienstleister nutzen, bei denen sie buchen möchten. Hier wäre eine Ausspielung via Widget (das sind kleine Mini-Websites, die sequenziell über ein Code-Snippet in eine andere Website integriert werden können) denkbar. Bezüglich des Endgerätes und der Art der Apps, die im Urlaub genutzt werden, wird die Interpretation, die weiter oben auf Basis der allgemeinen Nutzungsdaten getroffen wurde, durch die Bitkom Studie aus 2020 gestützt. Diese bestätigte sich in 2022 mit der Ausnahme, dass dort die Corona-App noch sehr stark im Einsatz war. Es kann angenommen werden, dass aufgrund dieser Sensibilität auch Angaben zum Füllstand bestimmter Ausflugsziele hier anschlussfähig wären. Interessant ist in diesem Zusammenhang auch die Bitkom e.V. Studie aus 2020, bei der erneut deutlich wird, wie wichtig Navigationsgeräte und Navigations-Apps wären (Bitkom e.V, 2020a). Hier dürfte ein Recommender aktuell jedoch nur schwerlich integriert werden können, da der Kontakt zu globalen Playern wie Google oder Apple eher schwierig ist und diese einen fremden Algorithmus höchstwahrscheinlich nicht akzeptieren würden. Potenzial besteht jedoch bei der Integration eines Recommenders in einer Reiseführer-App wie zum Beispiel outdooractive. Die Relevanz, die ein Recommender in einem solchen Rahmen haben könnte, wird durch die BVDW Studie von 2018 unterstützt. Selbst, wenn diese Studie bereits etwas veraltet ist, so kann angenommen werden, dass die Navigation (damals mit 37%) bei der Nutzung des Smartphones weiterhin eine große Rolle spielt und der Recommender deshalb in Reiseführer-Apps (bspw. auch bei Anwendungen wie der ADAC Trips App) eine entscheidende Rolle spielen könnten. Spannend sind auch jene Apps, bei denen Reisende Aktivitäten vor Ort nutzen können, da diese implizieren, dass Nutzende kurz davor sind, eine Leistung zu buchen (von 19% der Befragten in der BVDW Studie) und dann auch zum jeweiligen Ausflugsziel reisen.

Besuchsverhalten an ausgewählten Standorten im Rahmen des AIR-Projektes

Im Rahmen des Forschungsprojektes „AIR – AI-basierter Recommender für nachhaltigen Tourismus“ fanden u.a. Face to Face-Befragungen am Speicherkoog (n= 421), in Büsum (n= 396) und in St. Peter-Ording (n= 315) zum Besuchsverhalten an verschiedenen touristischen Hotspots statt (Bergler et al., 2023a; Bergler et al., 2023b; Bergler et al., 2023c). Befragt wurden ausschließlich Besucher*innen mit touristisch- bzw. freizeitmotiviertem Besuchsanlass, die das 14. Lebensjahr vollendet haben und über ausreichend deutsche Sprachkenntnisse verfügen.

Zunächst wurden die Teilnehmenden gefragt, welche Art von Informationsquellen sie generell nutzen, um sich über touristische Angebote/Ausflugsmöglichkeiten zu informieren (vgl. Abb. 80). Die Befragungen an allen Standorten ergaben, dass webbasierte Informationsquellen mit Abstand am häufigsten genutzt werden, gefolgt von persönlichen Kontakten. Druckerzeugnisse sowie Funk und Fernsehen werden am wenigsten, allerdings in Büsum dennoch von 12,16% der Befragten, genutzt. Der Anteil der Personen, die sich gar nicht informieren, ist mit Werten zwischen 2,7% und 4,13% sehr gering.

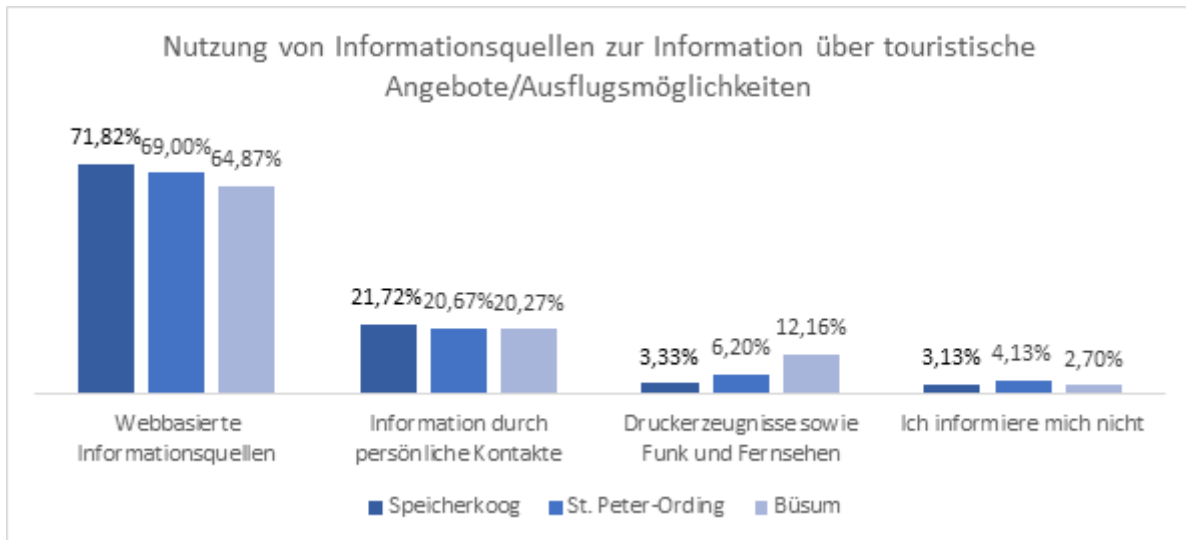


Abbildung 80: Nutzung von Informationsquellen zur Information über touristische Angebote/Ausflugsmöglichkeiten am Speicherkoog, in St- Peter-Ording und in Büsum

Quelle: Eigene Darstellung basierend auf Bergler et al., 2023a; Bergler et al., 2023b; Bergler et al., 2023c

Die anschließenden Befragungsergebnisse der drei Standorte wurden aufgrund der besseren Übersichtlichkeit zusammengefasst dargestellt, sodass sich die Ergebnisse auf die insgesamt 1.132 Befragten am Speicherkoog, in Büsum und in St. Peter-Ording beziehen.

In Abbildung 81 wurde die Nutzung der unterschiedlichen Informationsquellen innerhalb dieser Kategorien zunächst genauer untersucht. Die Teilnehmenden, die diese Art von Informationsquellen jeweils nutzen, wurden gebeten, anzugeben, wie häufig sie die folgenden spezifischen Informationsquellen nutzen. Sie zeigen deutlich, dass bei webbasierten Informationsquellen vor allem Internetsuchmaschinen sowie Websites der Region, der Unterkunft oder der Gastronomie und Freizeitanbieter häufig genutzt werden.

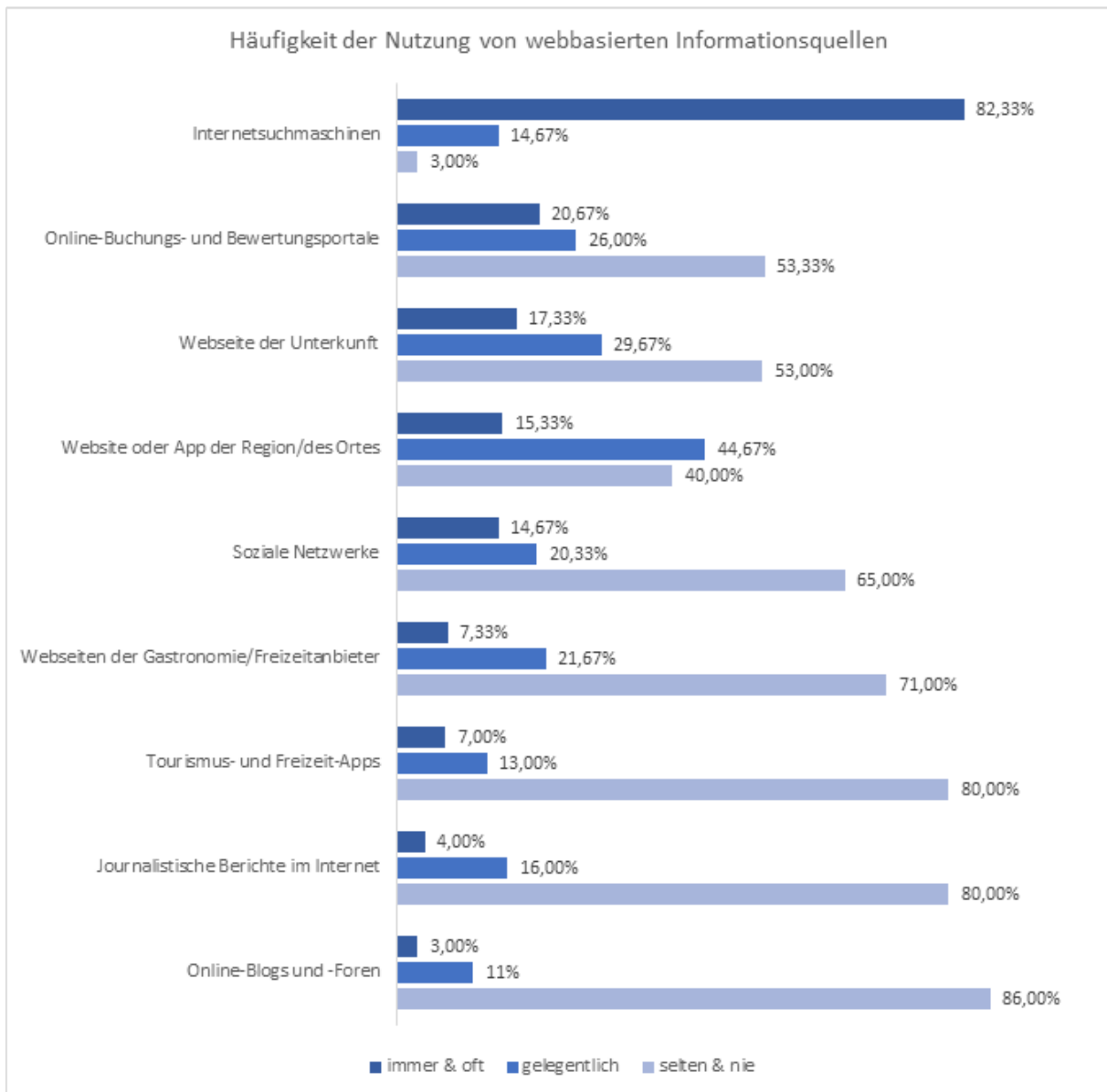


Abbildung 81: Häufigkeit der Nutzung von webbasierten Informationsquellen

Quelle: Eigene Darstellung basierend auf Bergler et al., 2023a; Bergler et al., 2023b; Bergler et al., 2023c

Bei den Druckerzeugnissen sowie Funk und Fernsehen werden vor allem Reiseliteratur, Reiseführer und Reisemagazine, Prospekte und Gastgeberverzeichnisse sowie Zeitungen und Nachrichtenmagazine genutzt (Bergler et al., 2023a; Bergler et al., 2023b; Bergler et al., 2023c).

Werden persönliche Kontakte als Informationsquelle genutzt, sind dies vor allem Verwandte, Bekannte und Freunde. Reisebüros sowie eine Tourist-Information der Region/des Ortes und Gastgeber*innen im Beherbergungsbetrieb werden ebenfalls gelegentlich genannt, jedoch weitaus weniger (Bergler et al., 2023a; Bergler et al., 2023b; Bergler et al., 2023c).

Die Teilnehmenden aller Befragungen wurden ebenfalls gefragt, ob sie sich über ihr heutiges Ausflugsziel informiert haben. Insgesamt haben sich 24% am Speicherkoog, 33% in Büsum und 22% der Befragten über ihr Ausflugsziel informiert (Bergler et al., 2023a; Bergler et al., 2023b; Bergler et al., 2023c).

Anschließend wurden die Befragten zudem gebeten, anzugeben, über welche Themen sie sich vor dem heutigen Ausflug informiert haben (vgl. Abb. 82). Die Ergebnisse zeigen, dass vor allem Informationen über das Wetter, Öffnungszeiten, Veranstaltungen sowie Angebote für Kinder relevant waren. Aber auch das gastronomische Angebot, die Anfahrtsroute sowie Parkmöglichkeiten, Angebote für sportliche Aktivitäten und Sehenswürdigkeiten wurden jeweils von 10% bzw. über 10% der Teilnehmenden genannt.

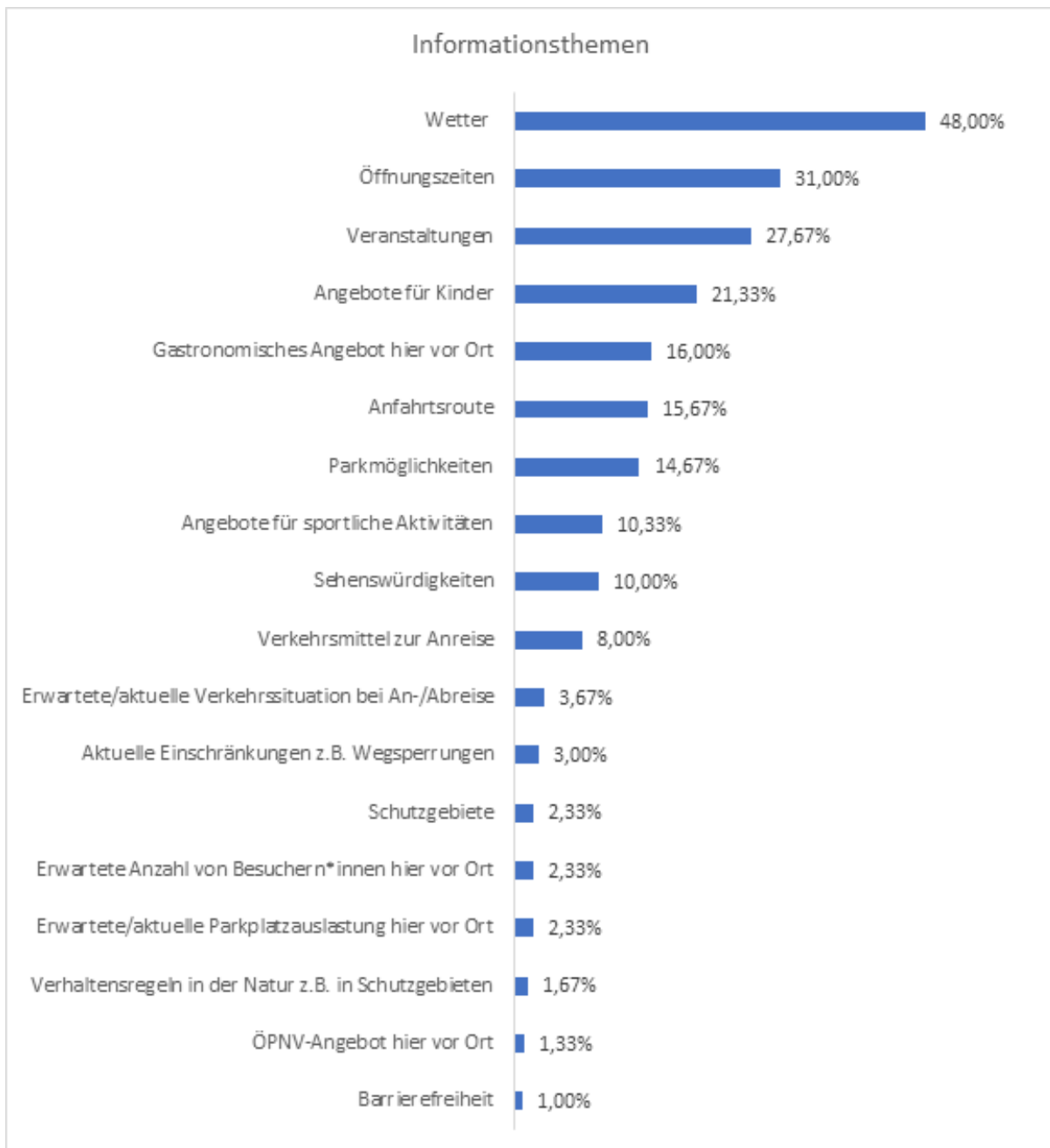


Abbildung 82: Informationsthemen

Quelle: Eigene Darstellung basierend auf Bergler et al., 2023a; Bergler et al., 2023b; Bergler et al., 2023c

Implikationen für die Relevanzabschätzung

Im Rahmen der Face to Face-Befragungen wurde deutlich, dass webbasierte Informationsquellen eine hohe Relevanz haben (82,33% der Befragten nutzen diese). Die konkreten Anwendungen, die genutzt werden sind (neben Suchmaschinen, die hier keine Relevanz haben, weil dort keine Recommender integriert werden können) insbesondere Online-Buchungs- und Bewertungsportale, aber auch Websites von Regionen, Orten oder Unterkünften, was wiederum für die Möglichkeit spricht, einen Recommender ergänzend mit Hilfe eines Widgets zu integrieren, oder aber diesen in bestehende Anwendungen (wie z. B. Karten) zu integrieren. Von besonderem Wert ist die Erkenntnis, dass sich Gäste informieren, die während ihres Urlaubs einen Ausflug machen. Hier scheint ein verstärktes Informationsbedürfnis zu bestehen, das mit einem Recommender befriedigt werden könnte. Da hier insbesondere die Wetter-Angaben sowie (neben Informationen zu POI oder Gastronomie) zur Anfahrtsroute von Interesse sind, könnte ein Recommender einem Abruf der Navigation vorgeschaltet werden, um das Reiseziel überhaupt auszuwählen. Erstaunlich und auffällig ist dabei, dass Live-Informationen zu Besuchsaufkommen, Parkplatzauslastung oder Verkehrssituation kaum eine Rolle spielen. Hier wäre jedoch auch die Frage der Ursache zu stellen, da es sich um eine Henne-Ei-Problematik handeln könnte (das, was nicht erwartet wird, wird auch nicht vermisst). Es kann insgesamt konstatiert werden, dass regionale und örtliche Websites durchaus eine hohe Relevanz haben und somit sehr gute Anknüpfungspunkte für die Integration eines Recommenders bieten.

Aufenthaltort während der Informationsabfrage

Die ARD/ZDF-Onlinestudie (ARD/ZDF-Forschungskommission, 2021) belegt, dass die Internetnutzung unterwegs eine immer größere Bedeutung erhält (siehe Anhang). In 2021 haben 79% der Befragten das Internet zumindest gelegentlich unterwegs genutzt. Insbesondere die jüngeren Altersgruppen nutzen das Internet viel unterwegs. Bei den 14-29-Jährigen liegt der Anteil bei 97% und bei den 30-49-Jährigen bei 94%. Dies kann darauf hindeuten, dass auch die Informationsabfrage während der An- bzw. Abreise, oder während des Aufenthaltes in der Destination an Bedeutung gewinnt. Ein weiterer Indikator dafür könnte die Studie von Bitkom Research (Bitkom e.V., 2020b) sein, die zeigt, dass 76% der Befragten ihr Smartphone mit in den Urlaub nehmen und diverse Apps für ihren Urlaub nutzen.

Die im Projekt AIR durchgeführte Befragung von Personen, die zuvor angegeben haben, sich informiert zu haben, zeigt jedoch, dass sich ein Großteil der Besucher*innen bereits mehrere Tage im Voraus über ihr Ausflugsziel informiert (vgl. Abb. 83). Ein ebenso nennenswerter Anteil informiert sich noch am Tag des Ausfluges, bevor sie losgefahren sind, während das Informationsverhalten auf dem Weg und am Ziel stark abnimmt (Bergler et al., 2023a; Bergler et al., 2023b; Bergler et al., 2023c).

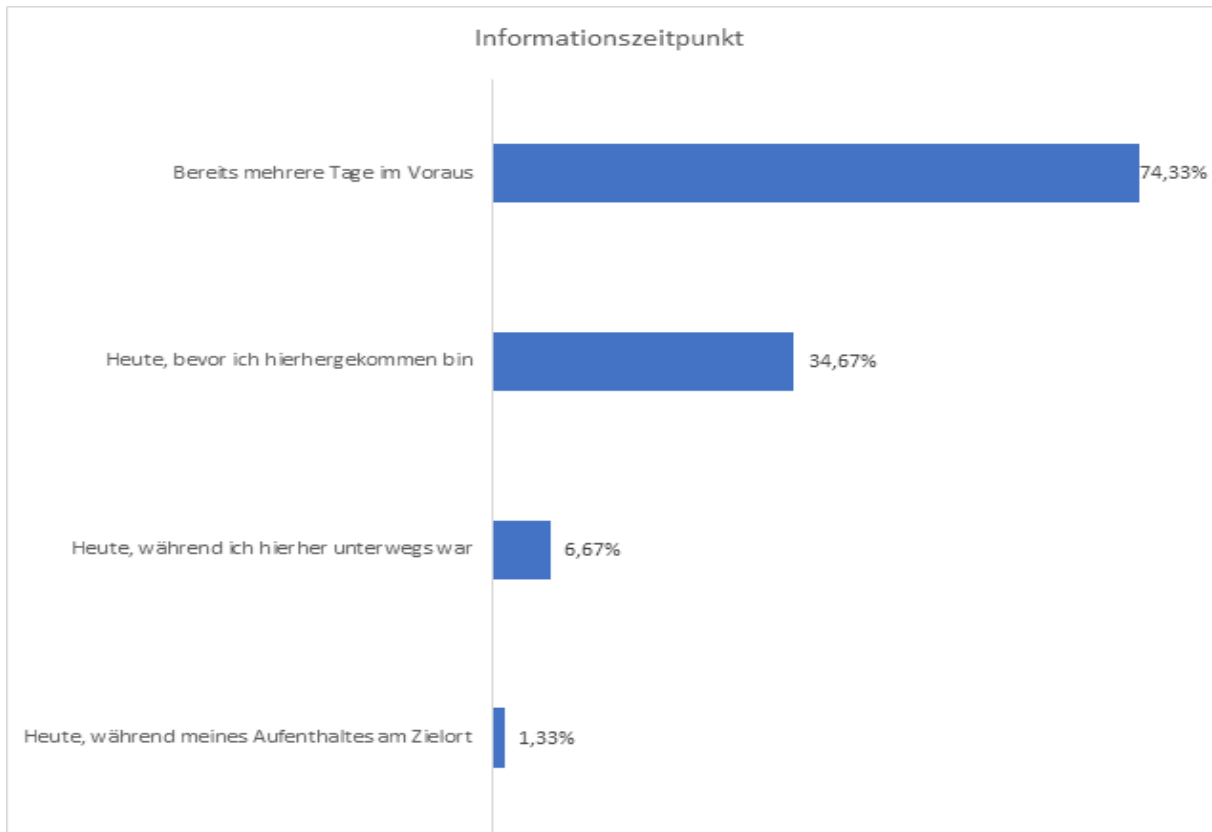


Abbildung 83: Informationszeitpunkt

Quelle: Eigene Darstellung basierend auf Bergler et al., 2023a; Bergler et al., 2023b; Bergler et al., 2023c

Implikationen für die Relevanzabschätzung

Die Relevanz des Smartphones als Ausspielungskanal konnte bereits durch die intensive Nutzung belegt werden und wird hier nochmals im Kontext des Urlaubs bestätigt. Dadurch, dass Gäste sich jedoch mehrere Tage oder am Anreisetag selbst informieren, ist anzunehmen, dass sie dafür nicht nur das Smartphone nutzen, sondern mitunter vermutlich auch das Laptop, das aber ebenfalls ein wichtiges Endgerät ist, was auch bisher in den Erhebungen deutlich wurde.

Informationsverhalten in Bezug auf Urlaubsreisen

In einer Studie des DI Tourismusforschung wurde das Informationsverhalten in der Inspirations- und Informationsreise bei der Campingplatzwahl von 1.084 Befragten untersucht (DI Tourismusforschung, 2023). In beiden Phasen sind bei den Offline-Kommunikationswegen demnach Gespräche mit Verwandten, Bekannten und Freunden, Campingführer und -zeitschriften und die Information beim Campingplatz am relevantesten, wie Abbildung 84 zeigt.

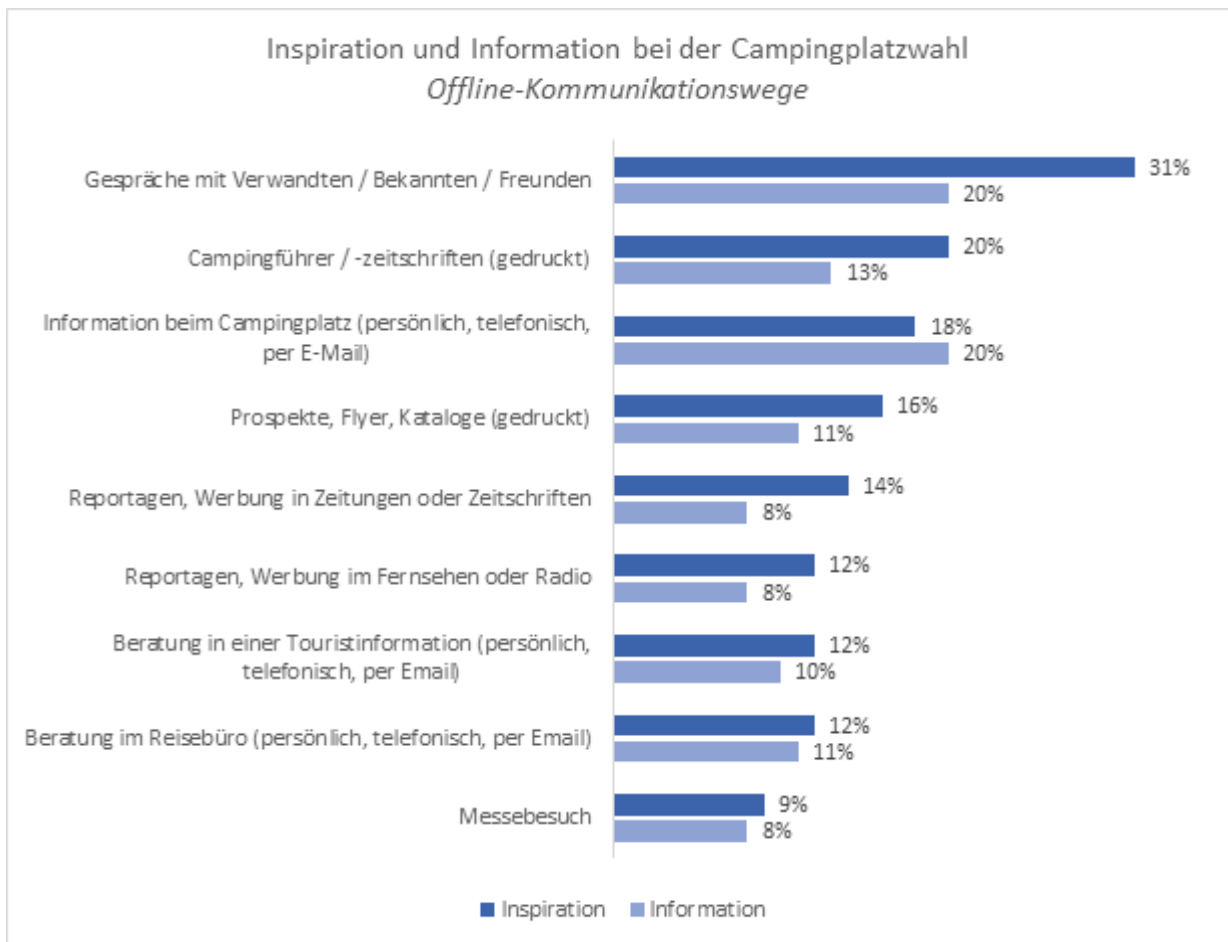


Abbildung 84: Inspiration und Information bei der Campingplatzwahl durch Offline-Kommunikationswege

Quelle: DI Tourismusforschung, 2023

Bei den Online-Kommunikationswegen werden in der Inspirationsphase Google und andere Suchmaschinen, Websites von Campingplätzen sowie Websites von Reisezielen am häufigsten genutzt. In der Informationsphase werden Websites von Online-Campingportalen am dritthäufigsten genutzt (vgl. Abb. 85 - 87).

Inspiration und Information bei der Campingplatzwahl Offline-Kommunikationswege – nach Planung Campingplatzbesuche

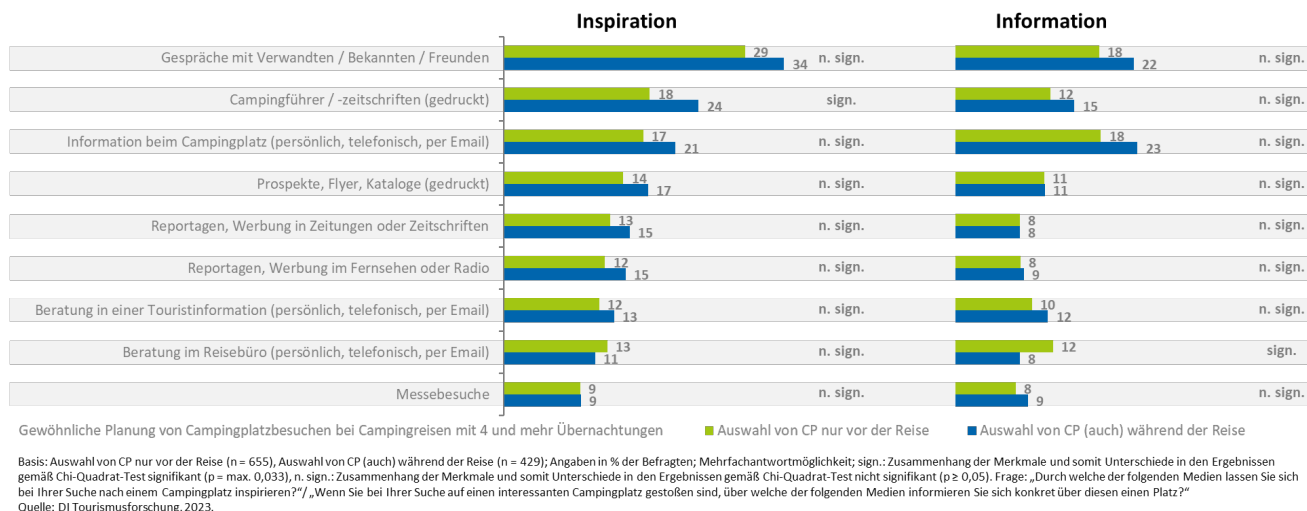


Abbildung 85: Inspiration und Information bei der Campingplatzwahl durch Offline-Kommunikationswege nach Planung Campingplatzbesuche

Quelle: DI Tourismusforschung, 2023

Inspiration und Information bei der Campingplatzwahl Online-Kommunikationswege – nach Planung Campingplatzbesuche

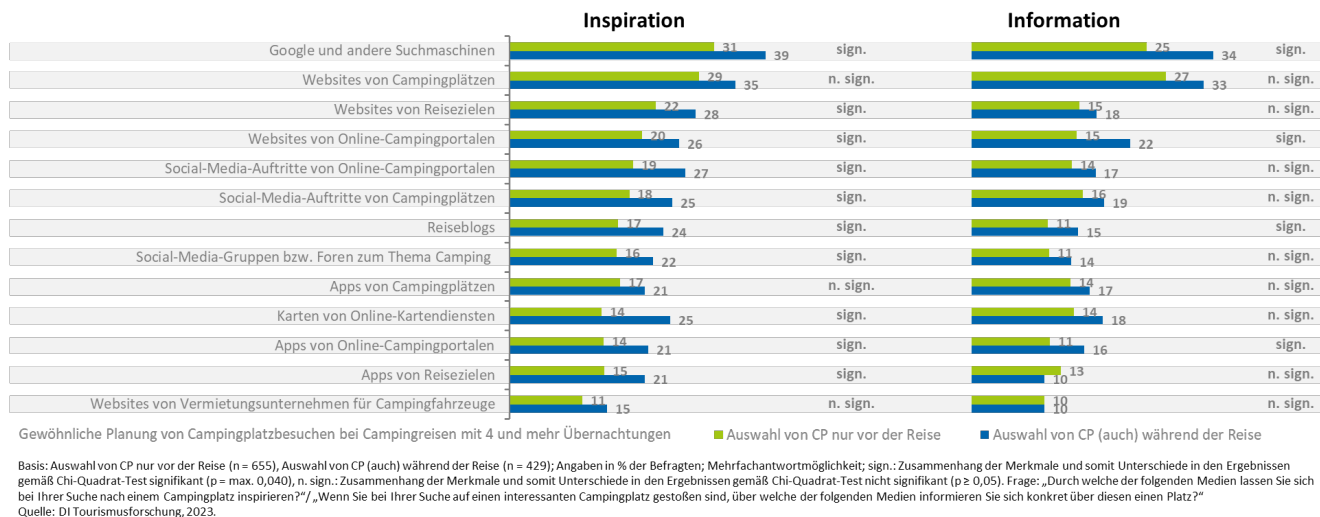
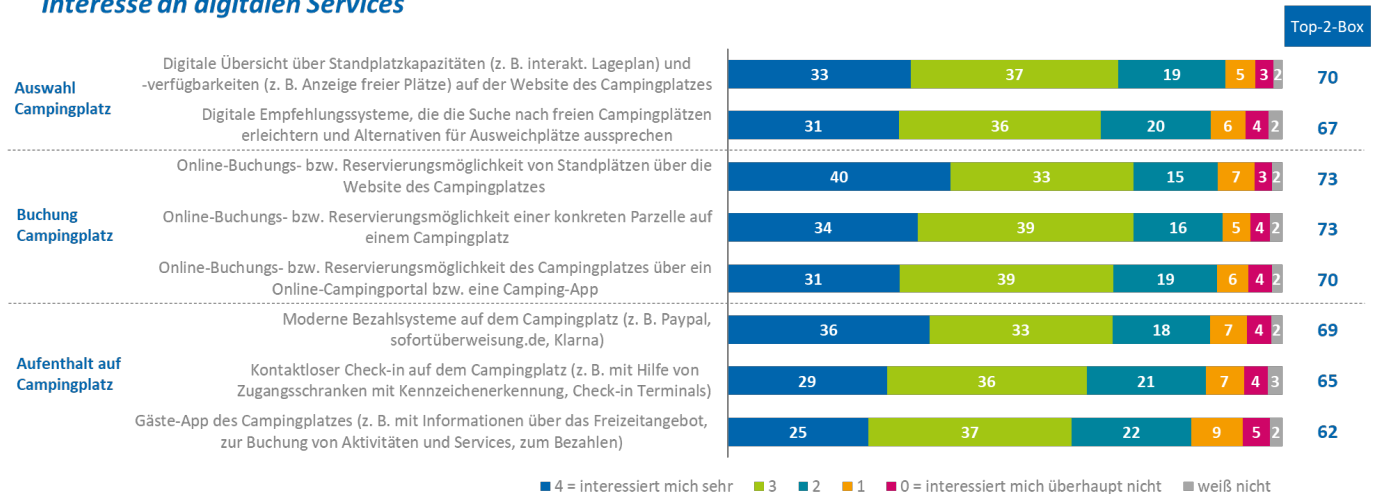


Abbildung 86: Inspiration und Information bei der Campingplatzwahl durch Online-Kommunikationswege nach Planung Campingplatzbesuche

Quelle: DI Tourismusforschung, 2023

Interesse an digitalen Services entlang der Customer Journey

Interesse an digitalen Services



Basis: Alle Befragte (n = 1.084); Angaben in % der Befragten; Abweichungen der Summe der Einzelwerte zu 100% sind rundungsbedingt.
 Frage: „Bitte geben Sie an, inwiefern Sie Interesse an den folgenden Angeboten / Leistungen haben.“
 Quelle: DI Tourismusforschung, 2023.

Abbildung 87: Interesse an digitalen Services entlang der Customer Journey

Quelle: DI Tourismusforschung, 2023

In der Campingplatzstudie vom DI Tourismusforschung wurde die Bekanntheit und Nutzung von Campingportalen bzw. -apps untersucht (DI Tourismusforschung, 2023). Am bekanntesten und am häufigsten genutzt wird der ADAC Camping- und Stellplatzführer, gefolgt von Camping.info und Campings (vgl. Abb. 88).

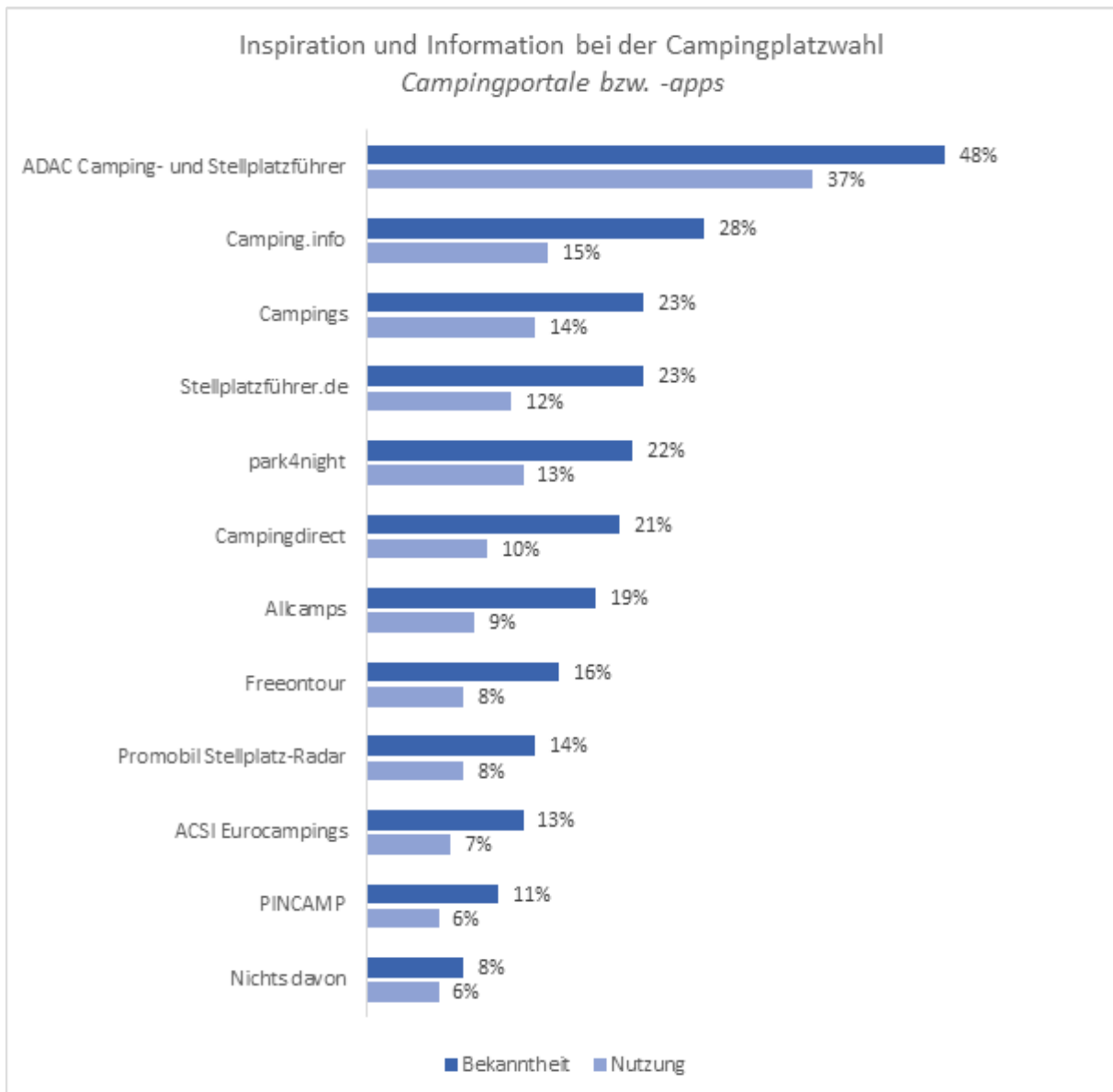


Abbildung 88: Inspiration und Information bei der Campingplatzwahl durch Camping-Portale bzw. -apps

Quelle: DI Tourismusforschung, 2023

Implikationen für die Relevanzabschätzung

Auch hier wird deutlich, dass ein Recommender die Möglichkeit offerieren könnte, Alternativen direkt anzubieten, wenn ein Campingplatz überfüllt ist und so würde der Parksuchverkehr reduziert werden. Gleichwohl bleibt zu prüfen, ob Campingplätze an der Integration solcher Informationen auf ihrer eigenen Website überhaupt Interesse hätten, oder ob es zu einem Konkurrenzdenken kommen würde (was aus wirtschaftlicher Perspektive anzunehmen wäre). Anders sieht es bei Online-Campingportalen aus, die naturgemäß eine Bandbreite an Auswahlmöglichkeiten zeigen. Hier wäre ein Recommender sehr hilfreich, um Campingplätze, die ausgebucht sind direkt auszublenden und andere zu empfehlen, die noch Kapazitäten haben. Die in der Studie des DI Tourismusforschung von 2023 genannten Websites könnten dabei wichtige Anlaufpunkte sein, um einen Recommender dort zu pilotieren – Voraussetzung wäre, dass die Campingplätze in einer Region (hier Schleswig-Holstein) nahezu flächendeckend

mit Sensorik ausgestattet sind, um Aussagen über den Füllstand des jeweiligen Campingplatzes treffen zu können.

Zusammenfassende Beurteilung der Relevanzabschätzung

Abschließend werden die Erkenntnisse der Analyse der Relevanzabschätzung hier noch einmal zusammengefasst. Die Kernergebnisse sind:

- **Endgeräte:** Es wurde deutlich, dass Smartphones im Kontext des digitalen Besuchermanagements das wichtigste Ausgabemedium ist. Aber auch Laptops und Tablets sind (noch) wichtig. Keine durchdringende Nutzung konnte hingegen für Wearables festgestellt werden. Die Ausgabe über Sprachassistenten sowie Virtual Reality spielen ebenfalls eine untergeordnete Rolle. Das Smart TV ist hingegen ebenfalls ein wichtiger Ausgabekanal und gewinnt zusätzlich an Relevanz, weil in Hotels hier oftmals ein personalisierter Startbildschirm implementiert werden kann, auf dem Angaben zu Besuchsaufkommen und Empfehlungen zu Alternativen sehr gut ausgespielt werden können.
- **Nutzungsdauer und Nutzungsart:** Fernsehen und Streaming wichtig. Gäste suchen Neuigkeiten. Daher Live-Daten. Wetter wird verstärkt gesucht und sollte auch mit angegeben werden, damit Nutzende einen Grund haben, die Anwendung aufzusuchen und zu nutzen. Navigation auch wichtig. Daher zumindest ein Link dorthin!
- **Medienformate:** Es konnte festgestellt werden, dass das Format Video wichtig ist und eine hohe Tagesreichweite hat. Die Ausspielung von Live-Bildern ist jedoch kritisch hinsichtlich des Datenschutzes. Gleichwohl sollte bei der Pilotierung eines Recommenders überlegt werden bestehende Webcams zu integrieren, damit sich die Besuchenden direkt einen Eindruck vor Ort machen können.
- **Zeitpunkt der Nutzung:** Herausgefunden wurde auch, dass die Nutzenden sich mehrere Tage im Voraus sowie am Anreisetag selbst informieren. Daher sollte ein Recommender sowohl Live-Informationen, aber auch Prognosen enthalten.
- **Drittanbieter:** Es wurde deutlich, dass von Nutzenden Reiseportale sowie Tourenplaner (outdooractive und komoot) und nicht zuletzt auch große Anbieter wie Google und Apple (Maps) genutzt werden. Realistischer ist daher die Einbindung auch bei themenspezifischen Apps wie z. B. der ADAC Trips App und weiteren.
- **Touristische Websites:** Relevant sind auch die Website von Regionen, Orten und Unterküften. Es kann mit einem Piloten daher auch die Strategie verfolgt werden, diesen auf vielen unterschiedlichen kleinen und weniger sichtbaren Websites zu integrieren. Grundsätzlich ist es daher sinnvoll, für alle Arten von externen Anbieter Lösungen anzubieten, die es ihnen ermöglicht, einen Recommender flexibel in ihr bestehendes Webangebot einzubinden. Widgets sind deshalb eine gute Lösung und eine Pilotanwendung sollte eine solche Form der Einbindung (i.d.R. via iFrame) anbieten - selbst, wenn eine Integration bei globalen Playern vermutlich unrealistisch ist, aber es gibt viele weitere, die auf ein solches Angebot zurückgreifen könnten (insb. Websites von Regionen, Orten und Leistungsträgern).

6.3 Analyse der Informationskanäle

Nachdem im ersten Schritt die Verbreitung von Endgeräten, die Nutzungsintensität unterschiedlicher Funktionen und die jeweiligen Ausspielkanäle allgemein analysiert wurden, wird im nächsten Schritt der Fokus auf konkrete Anwendungen und Kanäle gelegt, die auf Basis der Ergebnisse aus 6.1 infrage

kommen, um als Ausspielkanal zu fungieren. Relevante und im Vorfeld für das Projekt und auf den Use Cases beruhende Kategorien von Informationskanälen sind:

1. Digitale Tourenplaner (z. B. Komoot, Outdooractive)
2. Digitale Routenplaner für Individualverkehr und ÖPNV (z. B. Google Maps, DB Navigator)
3. Digitale Kartendienste (z. B. Google Maps, Digitaler Atlas Nord, OpenStreetMap)
4. Internetsuchmaschinen (z. B. Google)
5. Informationen der/über die Destinationen und Attraktionen (Indoor oder Outdoor)

In diesen Kategorien (Internetsuchmaschinen, digitale Routenplaner für den Individualverkehr und öffentlichen Personenverkehr, digitale Tourenplaner und digitale Kartendienste) wurden konkrete Anwendungen recherchiert sowie deren Relevanz und Verbreitung verglichen. Dies wurde vor dem Hintergrund der Use Case Spezifika (Camping, Radfahren, Surfen/Kiten, Parken) bewertet. Zudem wurden Apps, Websites und Anwendungen der Tourismusverbände in Schleswig-Holstein recherchiert und eine Sichtbarkeitsanalyse durchgeführt, da diese regional die höchste Relevanz besitzen.

Auf Basis der Erkenntnisse von 6.1 und 6.2 wird ein Relevanzbewertungsverfahren entwickelt. Ziel ist es dabei grundsätzlich, dass die Ausspielung eine möglichst große Wirkung entfalten kann. Dementsprechend sollte eine Relevanzbewertung auf dieses Ziel hin ausgerichtet werden.

Internetsuchmaschinen

Abbildung 89 zeigt eine Studie von StatCounter (StatCounter, 2023) die verdeutlicht, dass Google mit Abstand den größten Marktanteil der Suchmaschinen in Deutschland für Desktop, Mobiltelefon und Tablet hat. An zweiter Stelle folgt mit großem Abstand Bing. Andere Suchmaschinen haben nur einen geringen Marktanteil. Die Studie bezieht sich auf den Zeitraum Januar 2022 bis Mai 2023.

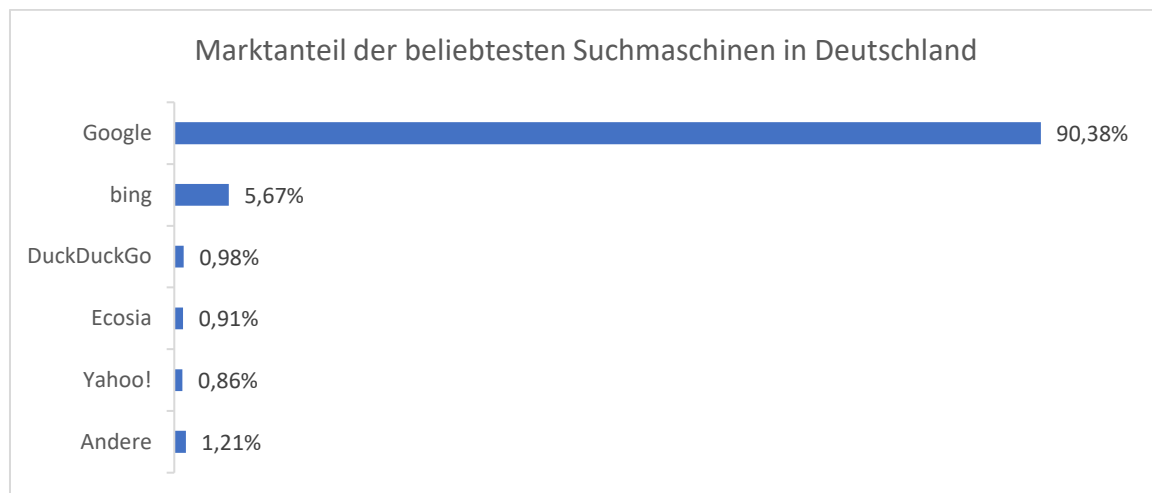


Abbildung 89: Marktanteil der beliebtesten Suchmaschinen in Deutschland.

Quelle: Eigene Darstellung basierend auf StatCounter, 2023

Die von GIM durchgeführte Mediengewichtungsstudie (die medienanstalten - ALM GbR, 2022) ergibt außerdem, dass Google ebenfalls die mit Abstand größte Tagesreichweite in Deutschland hat (vgl. Abb. 90). Bing folgt mit großem Abstand an zweiter Stelle. Yahoo hat ebenfalls einen geringen Anteil. Für

die Studie wurden 6.040 Personen ab 14 Jahren in Deutschland gefragt, welche Suchmaschine sie am Vortag der Befragung genutzt haben.

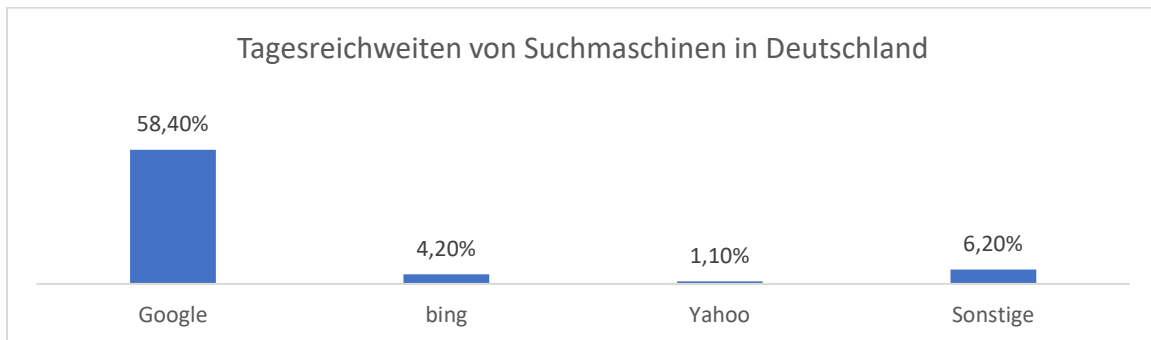


Abbildung 90: Tagesreichweiten von Suchmaschinen in Deutschland

Quelle: Eigene Darstellung auf Basis von die medienanstalten - ALM GbR, 2022

22,3% der Befragten einer weiteren Studie gaben an, eine Suchmaschine als Kontakthersteller zu Online-Infos zu nutzen, wie in Abbildung 91 ersichtlich wird. Dies bedeutet eine Steigerung von 2% im Vergleich zum Vorjahr 2020. Datenbasis ist eine Stichprobe von 3.660 Personen ab 14 Jahren in Deutschland. Von den 779 Personen, die eine Suchmaschine als Kontakthersteller genutzt haben, wurde Google mit Abstand am häufigsten als genutzte Suchmaschine genannt (91,80%). Mit großem Abstand folgen Suchmaschinen wie Web.de, Bing, Yahoo und T-Online.

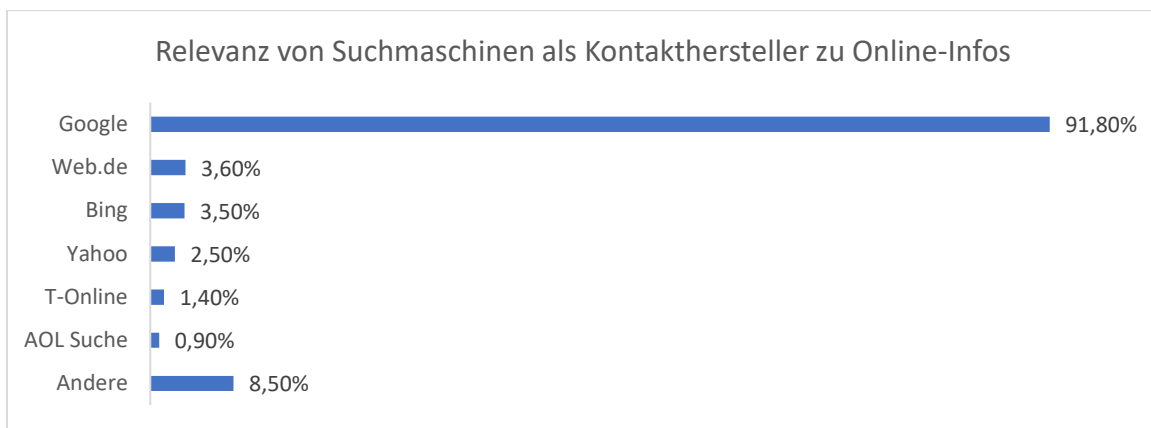


Abbildung 91: Relevanz von Suchmaschinen als Kontakthersteller zu Online-Infos

Quelle: Eigene Darstellung basierend auf die medienanstalten - ALM GbR, 2021

Implikationen für die Analyse der Informationskanäle

Grundsätzlich starten fast alle Suchen über das Internet bei Google, wenn es um die Nutzung von Suchmaschinen geht, wie die Zahlen von StatCounter zeigen. Dies muss jedoch nicht gleichbedeutend damit sein, dass auch die spezielle Suche nach Informationen zu Besuchsaufkommen und Besucherströmen ebenfalls über eine Suchmaschine startet. Dennoch wäre es vor dem Hintergrund dieser Dominanz und Sichtbarkeit von Google natürlich erstrebenswert, wenn die Informationen eines Recommenders

von Google berücksichtigt würden. Die Möglichkeit wäre hier zum einen eine Integration in die Suchmaschinenarchitektur, was jedoch als regionales und zeitlich begrenztes Forschungsprojekt unrealistisch erscheint, hier eine Verbindung hinzubekommen. Nichtsdestotrotz gilt dies für andere Projekte als erstrebenswert, um die Sichtbarkeit auf Google zu erhöhen. Dies kann auch gelingen, indem ein Recommender in eine Website integriert ist und mit SEO-Maßnahmen so beschrieben wird, dass eine Auffindbarkeit auf den Suchergebnisseiten (SERPs) möglich ist. Dies kann über eine entsprechende Verschlagwortung (aktuelles Besucheraufkommen in XY usw.) erfolgen. Gleichzeitig wäre es ein Ansatz, in der speziellen Suchmaschine "Google Things To Do" diejenigen POI, die mit Informationen zum Besuchsaufkommen ergänzt wurden, als strukturierte Daten zu annotieren und ihnen so dort eine erhöhte Sichtbarkeit zu verschaffen.

Digitale Routenplaner für den Individualverkehr und Öffentlichen Personennahverkehr (ÖPNV)

Im Jahr 2019 wurde eine Umfrage in Deutschland zur Nutzung von Apps für die Fahrplanauskunft für den ÖPNV durchgeführt (Statista, 2019). Es wurden 1.230 Personen im Alter von 18 bis 69 Jahren der deutschen Bevölkerung befragt, die den ÖPNV mindestens einmal im Monat nutzen, in einer der zehn größten Städte Deutschlands wohnen und das Smartphone für die Fahrplanauskunft nutzen. Am häufigsten werden demnach die Apps der Verkehrsbetriebe, die Deutsche Bahn App und die Websites der Verkehrsbetriebe genutzt (vgl. Abb. 92).

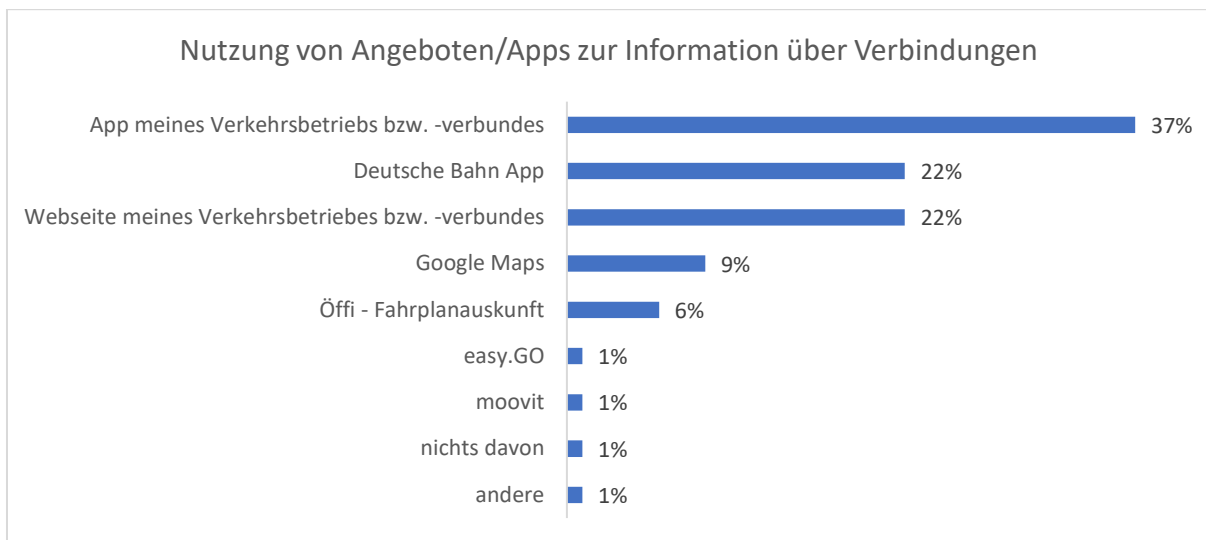


Abbildung 92: Nutzung von Angeboten/Apps zur Information über Verbindungen

Quelle: Eigene Darstellung basierend auf Statista, 2019

Eine kontinuierliche Marktforschung zum Öffentlichen Verkehr in Schleswig-Holstein aus dem Jahr 2021 (O.trend GmbH, 2022) bekräftigt die Ergebnisse (vgl. Abb. 93). Die meisten der 4.300 Befragten gaben an, „bahn.de“ als Informationsquelle im Internet für eine unbekannte Bahnverbindung in Schleswig-Holstein zu nutzen, gefolgt von Fahrplan-Apps, „hvv.de“ und „nah.sh“. Die Studie zeigt zudem, dass sich die Relevanz der Informationsquellen im Laufe der letzten Jahre geändert hat. Suchmaschinen haben an Bedeutung verloren, wohingegen Fahrplan-Apps, „hvv.de“ und „nah.sh“ an Bedeutung gewonnen haben.

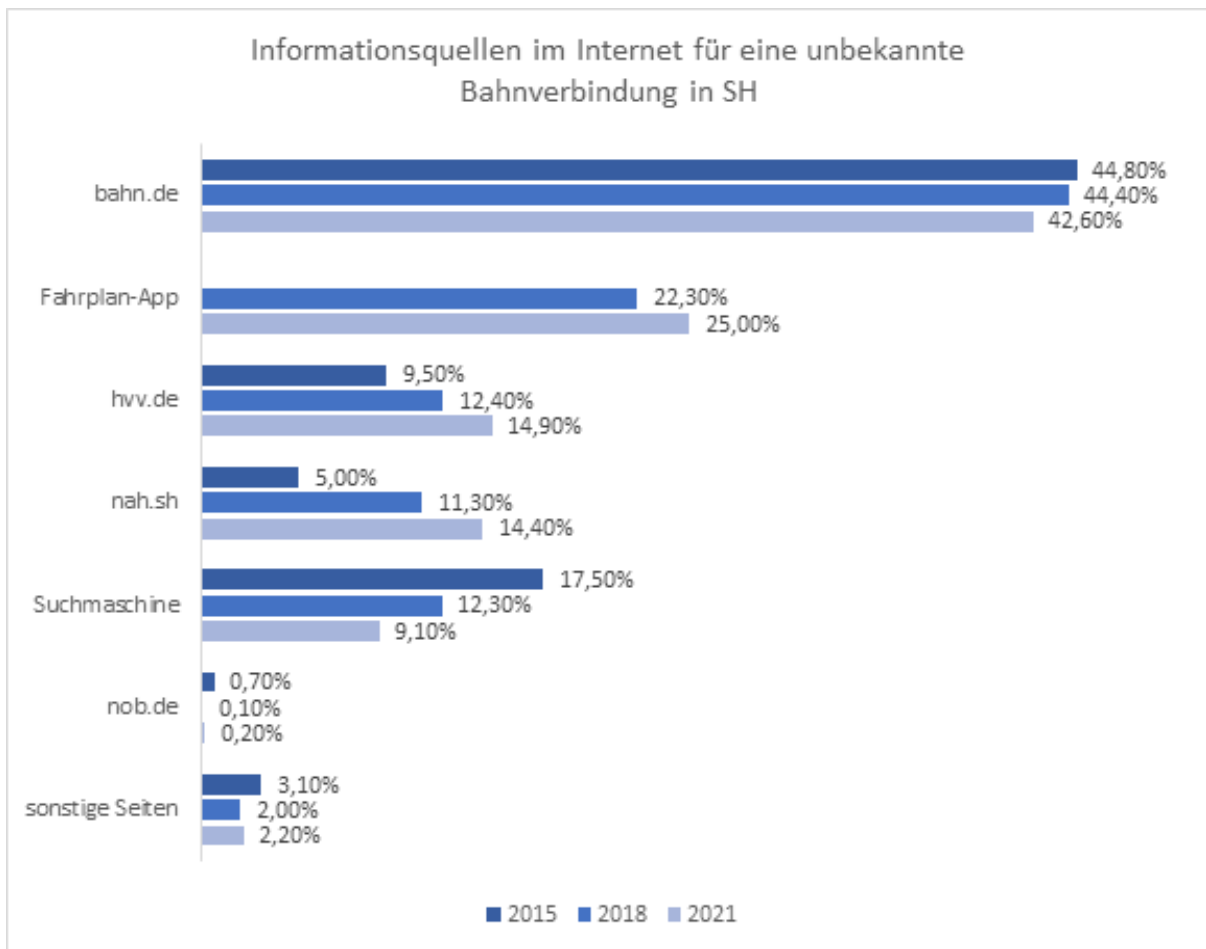


Abbildung 93: Informationsquellen im Internet für eine unbekannte Bahnverbindung in SH

Quelle: Eigene Darstellung basierend auf O.trend GmbH, 2022

Implikationen für die Informationskanäle

Es wird deutlich, dass die App der Bahn eine hohe Verbreitung hat. Dies ist insbesondere dann interessant, wenn ein Recommender Alternativen anzeigt, die mit öffentlichen Verkehrsmitteln gut erreichbar sind. Die Bahn entwickelte hier komplementär die DB Ausflug App. Diese ist dafür prädestiniert, dass POI angezeigt werden und zu diesen auch das aktuelle Besuchsaufkommen. Da diese Apps auch durch regionale Verkehrsbetriebe gepflegt werden können, ist es durchaus möglich, dass über einen direkten Kontakt eine solche Integration erfolgreich sein könnte - was zu einer sehr hohen Verbreitung und Sichtbarkeit des Recommenders führen würde.

Digitale Tourenplaner

Eine Online-Befragung des Allgemeinen Deutschen Fahrradclubs e.V., die ADFC-Radreiseanalyse aus dem Jahr 2021 (Allgemeiner Deutscher Fahrrad-Club e.V., 2022), ergibt, dass 75% der Befragten während einer Radreise die Wegweisung vor Ort als Informationsquelle nutzen (vgl. Abb. 94). Gefolgt wird dies von Apps mit 61,1% Nutzer*innen sowie mobiler Internetrecherche (58,4%) und gedruckten Radkarten (47,4%). Unter den App-Nutzer*innen während einer Radreise sind die Plattformen Komoot mit 66 %, gefolgt von Google Maps mit 63% und Outdooractive mit 13% führend. Die Daten der ADFC-

Radreiseanalyse wurden im Zeitraum November bis Januar über zwei geschlossene bundesweite, repräsentative Panels und eine offene Befragung, die über die Kommunikationskanäle des ADFC veröffentlicht wurde, generiert. Insgesamt standen 10.269 auswertbare Fragebögen zur Verfügung.

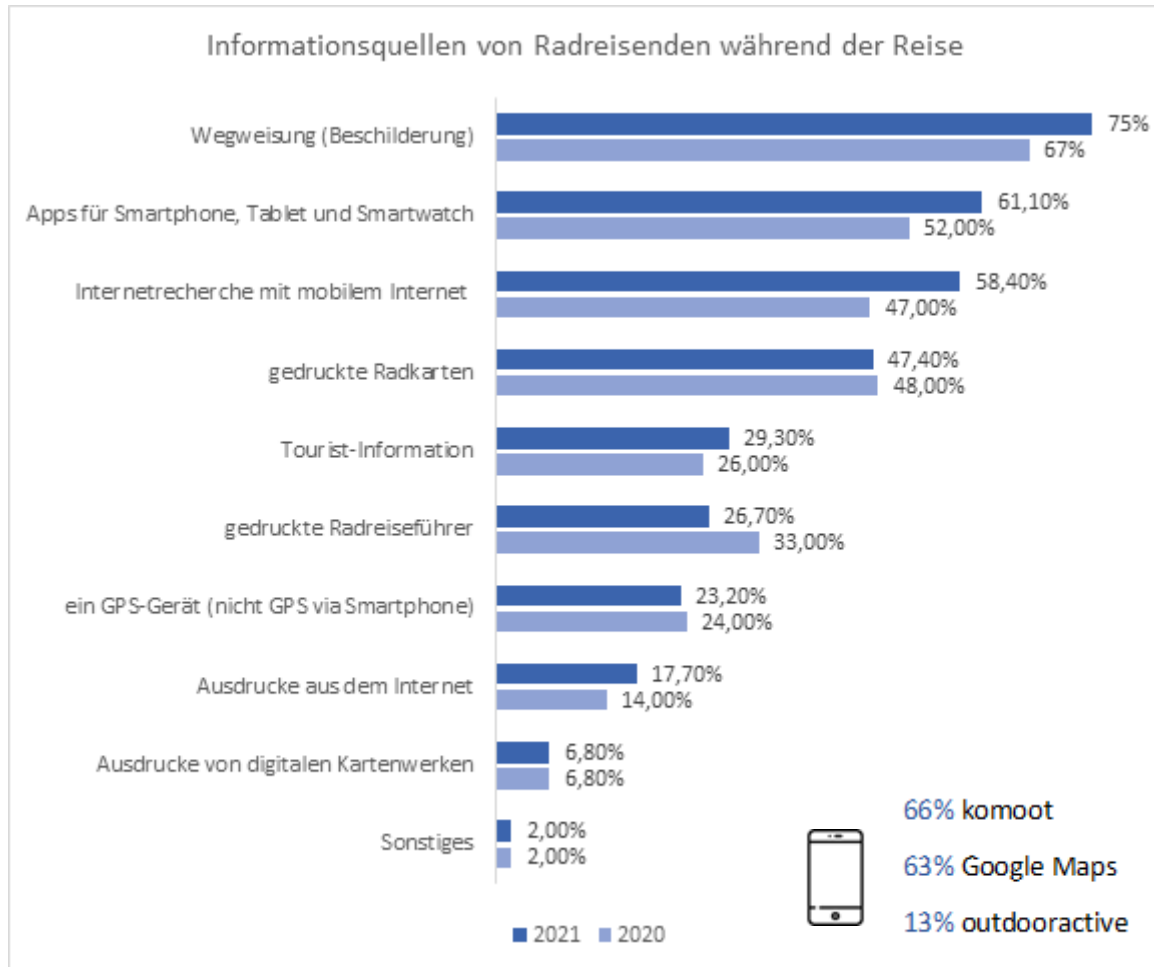


Abbildung 94: Informationsquellen von Radreisenden während der Reise.

Quelle: Eigene Darstellung basierend auf Allgemeiner Deutscher Fahrrad-Club e.V., 2022

Abbildung 95 zeigt die Ergebnisse der Bergzeit Wanderstudie 2022, bei der insgesamt 75% der Befragten angaben, dass sie Portale wie Outdooractive, Bergfex, Komoot & Co. als Inspirationsquelle für Wandertouren nutzen. Besonders bei den 26-45-Jährigen sind Wanderportale beliebt. Befragt wurden 3.478 Personen aus Deutschland, Österreich und der Schweiz. Die Umfrage richtete sich an wanderinteressierte Personen und wurde im Zeitraum 24.02.2022 bis 17.03.2022 durchgeführt.

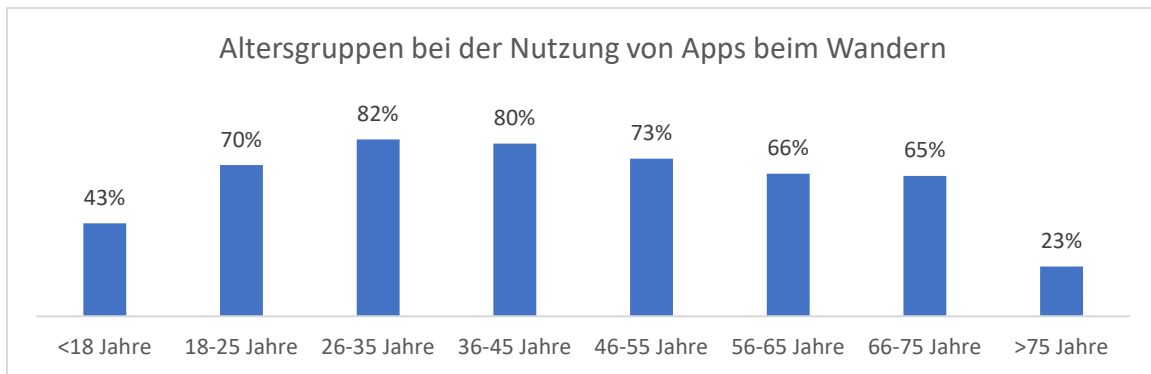


Abbildung 95: Nutzung von Apps beim Wandern nach Altersgruppe

Quelle: Eigene Darstellung basierend auf Bergzeit GmbH, 2022

In der Bergzeit Wanderstudie 2022 wurde ebenfalls untersucht, welche Apps die Befragten beim Wandern nutzen (vgl. Abb. 96). Die meisten Befragten nutzen Komoot. Mit einigem Abstand folgen Wetter-Apps, Outdooractive und Google Maps (Bergzeit GmbH, 2022).

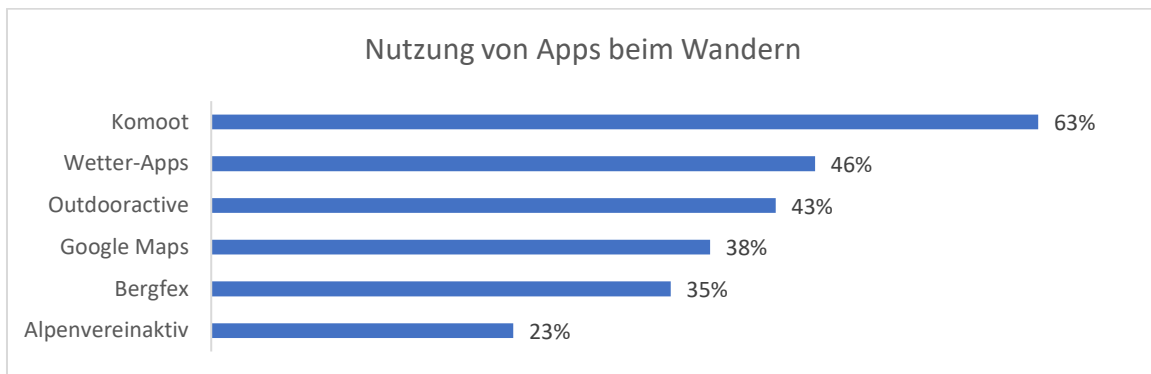


Abbildung 96: Nutzung von Apps beim Wandern nach Anbieter

Quelle: Eigene Darstellung basierend auf Bergzeit GmbH, 2022

Implikationen für die Informationskanäle

Mit Blick auf die Studien zu digitalen Tourenplanern, nimmt die Ad-hoc Recherche von Radfahrenden durch Beschilderung (dies können auch digitale Schilder sein, die eine Live-Auslastung anzeigen, ähnlich wie bei einem Parkleitsystem in der Stadt) und Apps zu. Besonders verbreitet ist dabei komoot (mit 66%), sodass eine Integration in diesen Dienst wünschenswert wäre. Möglich wäre dies durch den Einkauf von Sichtbarkeit in Form von Collections oder Ähnlichem, was komoot als Dienst den Destinationen anbietet. So könnte ein Prototyp des Recommenders prominent ausgespielt werden. Deutlich wird auch, dass insgesamt die Nutzung von analogen Formaten wie Radkarten und Reiseführern abnimmt (bei Reiseführern von 33% auf 26,7% in nur einem Jahr). Die Nutzung von digitalen Orientierungsangeboten wird beim Wandern noch einmal untermauert indem die Bergzeit Wanderstudie zeigt, dass Wander*innen zwischen 18 und 75 Jahren jeweils mindestens zu 62% Apps nutzen und davon mit Abstand die meisten komoot (mit 63%).

Destinationswebsites aus Schleswig-Holstein

Um eine erfolgreiche Ausspielung eines Recommenders gewährleisten zu können, ist es von großer Bedeutung, die Sichtbarkeit potenzieller Ausspielkanäle zu analysieren. Wie die vorherige Recherche eindeutig aufgezeigt hat, wird vornehmlich Google unter den Suchmaschine genutzt. Daher ist es in diesem Rahmen auch wichtig einordnen zu können, inwiefern diese Suchmaschine den Zielgruppen die für einen Recommender in Schleswig-Holstein wichtiger Domains anzeigt.

Für das Projekt LAB-TOUR SH wurden die relevanten Websites der Destinationen ausgewertet (vgl. Abb. 97). Grundlage dafür waren die Mitglieder des Tourismusverbandes Schleswig-Holstein (TVSH). Aus der Analyse der Relevanzabschätzung ging hervor, dass Widgets - also das Einbinden des Recommenders in eine bestehende Website - eine gute Möglichkeit der Implementierung sein könnte. Dementsprechend wird für die Websites der TVSH Mitglieder eine entsprechende Relevanz angenommen. Weiterführend wurden zudem Websites ausgewertet, bei denen von einer hohen Relevanz für den jeweiligen Use Case ausgegangen wird. Auch bei dieser Analyse ist es das Ziel, den Recommender auf spezifischen Ausspielkanälen zur Verfügung zu stellen, auf die die Nutzenden vor oder während ihrer Reise mit hoher Wahrscheinlichkeit zugreifen könnten.

Hinweise zum Analyse-Tool: Es wurde mit Hilfe von SISTRIX eine Sichtbarkeitsanalyse relevanter Informationskanäle durchgeführt. Die Berechnung dieses Index erfolgt anhand eines repräsentativen Querschnitts von 1 Million Suchanfragen bei Google, von denen jeweils die ersten 100 Treffer als Datengrundlage genutzt werden. Die daraus resultierenden 100 Millionen Datenpunkte werden anschließend nach dem jeweiligen Suchvolumen sowie der Klickwahrscheinlichkeit gewichtet und für alle Rankings der Domain aufsummiert. Der Sichtbarkeitsindex ergibt sich dann aus der Summe all dieser Werte. Je höher der Wert ist, desto sichtbarer sind die Websites und Informationskanäle, wodurch eine erhöhte Wahrnehmung bei den Nutzenden impliziert werden kann.

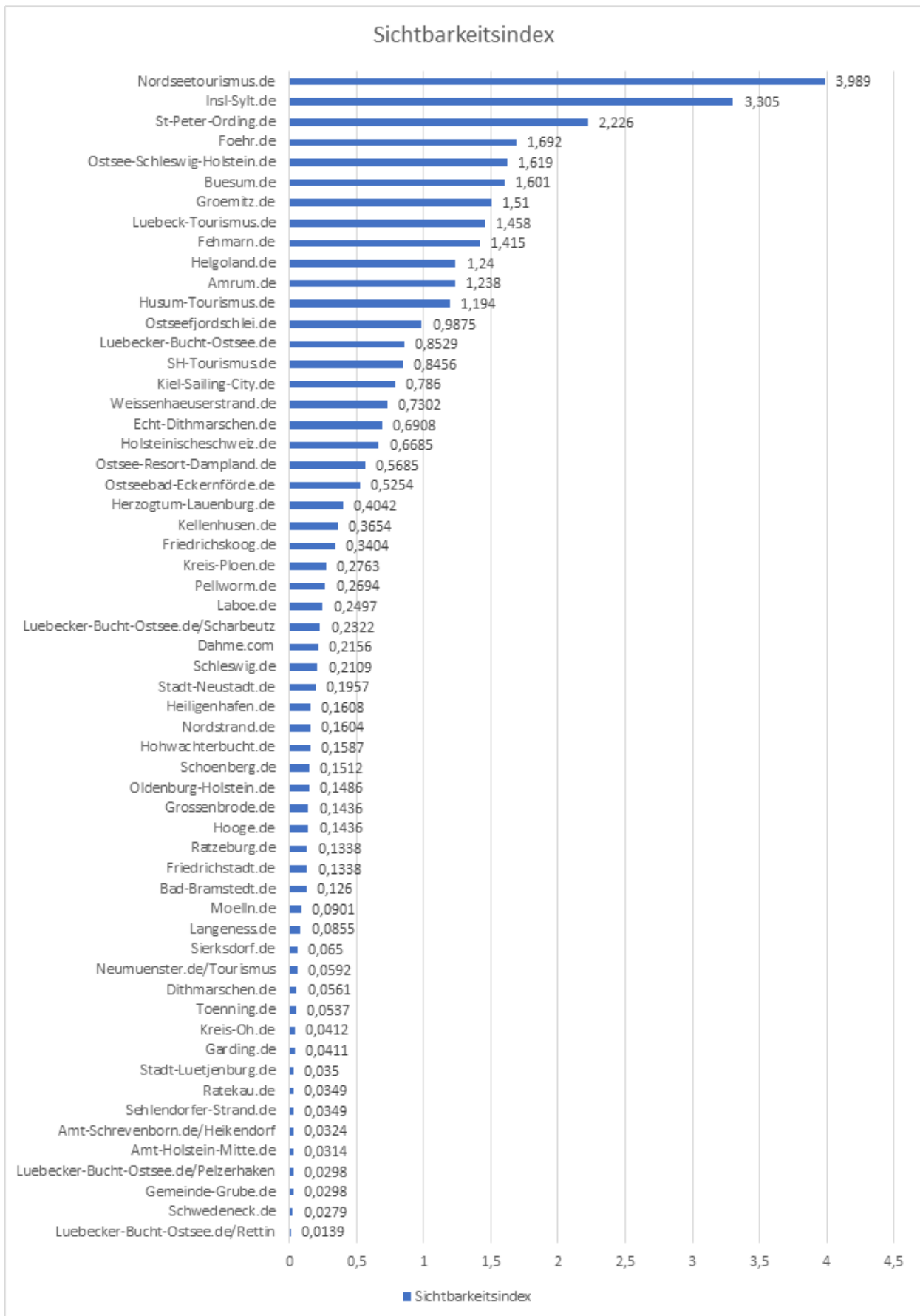


Abbildung 97: Sichtbarkeitsindex von Destinationswebsites in Schleswig-Holstein

Quelle: Eigene Darstellung basierend auf SISTRIX

Implikationen für die Analyse der Informationskanäle

Es wird deutlich, dass insbesondere die beiden Tourismusmarketingorganisationen (TMO) der Nordsee- und der Ostseeküste Nordseetourismus.de und Ostsee-Schleswig-Holstein.de eine hohe Reichweite besitzen. Daneben weisen die touristisch stark relevanten Destinationen wie St. Peter-Ording, Büsum und Grömitz sowie die Inseln Sylt, Föhr, Fehmarn und Helgoland eine hohe Sichtbarkeit auf. Mit Lübeck ist zudem eine Stadt unter den zehn sichtbarsten touristischen Websites in Schleswig-Holstein vertreten.

Parkplätze an der Ostsee

Vor allem in der Saison sind die Parkplätze in Grömitz und auf Fehmarn sehr stark ausgelastet bzw. überlastet. Es kommt zu einem erhöhten Parkplatzsuchverkehr, da es derzeit noch kein digitales Parkleitsystem oder andere mobile Lösungen gibt. Autofahrer*innen müssen aufgrund der fehlenden Auslastungsinformationen vermehrt verschiedene Parkplätze anfahren, bis eine freie Parkbuchung gefunden werden kann. Ein Ansatz für eine problemorientierte Ausspielung des Recommenders wären daher Websites, die Daten zu vorhandenen Parkplätzen führen oder die Online-Bezahlung von Parkplätzen ermöglichen. Aus diesem Grund wurde eine Sichtbarkeitsanalyse entsprechender Anbieter durchgeführt, die auf Basis einer Online-Recherche identifiziert wurden (vgl. Abb. 98).

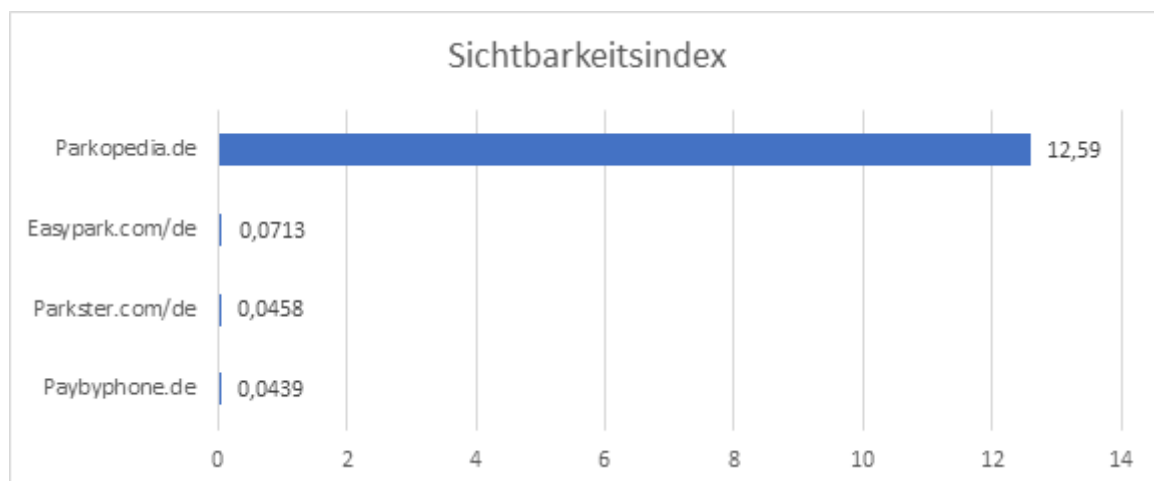


Abbildung 98: Sichtbarkeitsindex von relevanten Informationskanälen/Websites für Parkplätze an der Ostsee

Quelle: Eigene Darstellung basierend auf SISTRIX

Implikationen für die Analyse der Informationskanäle

Es wird deutlich, dass Parkopedia mit Abstand die höchste Relevanz besitzt. Gleichwohl analysiert der Sichtbarkeitsindex nur Websites und Anbieter wie Easypark oder Parkster sind vornehmlich app-basiert, was den geringen Sichtbarkeitsindex erklärt, aber nicht zwingend deren Relevanz insgesamt. Dementsprechend kann es ein lohnenswerter Ansatz sein, auf diese Anbieter zuzugehen. Insbesondere auch deshalb, weil Anbieter wie Parkplatz-Finder und HERE WeGo im Sichtbarkeitsindex von SISTRIX lediglich 0-Werte erhalten, aber dennoch eine entsprechende Relevanz und Verbreitung haben, die genutzt werden kann.

Wetter- und Windwebsites

Eine der beliebtesten Aktivitäten vor Ort ist nach Aussage mehrerer Praxispartner*innen das Windsurfen und Kitesurfen. Um diesen Sportarten erfolgreich nachgehen zu können, ist vor allem die Windsituation entscheidend. In Abbildung 99 wurde daher auch eine Sichtbarkeitsanalyse für entsprechende Website durchgeführt, die diese Informationen ausspielen, damit auch hier geeignete Ausspielkanäle bewertet und somit möglichst viele potenzielle Besucher*innen erreicht werden können. Die Auswahl wurde hier ebenfalls anhand einer Online-Recherche getroffen, bei der verschiedene Rankings zu den beliebtesten Wetter- und Windwebsites herangezogen wurden.

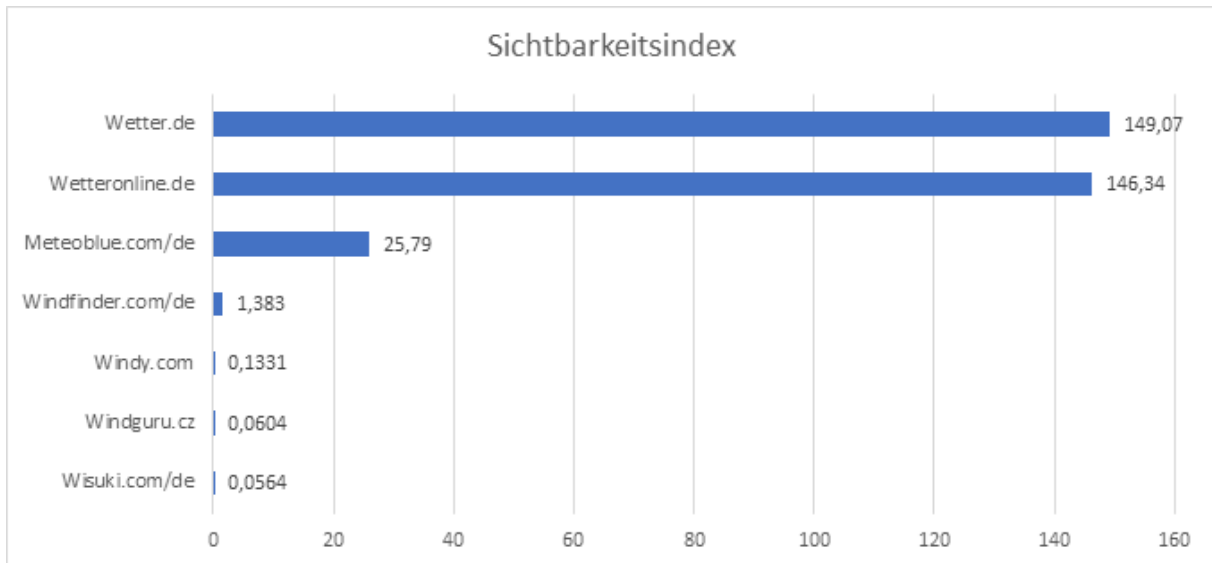


Abbildung 99: Sichtbarkeitsindex von relevanten Informationskanälen/Websites für Parkplätze an der Ostsee

Quelle: Eigene Darstellung basierend auf SISTRIX

Implikationen für die Analyse der Informationskanäle

Es wird deutlich, dass Wetter-Websites eine mit weitem Abstand höhere Sichtbarkeit haben als Wind-Websites. Dies ist vergleichbar mit der Nutzung von allgemeinen Diensten wie Google Maps für alle Formen von Navigation und die themenspezifische Navigation mit Hilfe von Tourenanbietern (outdooractive) für bestimmte Aktivitäten wie Wandern oder Radfahren. Da es vermutlich schwierig ist, einen lokal begrenzten Recommender auf Websites wie Wetter.de zu integrieren, wären Bemühungen der Integration auf speziellen Angeboten von Windfinder, Windy oder Windguru durchaus lohnend. Auch hier sollte beachtet werden, dass diese Anbieter stark von ihren Apps abhängig sind und darüber Sichtbarkeit erlangen und daher der SISTRIX Sichtbarkeitsindex hier nur bedingt aussagekräftig ist.

Wohnmobilstellplätze in der Holsteinischen Schweiz

Bei dem Wohnmobil Use Case besteht die Problemstellung darin, dass sowohl die Auslastungen der einzelnen Wohnmobil- und Campingplätze als auch mögliche Alternativen belegter Stellplätze derzeit nicht zentral zur Verfügung gestellt werden können. Dadurch fahren Besucher*innen in Zeiten hoher Nachfrage häufig verschiedene Standorte an, um einen verfügbaren Platz zu finden. Das hat zur Folge,

dass sich das Verkehrsaufkommen deutlich vermehrt. Damit der Recommender effektiv zu einer Lösung beitragen kann, wurden die zuvor recherchierten Inspirations- und Informationswebsites für die Campingplatzwahl einer Sichtbarkeitsanalyse unterzogen, um geeignete Ausspielkanäle zu identifizieren (vgl. Abb.100).

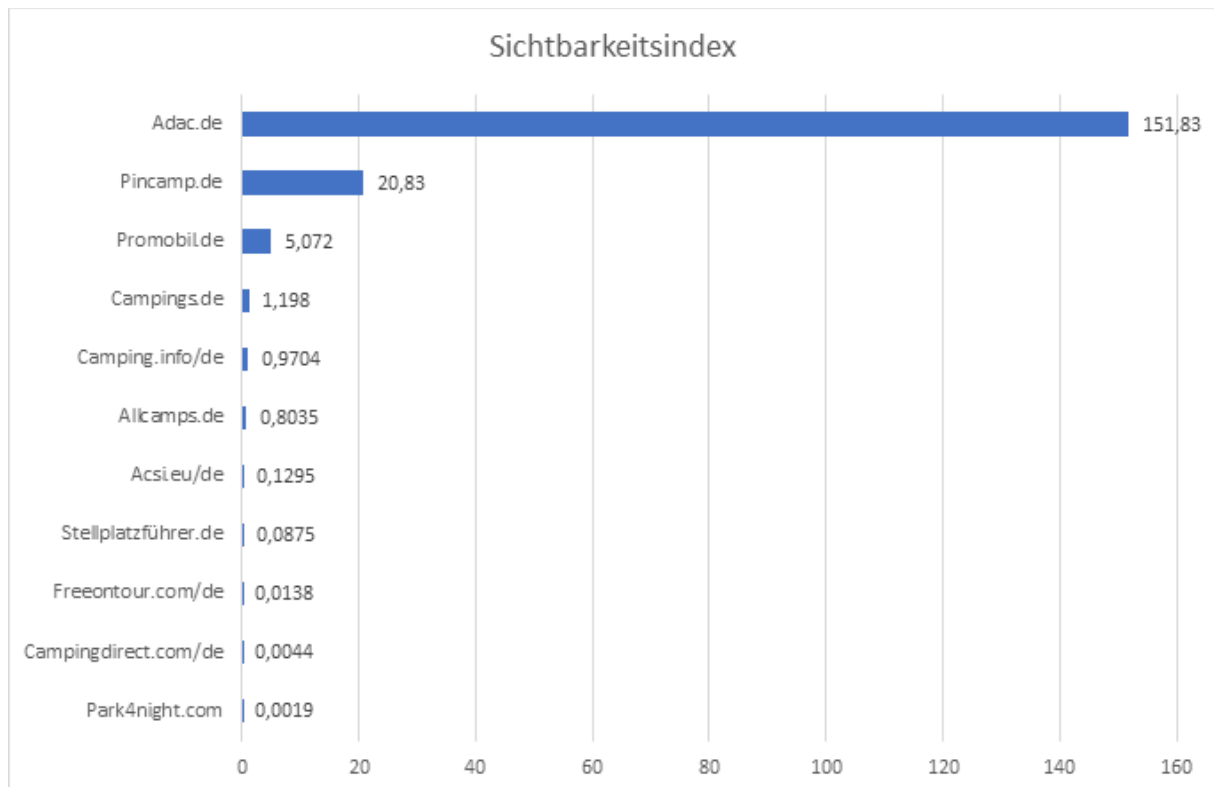


Abbildung 100: Sichtbarkeitsindex von relevanten Informationskanälen/Websites für Wohnmobilstellplätze an der Holsteinischen Schweiz

Quelle: Eigene Darstellung basierend auf SISTRIX

Implikationen für die Analyse der Informationskanäle

Bei den Anbietern für Wohnmobilstellplätze und Campingplätze kann ebenfalls konstatiert werden, dass es sehr große Anbieter gibt, die wenig differenzierte Angebote machen - wie in diesem Fall zum Beispiel der ADAC. Andere Anbieter haben eine eher geringe Sichtbarkeit, sind jedoch eventuell flexibel hinsichtlich der Integration einer Widget-Lösung eines Recommenders. Es kann daher versucht werden, auf große Akteure auch in diesem Segment zuzugehen und die Möglichkeit der Implementierung eines lokalen Recommenders anzubieten. Erfolgsversprechend können aber auch kleine Anbieter sein, die dann jedoch mitunter eine höhere Flexibilität aufweisen - aber auch begrenzte finanzielle Mittel haben, sodass die Implementierung entsprechend einfach möglich sein sollte.

Radrouten und Hotspots an der Schlei

In der Region Schlei gibt es eine Vielzahl an Themen- und Radwegen für Radfahrende, die jedoch eine stark variierende Nachfrage aufweisen. Die an der Schlei gelegenen Radwege sind mitunter so stark belastet, dass Radfahrende nicht aneinander vorbeifahren können, während andere Radwege, die im

ländlichen Raum gelegen sind, ein ausreichendes Platzangebot zur Verfügung haben. Eine ähnliche Problemstellung ergibt sich bei den Hotspots an der Schlei. Ein starker Anstieg der Übernachtungskapazitäten in der Region hat zu der Entwicklung von Orten mit hohem Besuchsdruck geführt. Das Ziel des Recommenders ist demnach bei beiden Use Cases, die Besucher*innen vermehrt in den ländlichen Raum zu verteilen und so den Besuchsdruck in Schleinähe zu verringern. Ermöglicht werden soll das unter anderem, indem den Besucher*innen alternative Ausflugsmöglichkeiten und Routen angeboten werden. Basierend auf den zuvor durchgeführten Recherchen zu Informationsquellen von Radfahrenden während der Reise und der Nutzung von Apps beim Wandern, wurde diese Websites daher einer Sichtbarkeitsanalyse unterzogen (vgl. Abb. 101).

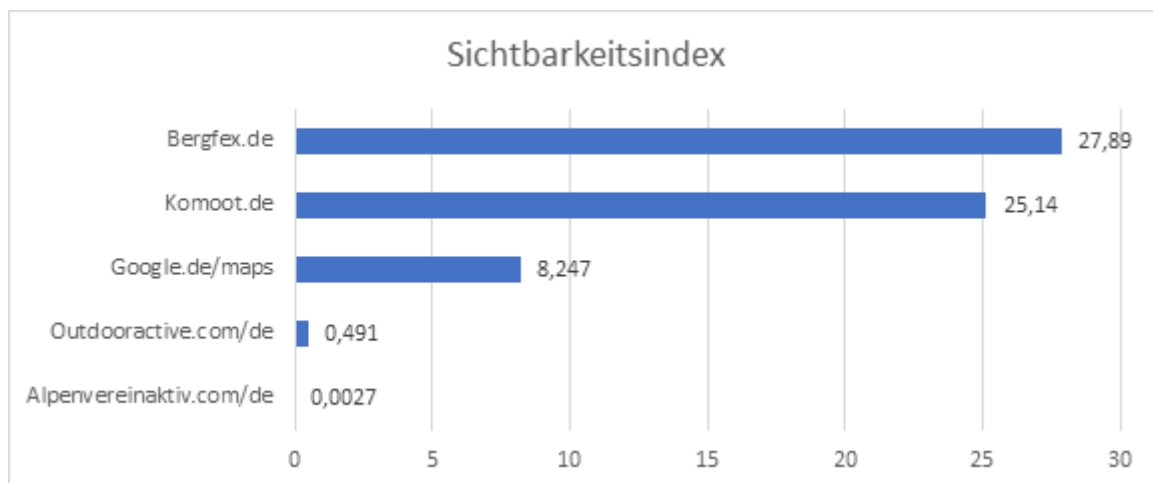


Abbildung 101: Sichtbarkeitsindex von relevanten Informationskanälen/Websites für Radrouten und Hotspots an der Schlei

Quelle: Eigene Darstellung basierend auf SISTRIX

Implikationen für Analyse der Informationskanäle

Da bereits bei der Relevanzabschätzung mehrere Studien zur Verbreitung und Nutzung von Tourenplanern durchgeführt wurde, sind diese Werte nun lediglich ergänzend. Dies gilt insbesondere auch deshalb, weil auch hier ein verzerrtes Bild im Rahmen einer Sichtbarkeitsanalyse entstehen kann. So ist die Sichtbarkeit von Google Maps allein schon deshalb höher, weil bei vielen Suchanfragen die Kartendarstellung in den Suchergebnisseiten mit ausgegeben wird. Dazu haben sowohl Komoot als auch Outdooractive eine komplementäre App und Touren von Outdooractive werden bei Anbietern wie Bergfex mit ausgegeben. Dennoch kann festgehalten werden, dass es mit dem Alpenverein und anderen Anbietern Anhaltspunkte gibt, wo der Versuch gestartet werden könnte, Live-Daten zur Auslastung auszuspielen.

Beispiele für die Darstellung von Ausspielkanälen

Es gibt bereits diverse Beispiele, wie touristische Daten in Form von Apps dargestellt werden können. In Teilen haben diese auch Informationen zum Besuchsaufkommen integriert und bieten daher eine gute Grundlage für die Gestaltung eines Piloten. Im Anschluss werden verschiedene, im Tourismus bekannte, diskutierte und verbreitete Systeme vorgestellt und diskutiert.

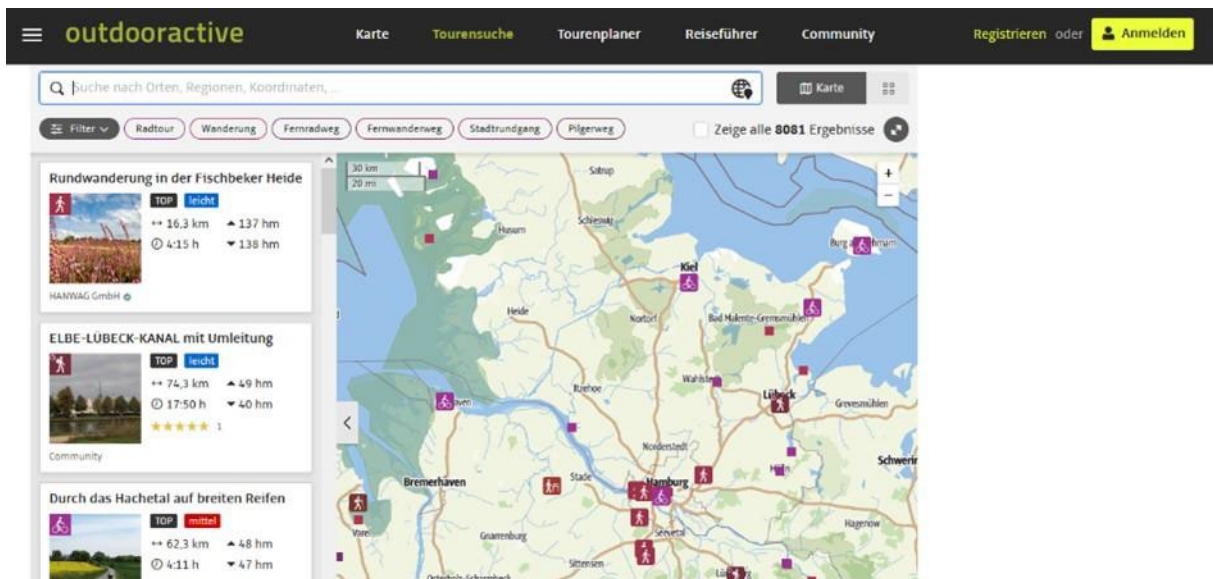


Abbildung 102: Tourensuche bei Outdooractive

Quelle: www.outdooractive.com/de

Die Anwendung von Outdooractive bietet die Möglichkeit, Touren in unterschiedlichen Regionen zu suchen. Abbildung 102 zeigt, wie anhand einer Karte sowohl angezeigt wird, in welchem geografischen Gebiet die Route verläuft als auch um welche Art (Radtour, Wandertour etc.) es sich hierbei handelt. Gleichmaßen werden die wichtigsten Eckdaten der Tour wie Länge, Dauer und Höhenmeter auf den ersten Blick angezeigt. Nutzer*innen können zudem Filter einsetzen, um die Suchergebnisse auf ihre Präferenzen einzuschränken.

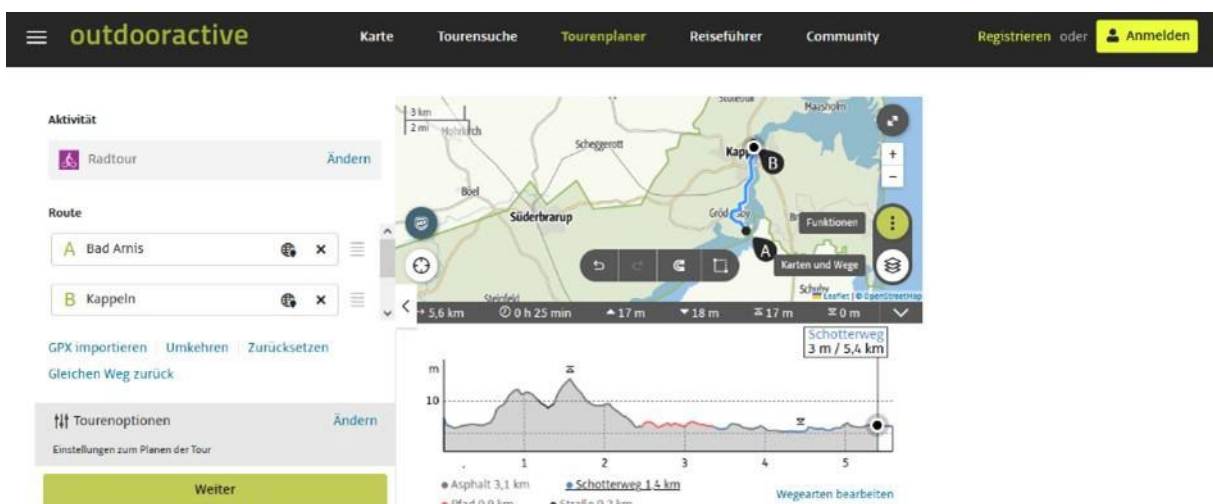


Abbildung 103: Tourenplaner bei Outdooractive

Quelle: www.outdooractive.com/de

Weiterhin ist es bei Outdooractive möglich, eine Tour mittels eines Start- und Endpunktes zu planen (vgl. Abb. 103). Auf Grundlage der Art der Tour wird eine passende Route erstellt und es werden zusätzlich Informationen zur Routenbeschaffenheit angegeben.

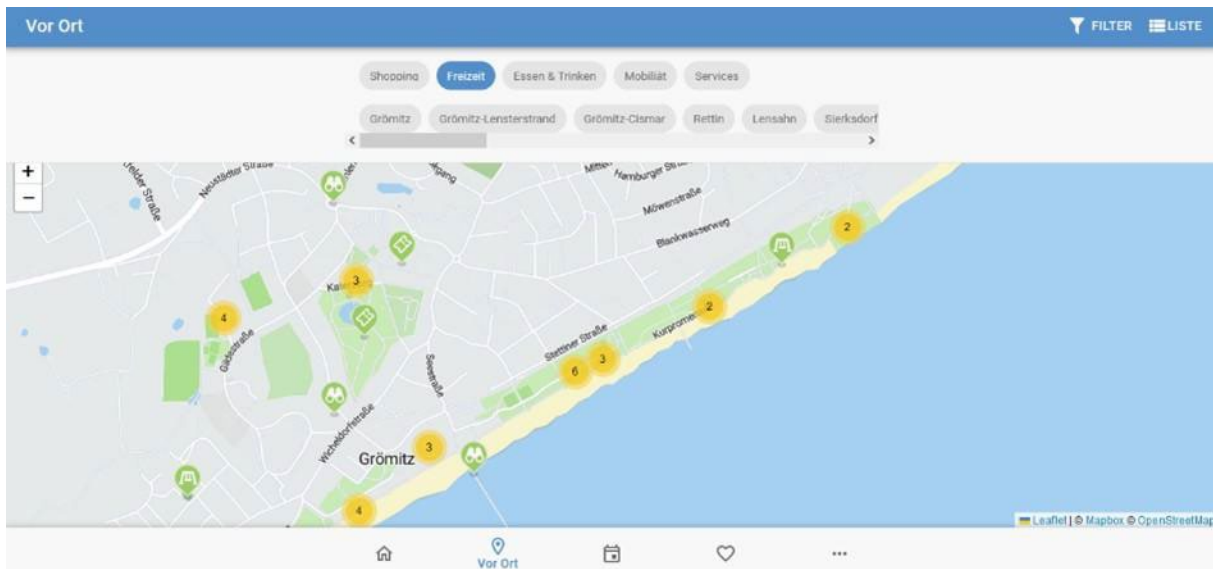


Abbildung 104: Lübecker Bucht Guide

Quelle: www.luebecker-bucht.guide

Der Lübecker Bucht Guide ist eine Anwendung, welche verschiedene POI innerhalb der Region ausspielt, wie in Abbildung 104 ersichtlich ist. Durch eine Filteroption kann nach verschiedenen Aktivitäten wie Shopping, Freizeit etc. gesucht werden, die anschließend auf eine Karte bzw. in Form einer Liste angezeigt werden können.

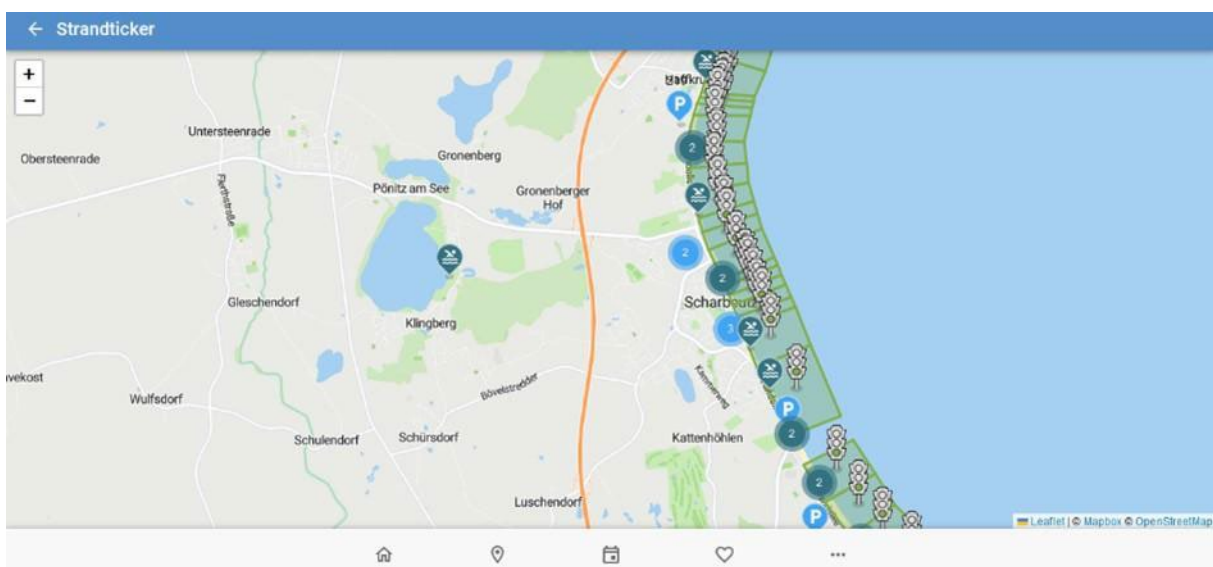


Abbildung 105: Strandticker des Lübecker Bucht Guides

Quelle: www.luebecker-bucht.guide/beachticker

Eine Besonderheit des Lübecker Bucht Guides ist der Strandticker, der Nutzer*innen die derzeitige Auslastung verschiedener Strandabschnitte in Niendorf, Scharbeutz, Haffkrug, Sierksdorf, Neustadt in Holstein und am Timmendorfer Strand auf Basis eines Ampelsystems darstellt (vgl. Abb. 105). Während eine grüne Ampel auf eine geringe Auslastung hindeutet, ist bei Orange von einer höheren Auslastung und bei Rot von einer kompletten Auslastung des entsprechenden Strandabschnittes zu rechnen.

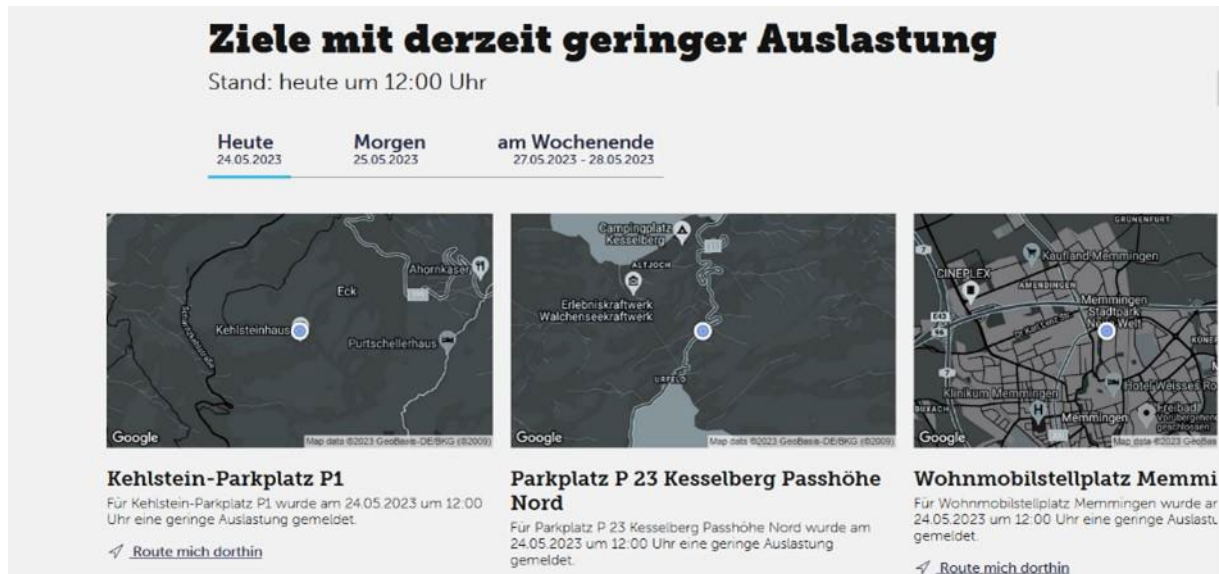


Abbildung 106: Ausflügsticker Bayern

Quelle: www.ausflügsticker.bayern

Eine weitere Anwendung ist der Ausflügsticker Bayern, der in Abbildung 106 dargestellt ist. Dieser zeigt zunächst Ziele in Bayern mit derzeit geringer Auslastung für den derzeitigen Tag an und gibt weiterhin Prognosen zur Auslastung für den darauffolgenden Tag sowie das Wochenende.

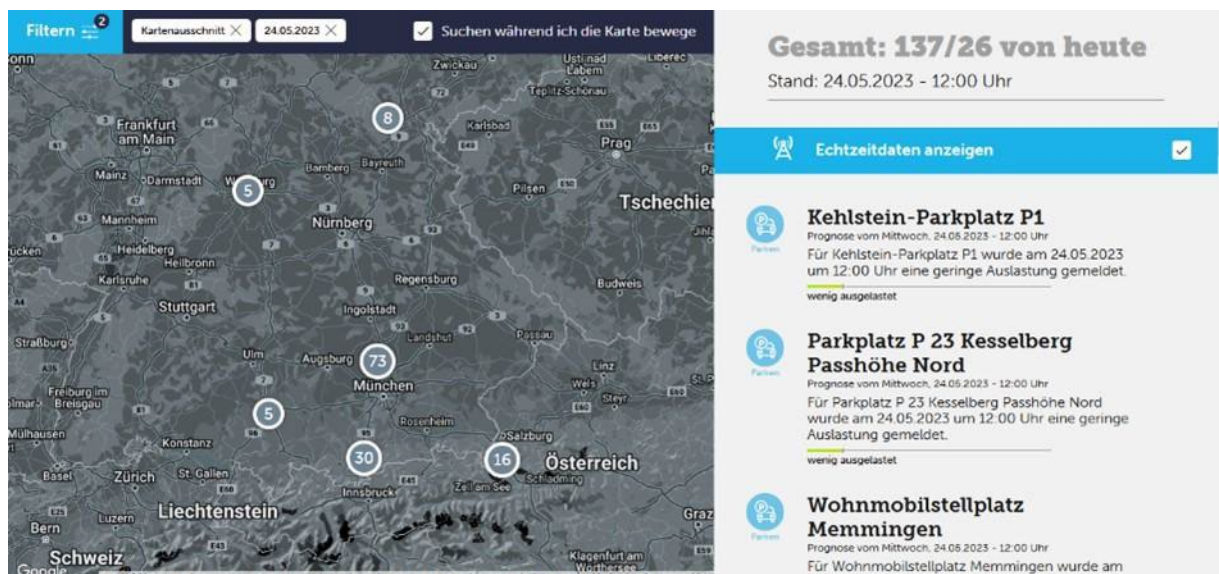


Abbildung 107: Ausflügsticker Bayern

Quelle: www.ausflügsticker.bayern/

Weiterhin bietet der Ausflugsticker Bayern die Möglichkeit, anhand von Echtzeitdaten die Auslastung verschiedener POI auf einer Karte anzuzeigen (vgl. Abb. 107). Um die Suche zu individualisieren, können die Meldungen gefiltert werden, beispielsweise nach der Auslastung (Art) auf Parkplätzen (POI).

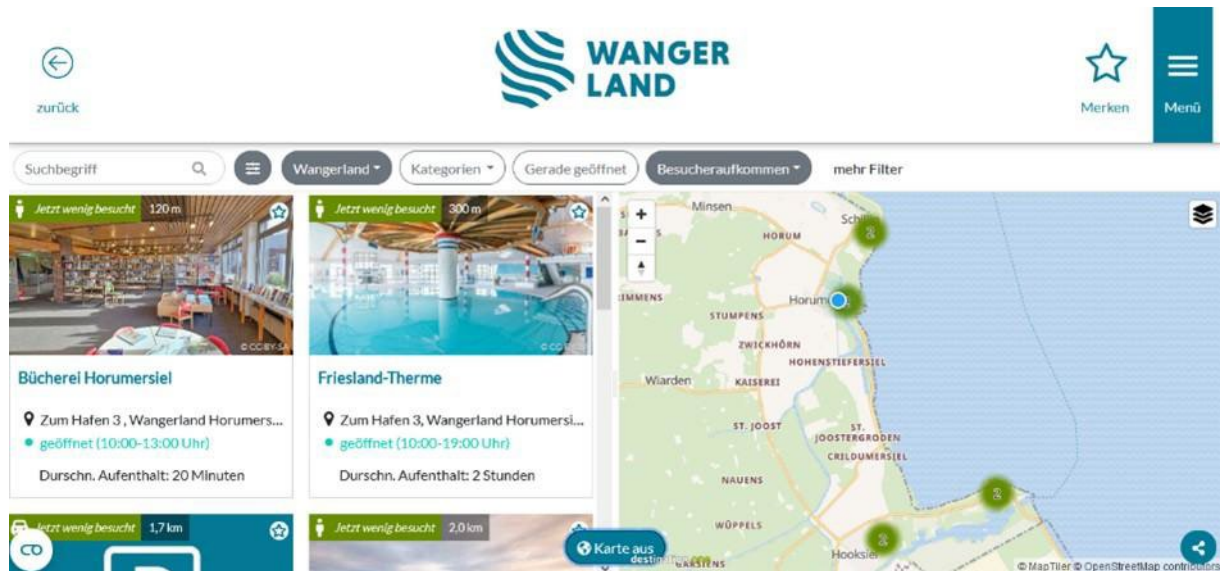


Abbildung 108: Digitaler Reiseführer Wangerland

Quelle: www.wangerland.de

Das Wangerland hat ebenfalls eine Anwendung, welche sowohl verschiedene Veranstaltungen in der Region als auch das aktuelle Besucheraufkommen ausspielt, wie Abbildung 108 zeigt. Auf einer Karte können die verschiedenen Vorschläge verortet werden und eine korrelierende Liste gibt zusätzlich Informationen zu dem vorgeschlagenen POI. Auch hier können die Aktivitäten zuvor wieder durch eine Filteroption eingeschränkt werden.

Seenotrettungsmuseum Fehmarn

Adresse: Burgstaaken 89, 23769 Fehmarn
 Öffnungszeiten: Geöffnet · Schließt um 17:00
 Telefon: 0171 5668873

4,4 ★★★★★ 103 Google-Rezensionen
 Museum auf Fehmarn, Schleswig-Holstein

Museum zum Thema "150 Jahre Seenotrettung" mit Rettungskreuzer und Anschauungsmaterial zur Rettungstechnik.

Stoßzeiten

MO DI MI DO FR SA SO

12 Uhr: Normalerweise mäßig besucht

09 Uhr 12 Uhr 15 Uhr 18 Uhr 21 Uhr

Abbildung 109: Google Popular Times für das Seenotrettungsmuseum auf Fehmarn

Quelle: www.google.com/de

Abbildung 109 zeigt Google Popular Times als eine Zusatzfunktion von Google, die die Auslastung an unterschiedlichen POI anzeigt. Die Stoßzeiten können für jeden Tag der Woche für einen ganzen Tagesverlauf eingesehen werden. Diese Auskunft wird den Nutzer*innen bei der Suche nach POI automatisch zu den allgemeinen Informationen angezeigt.

Zurück zur Auswahl Zum Wandern! Liste anzeigen

Jetzt Heute Filter

Pölvener Runde

Mittel Gehzeit: 2 Std
 Anreise: 37 Min

Über die Märchenwiese zur Gaudeamushütte

Mittel Gehzeit: 4 Std

Ab 12:59 An 13:11 Dauer 12 min

Abbildung 110: Tourensuche bei Naturtrip für die Region Wilder Kaiser

Quelle: <https://wilderkaiser.naturtrip.app/>

Naturtrip bietet Web-Apps an, um bei der Planung von Ausflügen mit Bus und Bahn in Niedersachsen zu unterstützen (vgl. Abb. 110). Es werden ausschließlich Ziele angezeigt, die unter der Berücksichtigung von Startpunkt und Wochentag gut zu erreichen sind. Naturtrip hat dieses Konzept unter anderem für die Region Wilder Kaiser umgesetzt. Zunächst kann ein Thema ausgewählt werden, z.B. wandern. Anschließend besteht die Möglichkeit weitere Filter, wie den Zeitpunkt, die Wander-Dauer, den Schwierigkeitsgrad, die maximale Reisezeit etc. auszuwählen. Daraufhin werden wie in der obigen Abbildung zu sehen Vorschläge angezeigt, entweder in der Karten- oder Listenansicht. In der Detailansicht sind zudem Informationen zum Ziel wie Öffnungszeiten und Einkehrmöglichkeiten enthalten. Zudem werden Tickettipps, die nächste Verbindung sowie die CO2 Einsparung im Vergleich zur Anreise mit dem Auto ausgegeben. Die Informationen zu den Ausflugszielen, Veranstaltungen, Wanderwegen etc. werden vom Tourismusverband bereitgestellt und die Fahrplandaten vom regionalen Verkehrsbund.

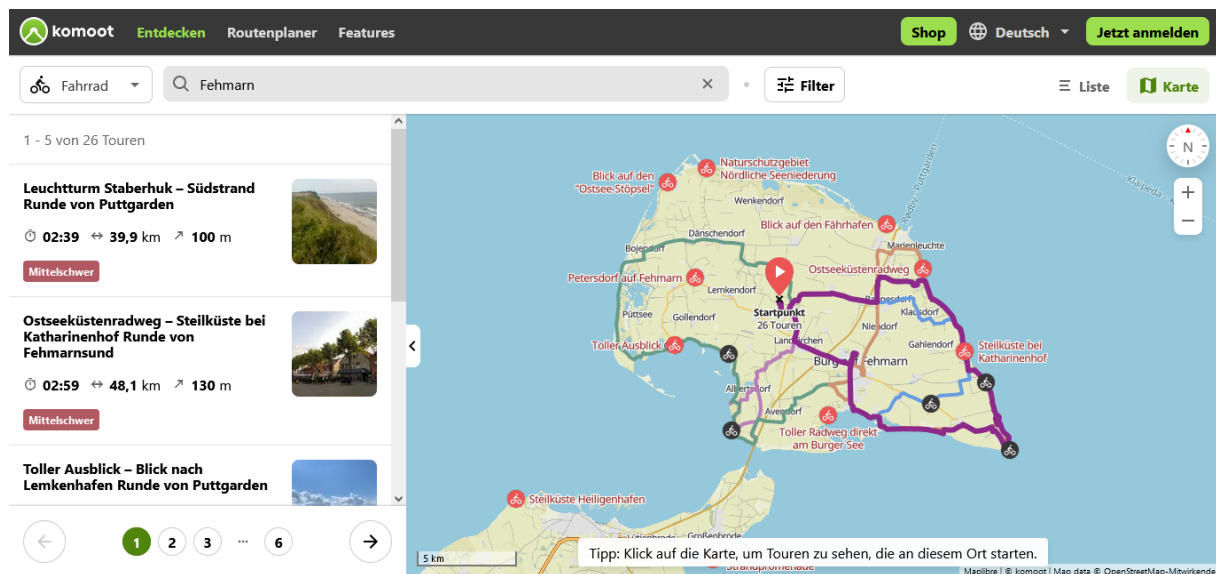


Abbildung 111: Tourensuche bei Komoot

Quelle: <https://www.komoot.de/>

Komoot hat sich der Tourenplanung und Navigation verschrieben. In ihrer Anwendung besteht die Möglichkeit, sich für eine Sportart und einen Ort Vorschläge für Touren anzeigen zu lassen (vgl. Abb. 111). Zudem können weitere Filter wie Dauer, Schwierigkeitsgrad, Höhenmeter, die bevorzugte Web-Oberfläche oder der Routentyp ausgewählt werden. Die Touren werden sowohl in Karten- als auch Listenansicht zur Verfügung gestellt. Zudem sind eine Sprachnavigation und Offline-Karten enthalten. Touren können ebenfalls gespeichert und mit Freunden geteilt werden. Komoot kann auf einer Vielzahl von Geräten, von GPS-Geräten über Smartwatches bis hin zu E-Bikes genutzt werden.

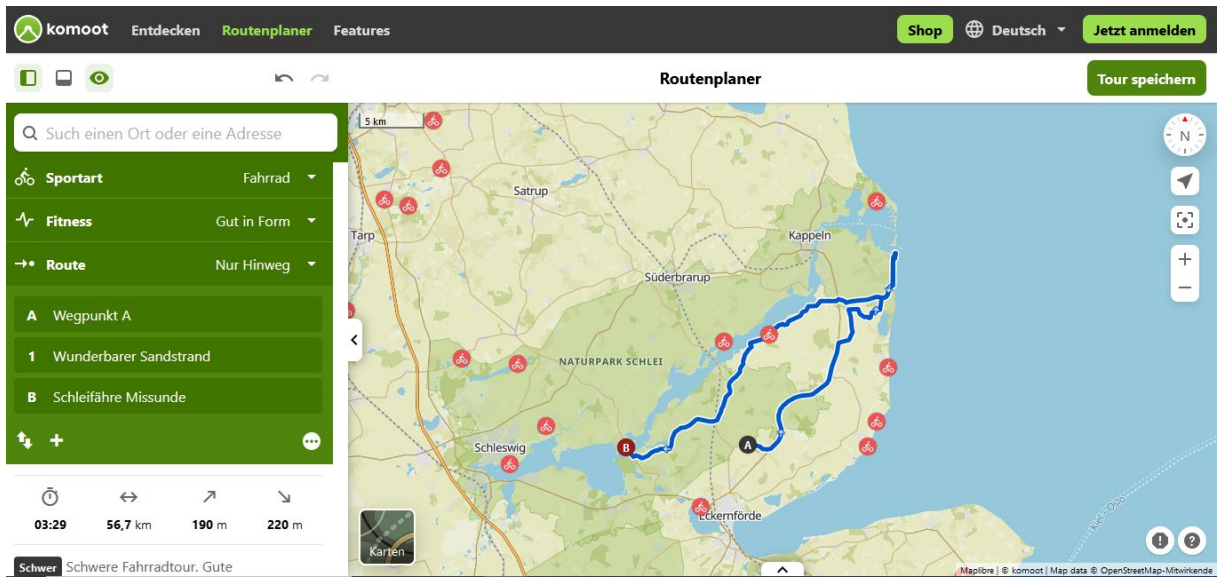


Abbildung 112: Routenplaner bei Komoot

Quelle: <https://www.komoot.de/>

Mit komoot besteht zudem die Möglichkeit, maßgeschneiderte Routen anzufertigen, wie Abbildung 112 zeigt. Auch hier können eine Sportart sowie weitere Filteroptionen ausgewählt werden. Neben dem Start- und Endpunkt können weitere individuelle Stopps hinzugefügt werden.

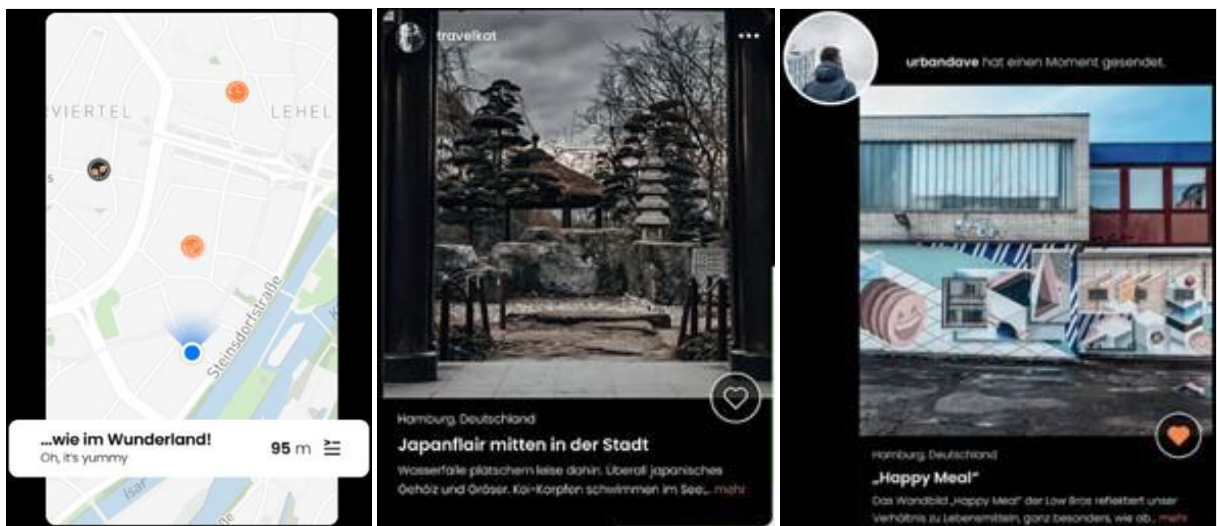


Abbildung 113: UrbnUps

Quelle: <https://urbnups.com/>

Abbildung 113 zeigt die App UrbnUps, die für iPhone und Android-Smartphones genutzt werden kann und sowohl Geheimtipps von Einheimischen als auch Geschichten zu bekannteren Orten bietet. Es besteht die Möglichkeit, Dinge zu speichern, Notizen zu machen und diese ebenfalls mit Freunden oder öffentlich zu teilen. Die Interaktion mit der Community besteht neben dem Teilen von Erlebnissen in der Vergabe von UPs.

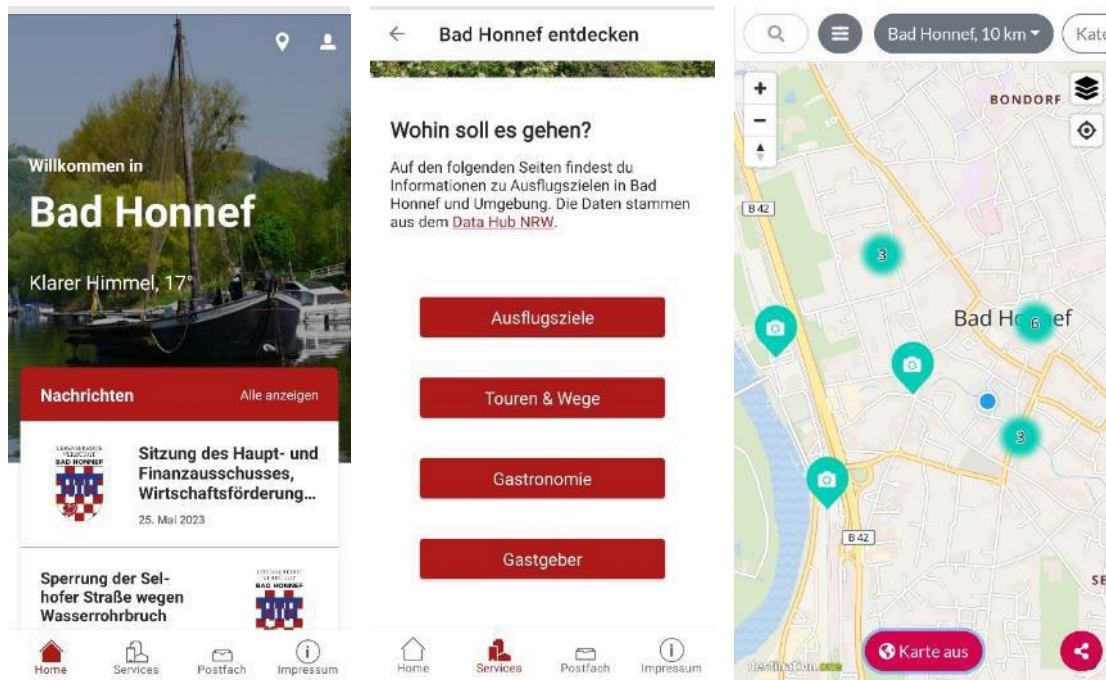


Abbildung 114: CityKey

Quelle: <https://citykey.app/>

Die CityKey App von Telekom bietet einen einfachen und digitalen Bürgerservice über das Smartphone. CityKey wird jeweils gemeinsam mit Städten und Gemeinden in ganz Deutschland entwickelt (vgl. Abb. 114). Enthaltene Funktionen sind unter anderem die digitale Verwaltung wie z.B. das Ausfüllen von Formularen, das Buchen von Terminen für das Amt, die Bürgerbeteiligung per Umfragen, Mängelmelder, Tipps & Infos zu Orten in der Stadt, ein Müllabfuhrkalender, Veranstaltungen und tagesaktuelle Nachrichten aus Rat und Stadtverwaltung. Die Informationen zu den Ausflugszielen in dem Beispiel Bad Honnef stammen aus dem Data Hub NRW. Die Ausflugsziele können nach dem Umkreis, Kategorien sowie „schlechtwettergeeignet“ gefiltert werden. Es kann entweder eine Karten- oder Listensicht ausgewählt werden.

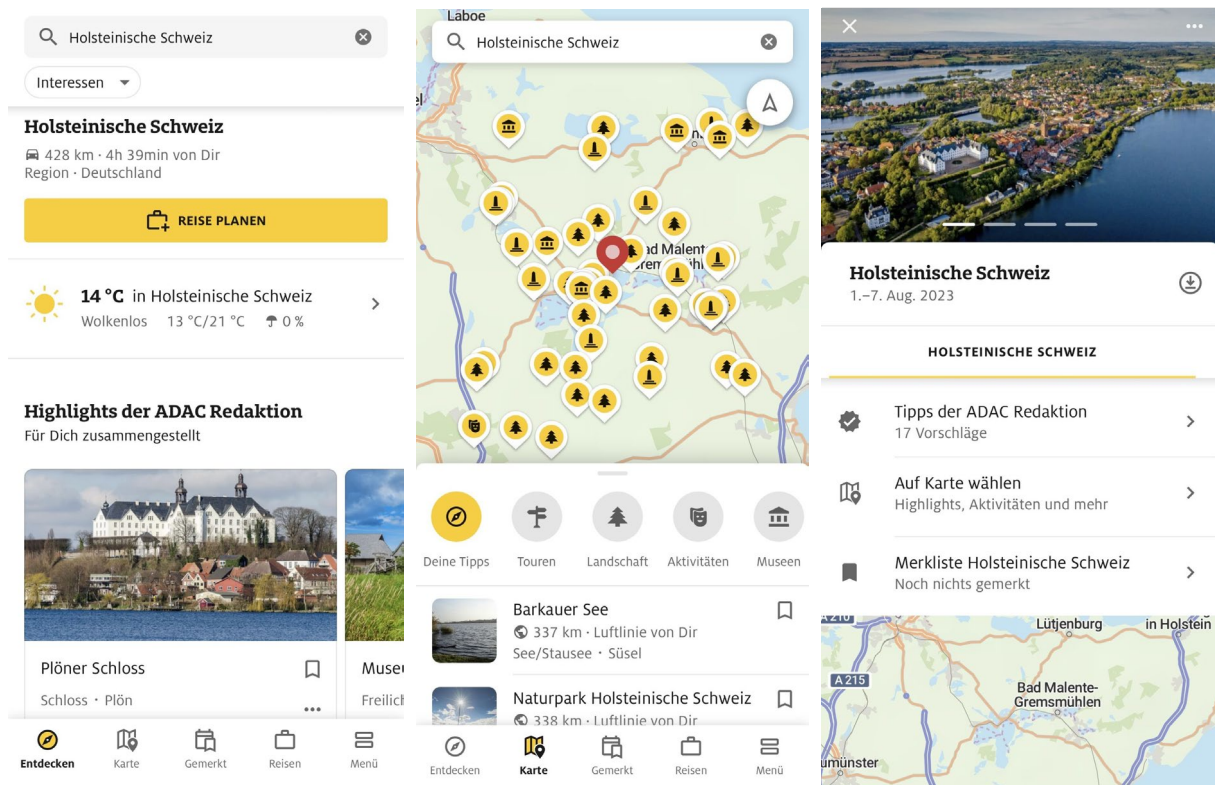


Abbildung 115: ADAC Trips App

Quelle: <https://www.adac.de/services/apps/trips/>

In der ADAC Trips App werden individualisierte Vorschläge für Sehenswürdigkeiten, Touren, Aktivitäten und Restaurant-Tipps, basierend auf den Interessen angezeigt (vgl. Abb. 115). Per Swipe nach rechts kann festgelegt werden, was einem gefällt. Auf Basis dessen generiert der Algorithmus Vorschläge. Die POI und Touren werden in einem selbst ausgewählten Umkreis von bis zu 100 km inklusive 14-Tage Wettervorhersage angezeigt. Die Informationen stammen von Partner*innen aus der Region. Das Urlaubsplanungs-Tool enthält zudem einen Offline-Modus und es besteht die Möglichkeit, Erlebnisse aus der App heraus zu buchen. Eine weitere Funktion besteht in dem Corona Radar. Es wird sowohl die gültige Warnstufe als auch die Krankenhaus-Ampel, Hotspots, der Inzidenz-Trend und geltende Maßnahmen angezeigt.

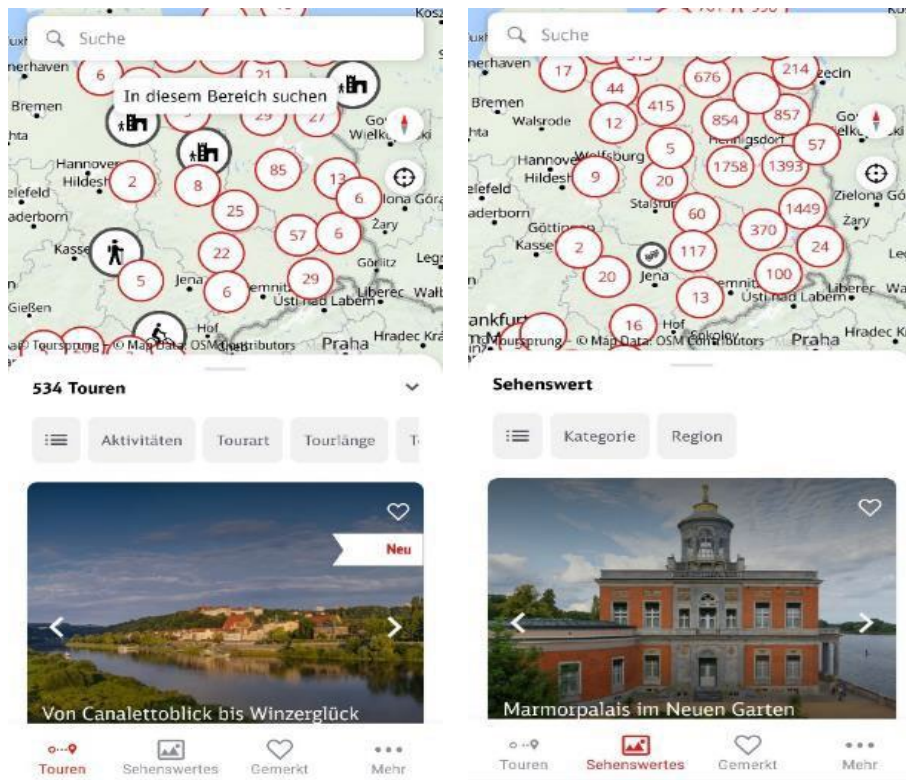


Abbildung 116: DB Ausflug App

Quelle: <https://www.dbausflug.de/>

Die DB Ausflug App ist ein Tourenplaner inklusive kultureller und kulinarischer Tipps, Offline-Karten zum Routing, Hinweise zu Sehenswerten und Öffnungszeiten sowie wichtigen Reiseinformationen (vgl. Abb. 116). Es sind über 500 vorgefertigte und geprüfte Touren enthalten. Es besteht die Möglichkeit die Touren neben der Region nach einer Reihe weiterer Kategorien zu filtern.



Abbildung 117: ByTheWays App

Quelle: <https://www.bytheways.io/>

ByTheWays ist eine App für die Routenplanung von Roadtrips abseits der Massen. Es sind sowohl Rundreisen, als auch Kurztrips und Ausflüge enthalten (vgl. Abb. 117). Es werden Sehenswürdigkeiten und Orte abseits der Hotspots vorgeschlagen. Nachdem ein Start- und Zielpunkt eingegeben wurde, werden Stopps entlang der Route vorgeschlagen. Während der Fahrt werden in Echtzeit Ziele entlang der Route vorgeschlagen.

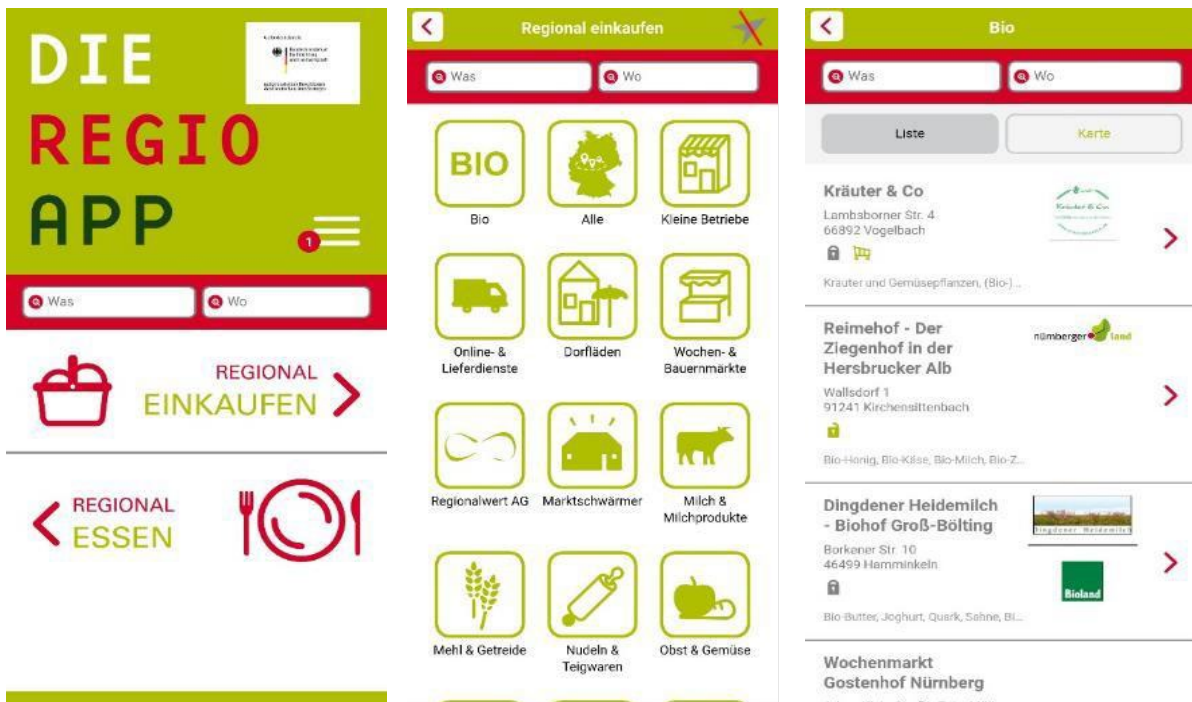


Abbildung 118: RegioApp

Quelle: <https://www.regioapp.org/>

Die RegioApp in Abbildung 118 bietet eine Umkreissuche nach regionalen Lebensmitteln und regionalem Essen. Angezeigt werden direktvermarktende Betriebe und Gastronomie, aber auch Verkaufsstellen wie Dorfläden, Lebensmitteleinzelhändler mit regionalen Produkten und Wochenmärkte. Nachdem angegeben wurde, was gesucht wird und wo es gesucht wird, werden alle Betriebe nach der Entfernung gelistet. Es sind ebenfalls detailliertere Infos zu den Betrieben, wie Öffnungszeiten, die angebotene Produktpalette usw. enthalten. Der Weg zu dem gewählten Betrieb ist auf einer Straßenkarte, mit integrierter Navigationsfunktion, ersichtlich.

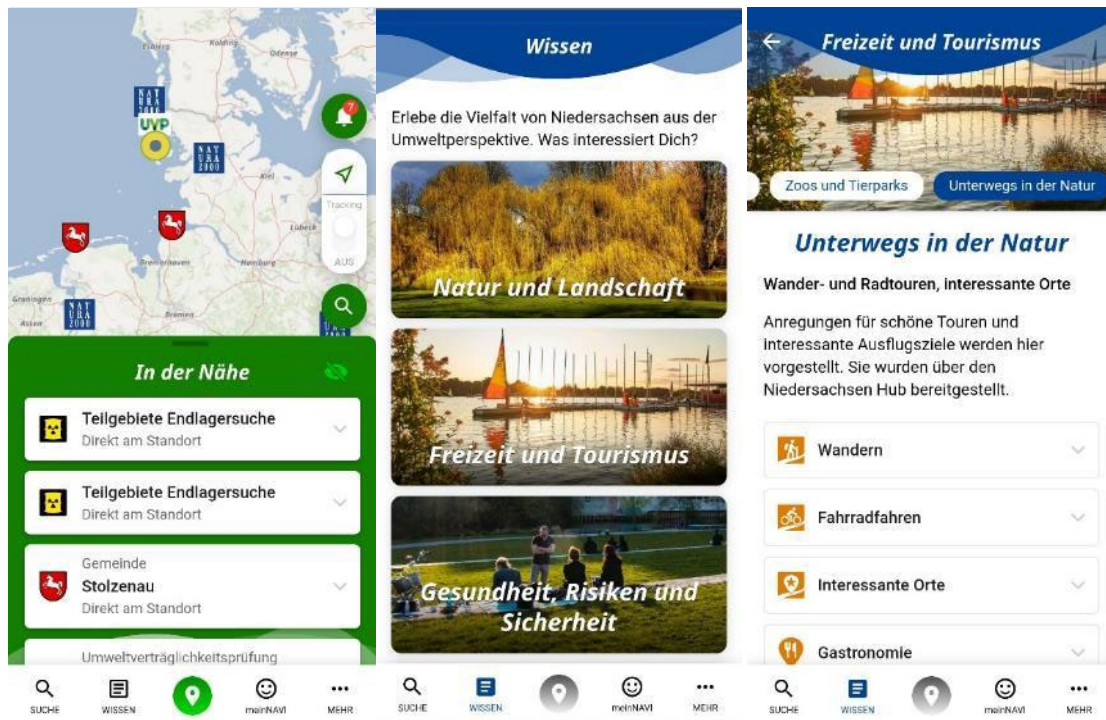


Abbildung 119: UmweltNAVI Niedersachsen App

Quelle: <https://umwelt-navi.info/>

Die UmweltNAVI Niedersachsen App bietet die Möglichkeit nach Umweltobjekten zu suchen und enthält ebenfalls eine Navigation zu den Objekten (vgl. Abb. 119). Enthaltene Informationen sind beispielsweise Hochwassergebiete, Messwerte zu Pegeln, Luft und Lärm, Badestellen oder Naturparke, Industrieanlagen, Wolfsvorkommen oder Radioaktivitätsmessstationen. Es besteht die Möglichkeit Fotos hochzuladen und diese mit anderen zu teilen. Zudem sind unter „Wissen“ Informationen zu unterschiedlichen Kategorien enthalten. Innerhalb der Kategorien gibt es weitere Unterkategorien. Bei Freizeit und Tourismus besteht beispielsweise die Möglichkeit, sich Wanderrouten anzeigen zu lassen. Die Informationen werden über den Niedersachsen Hub bereitgestellt.

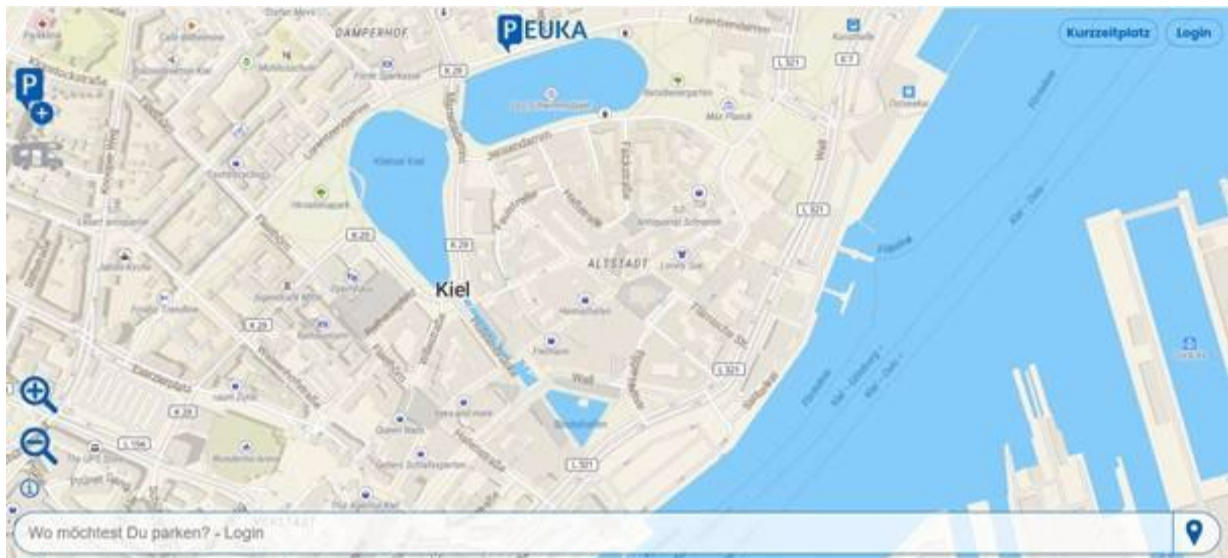


Abbildung 120: Peuka Parkplatz-Sharing App

Quelle: <https://www.peuka.com/>

Peuka ist eine App für Parkplatz-Sharing. Ziel ist es, den Parksuchverkehr einzudämmen, indem Zugang zu privaten Parkplätzen gewährt wird (vgl. Abb. 120). Privatpersonen können ihren Parkplatz anbieten, während sie ihn nicht nutzen. Die angebotenen Parkplätze können digital gebucht werden. Es können auch gewerbliche Parkplätze und Unterstellplätze angeboten und gebucht werden.

Implikationen für die Informationskanäle

Auf Basis der Analyse der ausgewählten Anwendungen sind diverse Funktionen zu erkennen, die in die Design-Überlegungen einer Pilotanwendung mit einfließen können:

1. Karten- und Listenansicht

Auffällig ist bei fast allen Anwendungen eine Kombination aus Karten- und Listenansicht in unterschiedlicher Darstellungsform. Daran sollte sich auch eine Pilotanwendung im Rahmen des Projektes orientieren, um den Konventionen anderer Anwendungen gerecht zu werden.

2. Filteroption

Viele Anwendungen bieten die Möglichkeit einer Filterung. Dies ist auch im Rahmen einer Pilotanwendung eine Option, um eine Vorfilterung vornehmen zu können und nur relevante Alternativen zu einem ausgelasteten Ziel anzuzeigen. Gleichzeitig ist es das Ziel nach Möglichkeit die geeigneten Alternativen direkt und ohne aktive Auswahl der Nutzenden zu zeigen. Daher ist eine Filteroption zwar möglich, aber nicht zwingend. Interessant sind jedoch die Art der Filterfunktionen. So sind Angaben zu Schlecht- oder Gutwettereignung direkt mit der Wettervorhersage verknüpfbar. Das bedeutet für einen Empfehlungsdienst, dass die vorliegende Datengrundlagen mit der Filterfunktion korrelieren sollte, damit entsprechende Empfehlungen auch zielgerichtet ausgespielt werden können.

3. Auslastung als Ampelsystem

Typisch ist die Anzeige der aktuellen Auslastung eines POI in Form eines Ampelsystems (grün = leer; rot = voll). Diese wohlbekannten Signalfarben haben einen hohen Wiedererkennungswert und können daher sehr gut eingesetzt werden, um den Füllstand eines POI zu signalisieren und können daher auch als Farbcodes im Rahmen einer Pilotierung eingesetzt werden.

4. Präferierte Anzeige nach Auslastungsgrad

Wie der Ausflugsticker Bayern zeigt ist es im Rahmen einer Listen- oder Kachelansicht möglich und in Teilen auch sinnvoll für die digitale Besucherlenkung, wenn wenig besuchte POI präferiert angezeigt werden. Gleichwohl ist es nicht ratsam, die Auslastung als einziges Filterkriterium heranzuziehen – dies wäre zu trivial und auch nicht der Anspruch der im Rahmen dieses Projektes pilotierten Anwendung. Vielmehr sollte die Rolle weiterer Parameter wie Entfernung, Ähnlichkeit der Aktivität, Erreichbarkeit via ÖPNV usw. geprüft werden.

5. Auslastungsprognosen

Ziel ist es auch, Prognosen für die Zukunft auszugeben. Diese können in unterschiedlicher Form ausgegeben werden. Sehr bekannt ist die Ausgabe in einem Balkendiagramm im Tagesverlauf, wie dies bei Google Popular Times erfolgt. Eine solche Darstellung kann auch für eine Pilotanwendung sinnvoll sein, da sie leicht verständlich ist.

6. Öffnungszeiten

Ein wichtiger Parameter, der stets angezeigt werden sollte ist die Öffnungszeit eines POI. Diese Information kann als kritisch erachtet werden, da eine Falscheingabe dazu führen kann, dass der gesamten Anwendung nicht mehr getraut wird.

7. Erreichbarkeit + Navigation

Viele Anwendungen bieten die Möglichkeit einer personalisierten Routenführung oder Navigation an. Auch die Ausspielung von Zielen, die besonders gut per ÖPNV zu erreichen sind (Naturtripp) ist eine Option. Dabei sollte jedoch bei einer Pilotierung überlegt werden, ob eine solche Funktion zwingend erforderlich ist. Etablierte Anwendungen wie Google Maps sind bei Nutzenden bekannt und bieten vielfältige Optionen. Wenn also ein Ziel seitens der Nutzenden ausgewählt wurde ist es auch möglich, dass die Route dorthin per Verlinkung an eine solche Navigationsanwendung ausgelagert wird.

8. CO2-Einsparung als Nudge

Eine Besonderheit der Anwendung Naturtripp ist die Berechnung der Einsparung des CO₂-Ausstoßes, wenn Reisende anstelle des PKW den ÖPNV wählen. Eine solche Anzeige, die auf die intrinsische Motivation der Nutzenden setzt kann auch im Rahmen einer Pilotierung angedacht werden, wenn sich dies anbietet. Als Icon für alles, was Emissionen einspart hat sich dabei das grüne Blatt etabliert, das auch im Rahmen einer Pilotierung dann Anwendung finden sollte.

9. Offline-Funktionalität

Viele Anwendungen haben eine Offline-Funktionalität, damit diese auch in Gebieten funktionieren, wo die Anbindung ans Internet ggf. nicht gegeben ist. Eine solche Funktionalität sollte im Rahmen einer Pilotierung allerdings nicht im Fokus stehen, denn dies würde vornehmlich die Routenführung betreffen, welche eben auch keine zentrale Funktion im Rahmen der in diesem Projekt anvisierten Pilotanwendung darstellt.

10. Interoperabilität für diverse Endgeräte

Die Anwendung von Komoot zeigt, dass auch die Ausspielung auf Smart Watches sinnvoll sein kann. Im Rahmen einer Pilotierung scheint dies zwar nicht sinnvoll, da der Einsatz eines Empfehlungsdienstes ein anderer ist als eine Routenführung. Gleichwohl ist anzustreben – und das zeigen auch die zuvor vorgestellten Studien – dass eine Pilotanwendung zumindest responsiv auf Desktop, Tablet und Smartphone genutzt werden sollte.

11. Speicherung und Teilen

Viele Anwendungen ermöglichen es, die POI mit anderen zu teilen. Eine solche Funktion kann auch für eine Pilotanwendung in diesem Projekt sinnvoll sein, weil dadurch eine gemeinschaftliche Auswahl von Zielen ermöglicht wird, während die Anwendung genutzt wird. Gleichzeitig ist es eine sinnvolle Ergänzung, wenn eine eigene Präferenzauswahl vorgenommen werden kann, indem bestimmte POI gespeichert werden (bspw. wie dies auch bei der Auswahl von Ferienwohnungen der Fall ist; das gängige Icon zur Speicherung ist hier ein Herz).

12. Mängelmelder

Die als Bürgerservice angelegte Anwendung „CityKey App“ hat die Funktion eines Mängelmelders integriert, die auch für eine Pilotierung sinnvoll sein könnte. Wenn falsche Angaben zu Öffnungszeiten und anderen Angaben gemeldet und korrigiert werden können, so hilft dies der Datenqualität und somit der Verlässlichkeit in die Pilotanwendung.

13. Umkreissuche

Eine Standard-Funktion ist es auch, dass der aktuelle oder anvisierte Standort in den Anwendungen frei- oder angegeben werden kann und dann im Umkreis des spezifizierten Punktes Angebote und Alternativen angezeigt werden. Diese Funktion (Standortauswahl oder -identifizierung) sollte auch in einer Pilotanwendung zwingend integriert sein.

14. Komplementäre Empfehlungen

Nicht immer müssen die angezeigten Empfehlungen als Alternative zu bereits überlasteten POI angezeigt werden. Vielmehr ist auch die Anzeige von ergänzenden POI möglich. Wenn also eine Wanderung ausgewählt wird, dann können entlang der Route (wie dies auch bei ByTheWays erfolgt) Möglichkeiten zum Einkehren oder ergänzende Aktivitäten angezeigt werden. Eine solche Funktion ist in einen Empfehlungsdienst, der auf die Anzeige von Alternativen ausgerichtet ist, nicht unmittelbar integrierbar und nicht das Kernziel, sollte aber mitgedacht werden – auch, um Besuchsströme ggf. durch verlängerte Aufenthalte an einem solchen komplementären POI zu entzerren.

15. Ergänzende (lenkende) Informationen

Die Integration von ergänzenden Informationen, die unabhängig von der Auslastung sind, können ebenfalls eine Lenkungswirkung haben und sollten bei einer Pilotierung mitgedacht werden. So ist es möglich, dass die Luftqualität oder die Lärmbelastung an einem Ort gemessen wird und – sollten diese Informationen zuverlässig vorliegen – ergänzend mit ausgespielt werden. Eine solche Ausspielung ist auch selektiv denkbar, je nachdem, ob eine Erhöhung und Verringerung der Besuchsfrequenz das Ziel ist.

16. Private Alternativen

Wie die App Peuka zeigt ist es möglich auch im Privatbesitz befindliche Alternativen anzuzeigen. Bei Peuka bezieht sich dies explizit auf private Parkplätze, was zu einer Entlastung der Parkplatzsituation beitragen könnte. Je nachdem, wie gut das Angebot von Peuka-Parkplätzen in einem bestimmten Umkreis ist, könnte ein Link zu dieser Anwendung in einer Pilotierung ausgespielt werden.

17. Beispielanwendungen

Neben dem Funktionsumfang zeigen die Beispiele, dass es diverse digitale Anwendungen gibt, in die Auslastungsdaten mit integriert werden können, um den Nutzenden der jeweiligen Anwendungen direkt eine Live-Auslastung oder eine Prognose der zukünftigen Auslastung zu geben. Bei einigen An-

wendungen wie dem Lübecker Bucht Guide, dem Ausflugsticker Bayern, oder dem digitalen Reiseführer Wangerland sind Auslastungsdaten bereits integriert und werden mit ausgeliefert. An solchen Anwendungen kann sich ein Pilot direkt orientieren.

18. Ergänzung von bestehenden Anwendungen

Eine weitere Strategie könnte es sein, dass anstelle einer eigenen Pilotanwendung die Auslastungsdaten in etablierte Anwendungen integriert werden:

- **Globale Player:** Auch allgemeine Anwendungen wie Google Maps wären prädestiniert für eine Integration, wobei dies eher unrealistisch ist, dass ein globaler Player auf eine solche Integration ermöglicht.
- **Tourismusspezifische Anwendungen mit hoher Verbreitung:** Hier sind insbesondere jene Anwendungen von Interesse, die eine hohe Verbreitung haben. Dies sind im touristischen Kontext – wie im Rahmen von 6.1 gezeigt – insbesondere Komoot und Outdooractive.
- **Weitere Anwendungen:** Erfolgsversprechender sind daher zum einen tourismusspezifische Anwendungen mit hoher Reichweite wie die ADAC Trips App oder der DB Ausflug Ticker.
- **Startups:** Gleichzeitig könnten auch Startups wie urbnups, ByTheWays oder Naturtripp Interesse an solchen Informationen haben, wenn diese kostenfrei zur Verfügung stehen.
- **Special Interest Anwendungen:** Dies kann auch bedingt für spezielle Anwendungen wie die CityKey App, die Region App oder das UmweltNAVI gelten.

Zusammenfassende Beurteilung der Analyse der Informationskanäle

Google: Suchen zum Besuchsaufkommen starten, wie viele andere Suchen auch, bei Google. Für Destinationswebsites sowie alle anderen, die Informationen zum Besuchsaufkommen vorhalten ist es wichtig, eine entsprechende Sichtbarkeit bei Google zu haben. Dies kann durch entsprechende SEO-Maßnahmen gefördert werden. Eine weitere Möglichkeit besteht durch die Integration von POI, zu denen es Informationen zum Besuchsaufkommen gibt beim Dienst Google Things To Do.

Die Bahn: Die App der Bahn hat eine hohe Verbreitung bei Personen, die mit dem öffentlichen Nahverkehr verreisen. Zudem ist die DB Ausflug App relevant, da über diese regionale Verkehrsbetriebe auch POI einpflegen können, die dann auf aktuelle Auslastungsinformationen verweisen können.

Informationen vor Ort: Da Beschilderung wichtig ist, können Live-Auslastungsdaten auch über digitale Schilder ausgespielt werden. Zusätzliche Sichtbarkeit kann vor Ort über die wichtige App von Komoot in Form von Collections eingekauft werden, damit POI mit Auslastungsinformationen mehr Sichtbarkeit erhalten.

Destinationswebsites: Das größte Potenzial haben Destinationswebsites. Hier ist die Sichtbarkeitsanalyse entscheidend um die Wichtigkeit der jeweiligen Website ermitteln zu können. In Schleswig-Holstein sind dies insbesondere die 10 mit der höchsten Sichtbarkeit, darunter die TMOs Nordseetourismus.de und Ostsee-Schleswig-Holstein.de sowie St. Peter-Ording, Büsum und Grömitz, aber auch die Inseln Sylt, Föhr, Fehmarn und Helgoland sowie die Website von der Stadt Lübeck.

Wetter- und Windwebsites: Wetterwebsites haben eine mit Abstand höhere Sichtbarkeit als Windwebsites. Möglich wäre es, wenn die aktuellen Wetterdaten vorliegen, dass diese direkt in eine Pilotanwendung eingebunden werden, damit Nutzende gar keine andere Anwendung aufrufen müssen. Für die speziellen Use Cases können auch die Windangaben wichtig sein und direkt in der Pilotanwendung angezeigt werden.

Weitere Use Case spezifische Anwendungen: Bei Parkplatzanbietern hat Parkopedia mit Abstand die höchste Sichtbarkeit, aber auch Easypark, Parkster, Parkplatz-Finder oder HERE WeGo sollten für eine Ausspielung in Betracht gezogen werden, da diese vornehmlich app-basiert sind. Ähnliches gilt für Anwendungen, bei denen Camping- und Wohnmobilstellplätze recherchiert werden können.

6.4 Deployment des Recommenders in einem exemplarischen Frontend

Die beiden in Abschnitt 5 eingeführten Module der Auslastungsprognose und Alternativengenerierung werden im Recommender zusammengeführt, sodass für eine Ausspielung nur dieser Service verantwortlich ist. Über den Recommender werden Informationen so zusammengefasst, dass über eine einzige Anfrage des Recommenders für einen beliebigen POI Kontextinformationen zurückgespielt werden. Kontextinformationen sind beispielsweise Öffnungszeiten, Fließtextbeschreibungen und Live-Auslastung oder Prognosen über die zukünftige Auslastung des POI. Darüber hinaus besitzt der Recommender die Fähigkeit, Alternativen für einen bestimmten POI zu erzeugen. Diese können für unterschiedliche Zeiträume, beispielsweise für den aktuellen Tag, den folgenden Tag oder die beiden darauffolgenden Tage, generiert und bereitgestellt werden. Über die Schnittstellen-Dokumentation der FIWARE Plattform und der verwendeten Datenmodelle (siehe: <https://github.com/LAB-TOUR-SH/dvm>) kann ein Endgerät diese Informationen nach Belieben ausspielen. In LAB-TOUR SH wurde hierfür ein prototypisches Frontend entwickelt und an den Context Broker angeschlossen, damit die bereitgestellten Funktionalitäten des Backends (Context Broker, Prognose- und Recommendermodul) als auch die angeschlossenen Daten zu den Frequenzen getestet und das Zusammenspiel der einzelnen Komponenten analysiert werden kann. Zudem dient das Frontend als exemplarische Umsetzung (PoC), mit deren Hilfe die Eignung und Fehlerfreiheit der für den Endkunden bereitgestellten Funktionalitäten überprüft werden kann.

Die Benutzeroberfläche des Frontends konnte dabei entlang der Analyse der Informationskanäle (6.2) erstellt werden. So wurden wichtige Elemente wie die Darstellung als Karten- und auch als Listenansicht, Eine Adaption an Google Popular Times, Farbcodes für die Anzeige der Auslastung (grün = leer; rot = voll) usw. direkt integriert. Weitere Angaben wie Öffnungszeiten waren für das Recommendermodul relevant, aber in der Darstellung nicht vordergründig. Andere Anzeigen wie Wetterangaben, ein Mängelmelder, Angaben zu Luftqualität usw. sind zwar in Betracht gezogen worden, aber konnten aufgrund ihrer komplexen Form der Einbindung nicht in diesem Projekt umgesetzt werden. Sie sind jedoch möglicher Bestandteil einer Weiterentwicklung, die im Anschluss an das Projekt LAB-TOUR SH noch erfolgen kann.

Das Frontend ist die Benutzeroberfläche, über die Benutzer*innen mit der implementierten Anwendung interagieren können. Das Frontend basiert auf Webtechnologie und ist daher auf jedem beliebigen Endgerät mit einer Internetverbindung und einem aktuellen Browser erreichbar. Ziel ist es dem/der Benutzer*in die Möglichkeit zu bieten, Informationen zu POI in Schleswig-Holstein aufzuzeigen und den Standort der POI örtlich einzuordnen. Zusätzlich wird der/die Nutzer*in über die Live-Auslastung und zukünftige Auslastung von POI informiert. Somit können Nutzende des Frontends bewerten, welcher POI interessant ist und bei welchem die Auslastung nicht besonders hoch ist. Bei dem Prozess des Wählens einer Alternative von eine/n Benutzer*in unterstützt das Frontend intelligent, indem es sowohl sinnvolle Alternativen zu einem POI als auch geeignete und freie Parkplatzmöglichkeiten zu diesem aufzeigt. Die Startseite zeigt Schleswig-Holstein und verschiedene Cluster von POI, wie in Abbildung 121 zu sehen ist.

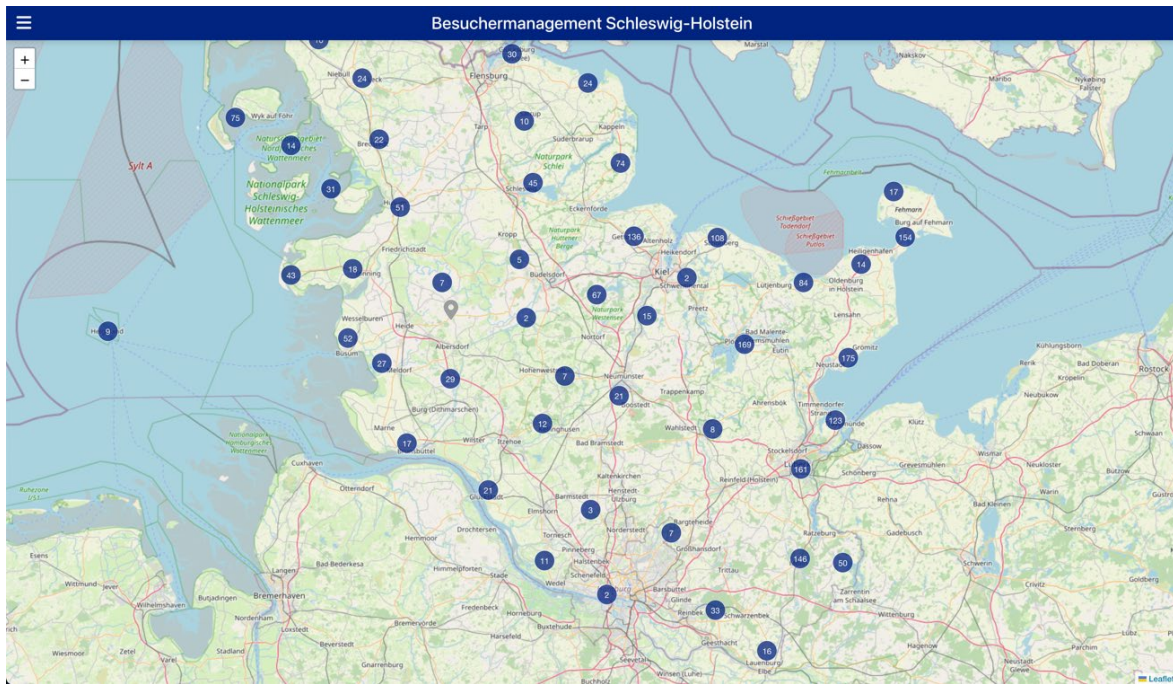


Abbildung 121: Frontend: Karte mit POI in Schleswig-Holstein

Abbildung 122 zeigt das Frontend, wenn Nutzer*innen die Landkarte heranzoomen. Am Beispiel von Fehmarn kann gesehen werden, wie das Frontend die Cluster Darstellung erweitert, sodass einzelne POI erkennbar sind.



Abbildung 122: Frontend: Detailansicht

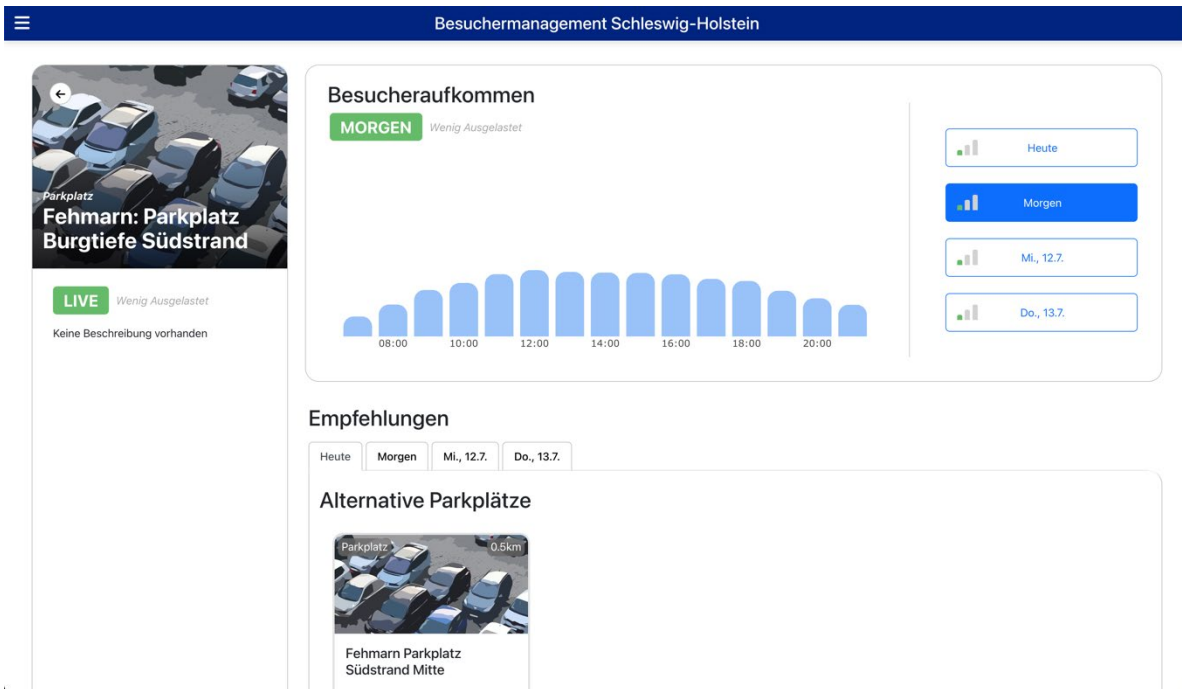


Abbildung 123: Frontend: Detailansicht zum Besucheraufkommen und generierten Alternativen am Beispiel Fehmarn Parkplatz Burgtiefe Südstrand

Klicken Nutzer*in beispielsweise auf einen Parkplatz, öffnet sich die Detailansicht für den Parkplatz (vgl. Abb. 123). Die äußere linke Spalte enthält in dieser Ansicht Informationen über den angeklickten POI. Diese enthalten den Namen, die Liveauslastung (farblich markiert) als auch eine Freitext Beschreibung. Das Besucheraufkommen wird stündlich in einem Diagramm angezeigt. Sind die Balken hoch, ist die Auslastung zur entsprechenden Zeit hoch. Erreicht der Spitzenwert eine Auslastung von 80% an einem Tag, wird der POI als ausgelastet rot markiert. Analog können die Auslastungen für heute, morgen und die darauffolgenden Tage eingesehen werden. Weiterhin werden dem/der Nutzer*in mögliche Alternativen angezeigt. Wenn Nutzende auf der Landkarte auf einen Parkplatz klicken, wird angenommen, dass ein Parkplatz gesucht wird und es werden Alternativen für diesen Parkplatz angezeigt und anschließend mögliche Erlebnis-POI in der Nähe. Klicken Nutzende hingegen auf einen Erlebnis-POI, beispielsweise auf einen Strand, werden die Informationen andersherum angezeigt. Basierend auf den Auslastungsprognosen der POI in der Nähe ändern sich die Alternativvorschläge für jedes Zeitintervall.

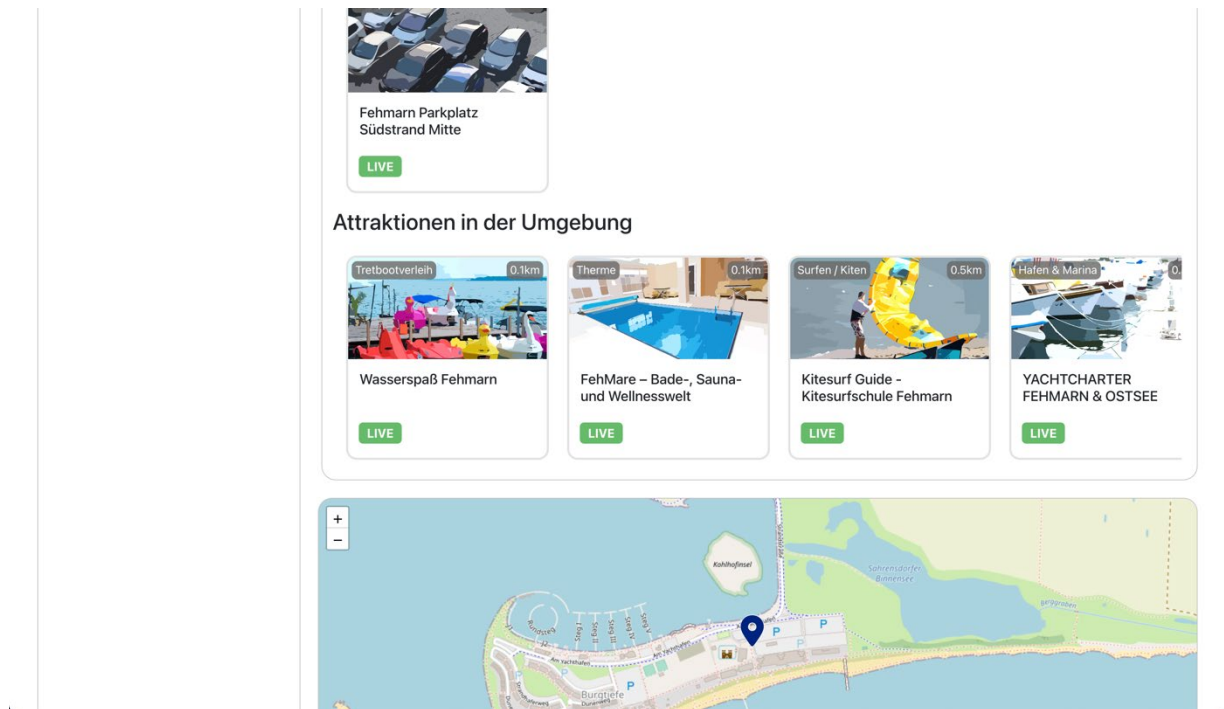


Abbildung 124: Frontend: Detailansicht zu Attraktionen in der Umgebung

Die Parkplatzinformationen im Frontend zeigen die umliegenden Attraktionen an (vgl. Abb. 124). Jede Attraktion hat eine Spitzenauslastung, die innerhalb eines bestimmten Zeitintervalls als farblich markierte Badges angezeigt wird. Dadurch erhalten Nutzer*innen schnell die benötigten Informationen, um fundierte Entscheidungen zu treffen. In der Detailansicht erhält eine Nutzerin oder ein Nutzer den Auslastungsverlauf des aktuellen Tages für eine ausgewählte Attraktion. Dieser Verlauf setzt sich aus vergangenen, aktuellen und prognostizierten Daten zusammen. Ähnlich wie bei einer Wettervorhersage kann der Verlauf auch für die nächsten drei Tage abgerufen werden, indem die entsprechenden Schaltflächen angeklickt werden. Für jeden dieser Tage werden drei bis acht alternative Attraktionen vorgeschlagen. Es gibt zwei Arten von Detailansichten im Frontend: Beim Klicken auf eine Attraktion öffnet sich die Detailansicht und es werden alternative Erlebnis-POI für die ausgewählte Attraktion angezeigt. Zudem werden, falls vorhanden, Parkmöglichkeiten für aufgelistet. Wenn auf einen Parkplatz geklickt wird, werden alternative Parkplätze angezeigt und es werden auch Attraktionsempfehlungen aufgelistet, die von diesem Parkplatz aus erreichbar sind.

Das Frontend wurde unter Verwendung der Programmiersprache React implementiert und bezieht seine Daten ausschließlich vom Context Broker. Beim Öffnen der Karte lädt der jeweilige Browser oder Client die benötigten Attraktionsdaten herunter. Ebenso lädt das Frontend beim Öffnen einer Detailansicht die benötigten Attraktionsdaten sowie alternative Vorschläge für diese Attraktion vom Context Broker herunter. Die Bilder im Frontend stammen aus der Datenbank der TASH. Falls kein Foto vorhanden war, wurde für jede Kategorie ein Beispielfoto im Comic-Stil verwendet. Die farbliche Markierung der Auslastungs-Badges basiert auf den Auslastungswerten der Attraktionen, die auf einer Skala von 0 bis 100 Prozent angegeben sind. Dabei stellt jedes Drittel des Farbverlaufs einen Schritt von Grün über Gelb bis Rot dar.

6.5 Pilotierung des Recommenders

Neben den in Abschnitt 6.2 gezeigten Möglichkeiten der Verbreitung und einem Live-Test des Recommenders (vorheriger Abschnitt 6.4), gibt es weitere Optionen, die im Rahmen dieses Projektes realisiert werden konnten. Eine naheliegende war die Integration des Prototypen in das #sh_wlan, das vom Verbundpartner ADDIX betrieben wird. Das Ziel im Rahmen des Projektes besteht somit darin, innerhalb des Captive Portals des #sh_wlan Empfehlungen des Recommender-Systems bereitzustellen. Ein Captive Portal ist üblicherweise eine Website, die vor der Nutzung eines öffentlichen WLANs wie des #sh_wlan aufgerufen werden muss. Dieser Vorgang wird normalerweise durch einen Captive Network Assistant (CNA) automatisiert, ein Programm, das auf den Clients vorhanden ist. Bei der Verbindung mit einem WLAN erkennt der CNA, dass ein Captive Portal vorhanden ist, und zeigt dies automatisch an. Im Fall des #sh_wlan wird bei der ersten Verbindung mit dem WLAN eine Anmeldeseite angezeigt, auf der die Nutzer*innen die Allgemeinen Geschäftsbedingungen (AGB) und Nutzungsbedingungen akzeptieren müssen, um das WLAN kostenfrei nutzen zu können. Zukünftige Anmeldungen erfordern kein erneutes Captive Portal. An diesem Punkt kann das Ziel erreicht werden, indem gelegentlich ein Captive Portal angezeigt wird, um den Nutzenden mit hilfreichen Informationen weiterzuhelfen

6.5.1 Das Kampagnensystem

ADDIX hat ein flexibles System entwickelt das es ermöglicht, Kampagnen in einem bestimmten Bereich und innerhalb eines festgelegten Zeitfensters gezielt einzusetzen. Eine Kampagne stellt dabei eine kleine Website dar, die einem Nutzer oder einer Nutzerin beim Login in das #sh_wlan angezeigt wird, vorausgesetzt, dass die betreffende Person sich zuvor über die Anmeldeseite "registriert" hat.

Die Entscheidung, ob eine Kampagne einem Nutzenden angezeigt wird, wird wesentlich von drei Faktoren beeinflusst: dem Ort, dem Zeitpunkt und der Frage, ob dem Nutzenden bereits dieselbe oder eine andere Kampagne zuvor angezeigt wurde.

- **Ort:** Eine Kampagne wird ausgespielt, wenn sich ein Nutzer mit einem Access Point verbindet. Bei der Erstellung einer Kampagne wird ein bestimmter Bereich ausgewählt, in dem die Kampagne angezeigt werden soll. Basierend auf dem Standort des Access Points kann zum Zeitpunkt der Verbindung eines Nutzers mit dem WLAN entschieden werden, ob diese Kampagne ausgespielt werden soll. Die genaue Position des Nutzers ist dabei nicht bekannt, außer dass er sich in Reichweite des betreffenden Access Points befindet.
- **Zeitpunkt:** Eine Kampagne ist nur innerhalb eines bestimmten Zeitraums aktiv. Es wird ein Start- und Enddatum festgelegt.
- **Kampagnenbesuche:** Bei der Erstellung einer Kampagne kann festgelegt werden, wie oft ein Nutzer diese angezeigt bekommen soll. Dies kann einmalig oder in regelmäßigen Abständen erfolgen, beispielsweise alle drei Stunden. Dadurch können Nutzern beim Verbinden mit dem #sh_wlan Informationen über den aktuellen Standort oder ähnliches einmalig oder regelmäßig zur Verfügung gestellt werden.

Neben diesen drei Faktoren spielt natürlich auch der Inhalt der Kampagne eine Rolle, also das, was der Nutzer letztendlich auf seinem Gerät sieht. Hierbei kommen vorgefertigte Vorlagen mit Platzhaltern zum Einsatz, die bei der Ausspielung durch relevante Informationen ersetzt werden. Die Werte für die Platzhalter werden entweder beim Erstellen einer Kampagne fest definiert oder zur Laufzeit durch ein Skript dynamisch generiert. Letzteres ist wichtig, um Inhalte von externen APIs wie dem Recommender einzubinden. Am Ende entsteht eine herkömmliche Website, die dem Nutzer angezeigt werden kann (vgl. Abb. 125).

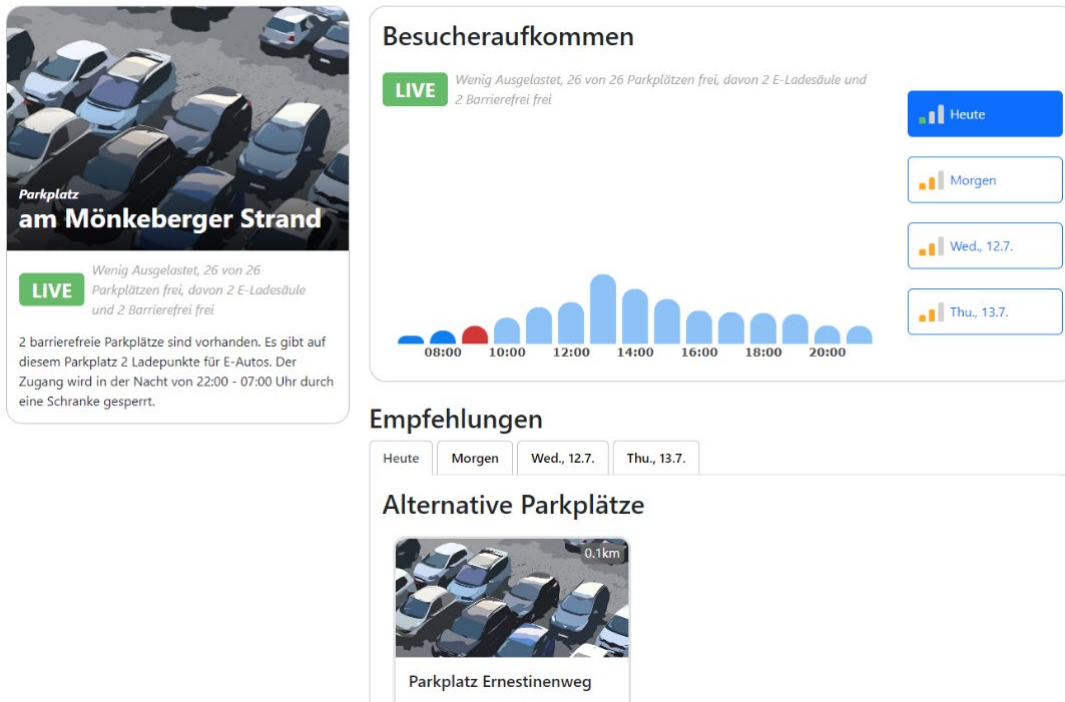


Abbildung 125: Website (Kampagne #sh_wlan) mit den Platzhaltern und per Skript generierten dynamischen Inhalten der aktuellen Parkplatzauslastung

6.5.2 Funktionsweise #sh_wlan und Kampagnen

Die grundlegende Funktionsweise des #sh_wlan lässt sich auf folgende Weise vereinfacht darstellen: Ein Client, beispielsweise ein Smartphone, verbindet sich mit dem WLAN, wodurch eine Verbindung zu einem Access Point hergestellt wird. Der Access Point sendet eine Anfrage an einen Authentifikationsserver, wie beispielsweise einen Radius-Server, um zu überprüfen, ob der Client Zugang zum WLAN haben darf. Der Radius-Server prüft in seiner Datenbank, ob der Client bereits zuvor eine "Registrierung" durchgeführt hat, und sendet eine entsprechende Antwort zurück an den Access Point. Wenn die Antwort positiv ist, wird der Client autorisiert und die Anmeldung im WLAN ist abgeschlossen. Falls dem Client jedoch der Zugang verwehrt wird, wird ein separater Prozess in Gang gesetzt (siehe auch den opt-Block im Diagramm unten). Der Access Point versucht, den Client auf ein Captive Portal umzuleiten. Der Client ruft das Captive Portal auf und erhält eine Anmeldeseite. Nachdem der Client die Allgemeinen Geschäftsbedingungen, Datenschutzbestimmungen und Nutzungsbedingungen akzeptiert hat, wird der Client registriert und das Captive Portal bewirkt, dass der Access Point erneut den Authentifikationsserver abfragt, ob der Client das WLAN nutzen darf. Nun erhält der Access Point die Bestätigung und der Client wird autorisiert.

Um eine Kampagne auszuspielen, mussten zwei Anpassungen vorgenommen werden: die Autorisierungsprüfung durch den Radius-Server und das Captive Portal. Bei der Autorisierung wird nun nicht nur entschieden, ob der Client registriert ist, sondern auch, ob zum aktuellen Zeitpunkt am Access Point, mit dem sich der Client verbunden hat, eine Kampagne aktiv ist. Zudem wird überprüft, ob der Client die Kampagne bereits gesehen hat und wie oft die Kampagne Nutzern angezeigt werden soll. Wenn dem Nutzer eine Kampagne angezeigt werden soll, sendet der Access Point eine Zugriffsablehnung (Access Reject) zurück, als ob der Client nicht registriert wäre, was dazu führt, dass der Client auf das Captive Portal umgeleitet wird. Das Captive Portal überprüft beim Aufruf, ob es sich um einen

neuen oder wiederkehrenden Nutzer handelt. Diese Prüfung ist einfach: Wenn der Client in unserer Datenbank registriert ist, handelt es sich um einen wiederkehrenden Nutzer. Wenn es sich um einen neuen Nutzer handelt, wird wie bisher die Anmeldeseite angezeigt. Wenn es sich nicht um einen neuen Nutzer handelt, wird die Kampagne angezeigt. Abbildung 126 zeigt ein Diagramm, welches den kompletten Anmeldeprozess an das #sh_wlan inklusive der eventuellen Ausspielung der Anmelde- oder Kampagnenseite veranschaulicht:

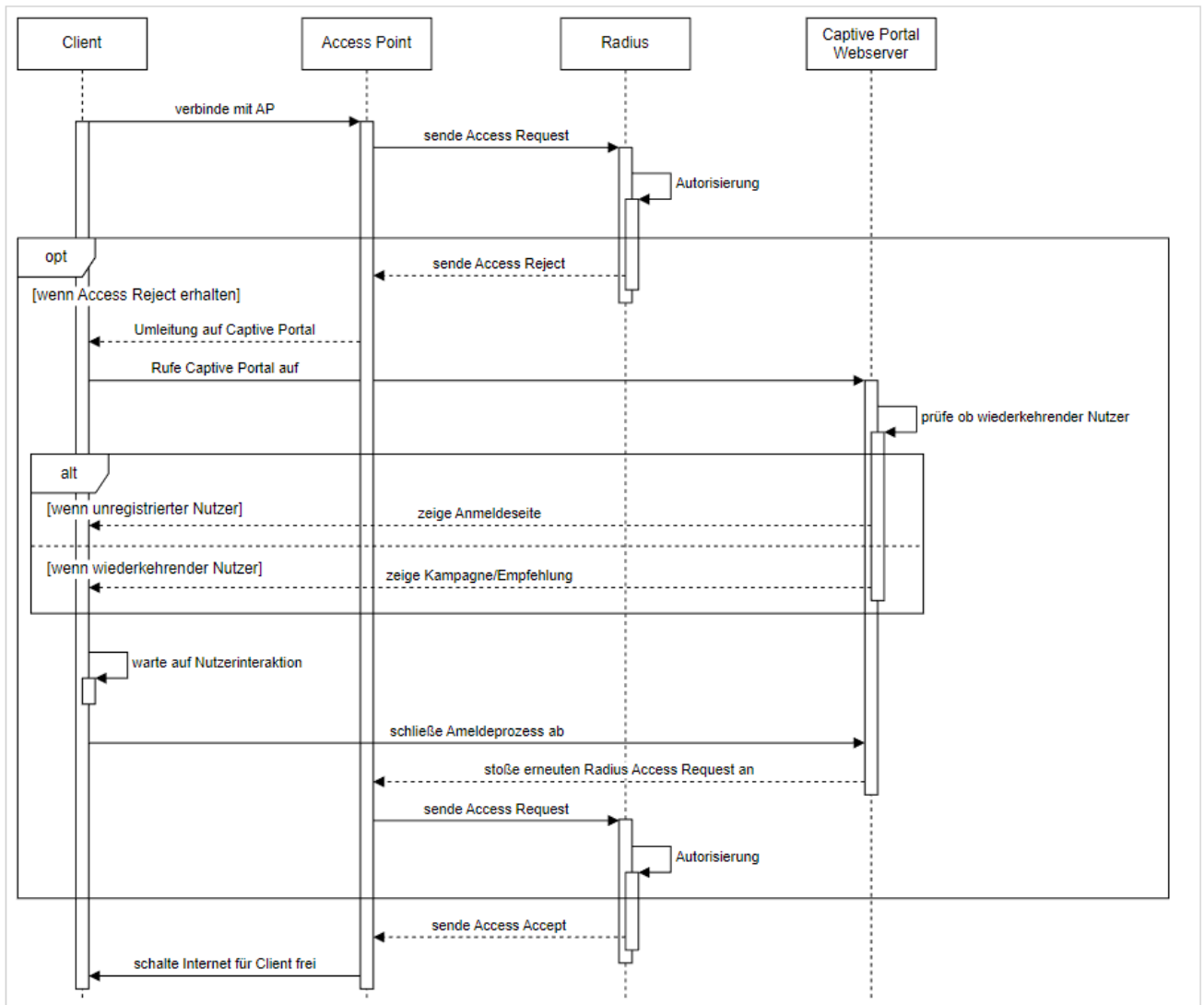


Abbildung 126: Anmeldeprozess an das #sh_wlan inklusive der eventuellen Ausspielung der Anmelde- oder Kampagnenseite

6.5.3 Dynamischer Inhalt und das Ausspielen von Empfehlungen

Das System zur Ausspielung von Kampagnen war bislang statisch. Es basierte auf Vorlagen mit Platzhaltern, die bei der Erstellung einer Kampagne durch festgelegte Werte ersetzt wurden. Es gab nur begrenzt Raum für dynamischen Inhalt. Um personalisierte Empfehlungen anhand von Ort und Zeitpunkt ausspielen zu können, musste eine dynamische Komponente hinzugefügt werden. Diese Weiterentwicklung fand im Rahmen des Projektes LAB-TOUR SH statt. Die Empfehlungen werden über eine gesonderte API vom Recommender bereitgestellt. Daher musste das Captive Portal in der Lage sein,

die erforderlichen Informationen von der gesonderten API des Recommenders abzurufen und anzuzeigen. Hierfür wurde das Vorlagensystem erweitert, um Platzhalter zu ermöglichen, die mit dynamischen Daten gefüllt werden können. Das Captive Portal führt auf der Serverseite ein Skript aus, das die benötigten Daten von der API des Recommenders abrufen, sie aufbereitet und damit die Platzhalter für die dynamischen Daten füllt.

Der Recommender benötigt für seine Funktion einen POI. Da der Nutzer im Captive Portal keine Möglichkeit hat, einen POI auszuwählen, im Gegensatz zur Benutzeroberfläche im Browser, muss der POI anhand der aktuellen Position des Nutzers bestimmt werden. Das Captive Portal verfügt über Informationen darüber, mit welchem Access Point sich der Client verbunden hat, und auch die Position der Access Points ist bekannt. Die Koordinaten werden an das Skript übergeben, das anschließend die API nach einem POI anhand der Koordinaten abfragt. Mit dem erhaltenen POI wird dann der Recommender aufgerufen, der Empfehlungen und Alternativen generiert. Diese Daten werden an das Captive Portal zurückgegeben, das daraufhin die Seite generiert und an den Client weiterleitet.

6.5.4 Pilotierung des Recommenders Strandweg Mönkeberg

Um eine gezielte Lenkung der Besucherströme im Tourismus zu erreichen, ist es erforderlich, die entsprechenden Empfehlungen und Informationen zielgerichtet bereitzustellen. Im Rahmen des Projektes LAB-TOUR SH führt ADDIX eine Pilotierung durch, bei der ein konkretes Beispiel zur Ausspielung solcher Informationen zum Einsatz kommt. Der Parkplatz am Strandweg in Mönkeberg wurde im Rahmen des Projektes mit Parkplatzsensorik ausgestattet. In unmittelbarer Nähe des Parkplatzes befindet sich ein #sh_wlan Access Point am Strandkiosk. Zusätzlich ist in Mönkeberg bereits der DorfFunk im Einsatz, wobei die Anbindung an den DorfFunk in Zusammenarbeit mit dem Verein Mönkeberg-ist.net realisiert wurde. Die relevanten Informationen werden sowohl über den #sh_wlan Access Point als auch im DorfFunk für die Zielgruppe bereitgestellt. Die Parkplatzsensorik besteht aus bodengebundenen Parkplatzsensoren des Herstellers Libelium (Typ "Smart Parking Node"; Version 2 Smart Parking EU). Diese Sensoren verwenden hauptsächlich Radartechnologie und verfügen als Backup über einen Magnetsensor.



Abbildung 127: Sensorik am Parkplatz Mönkeberg

Die im Rahmen der Pilotierung verwendeten Informationen beziehen sich auf den Belegungsgrad des Parkplatzes am Strandweg in Mönkeberg, der durch die Parkplatzsensorik erfasst wird. Die Sensoren übermitteln den Belegungszustand per LoRaWAN über ein Gateway an die Datenplattform, die im Projekt verwendet wird. Auf der Datenplattform werden spezifische Datenmodelle für die Belegungsinformationen verwendet. Jeder Parkplatzsensor wird auf der Datenplattform als Datenmodell "Device" betrachtet. Aus diesem Device wird ein Datenmodell "ParkingSpot" erstellt. Ein ParkingSpot kann entweder einer ParkingGroup, einem ParkingArea oder direkt einem Datenmodell "OffStreetParking" zugeordnet werden. Das Datenmodell "OffStreetParking" liefert kontinuierlich die Belegungsinformationen für die Parkflächen am Strandweg. Diese Informationen werden zur Berechnung von Auslastungsprognosen verwendet und von der Plattform verarbeitet (vgl. Abb. 127).

Als Ergebnis werden jede Stunde vier Prognosen und Empfehlungen generiert. Diese werden als Datenmodell "TouristRecommendation" bereitgestellt, jeweils eine Empfehlung für den aktuellen Tag bis Mitternacht und drei weitere Empfehlungen für die folgenden Tage. Der Belegungsverlauf und die Prognose werden in einem Datenmodell "OccupancySeries" für einen Zeitraum von 96 Stunden gespeichert. Dieses Datenmodell enthält alle relevanten Informationen und ermöglicht die Auslieferung von Empfehlungen an den Kunden.

Für die Pilotierung des Recommenders wird der #sh_wlan Access Point am Kiosk "Zantopps Fährhaus" genutzt, der sich in unmittelbarer Nähe des mit Parkplatzsensorik ausgestatteten Parkplatzes am Strandweg befindet. Auf dem #sh_wlan Access Point ist eine dynamische Kampagne konfiguriert, die systemseitig aktiv ist. Wenn sich ein #sh_wlan Benutzer über das Captive Portal anmeldet, wird der Recommender basierend auf dem POI des Parkplatzes am Strandweg eingesetzt und berücksichtigt die erfasste Belegung durch die Sensoren. Das Captive Portal erkennt den mit dem Client verbundenen

Access Point sowie dessen Position. Die Koordinaten des entsprechenden Access Points werden an das Skript übergeben, das dann die API anhand der Koordinaten nach einem POI abfragt. In diesem Fall ist der Access Point am Kiosk "Zantopps Fährhaus" der relevante Punkt in unmittelbarer Nähe des POI Parkplatz Strandweg.

Mit Hilfe des POI wird der Recommender angesteuert, der daraufhin Empfehlungen und Alternativen generiert. Diese Daten werden dann an das Captive Portal zurückgegeben, das daraufhin die entsprechende Seite generiert und an den Client sendet. Abbildung 128 zeigt, wie den #sh_wlan Benutzer*innen auf den Endgeräten die relevanten Informationen des Recommenders über die Ausspielung einer dynamischen Kampagne angezeigt wird.



Abbildung 128: Ausspielung einer Kampagne am #sh_wlan Access Point Mönkeberg auf mobilem Endgerät

Pilotierung Strandweg Mönkeberg via DorfFunk

Die dynamische Kampagne, die im #sh_wlan ausgespielt wird, kann technisch nur auf dem jeweiligen Access Point gesteuert werden. Um die entsprechende Kampagne anzuzeigen, muss das Gerät mit dem Access Point verbunden sein und eine funktionierende Übertragungsverbindung für die Inhalte bestehen. Um die Empfehlungen des Recommenders auch auf anderen Plattformen beispielhaft zu steuern, wurde im Rahmen des Projektes eine Integration mit dem DorfFunk vorgenommen.

Der DorfFunk ist eine von Fraunhofer IESE entwickelte Kommunikationsplattform für regionale Gemeinschaften und wurde im Jahr 2015 im Rahmen des Projektes "Digitale Dörfer" entwickelt. Mit der DorfFunk-App können Bürger*innen ihre Hilfe anbieten, Gesuche veröffentlichen und miteinander kommunizieren. Auch Neuigkeiten aus den DorfNews werden angezeigt, wodurch bereits in der Grundkonfiguration des DorfFunks die Möglichkeit zur Anzeige von Informationen besteht. Um die Informationen, ähnlich wie im #sh_wlan, auch im DorfFunk anzeigen zu können, werden die Inhalte der dynamischen Kampagne dupliziert und auf einem öffentlich zugänglichen Webserver gespeichert. Von diesem Webserver werden die Inhalte abgerufen und in Zusammenarbeit mit dem Account von "schrevenborn-aktuell.de" (ehemals "moenkeberg-ist.net") in den DorfFunk von Mönkeberg integriert. Dadurch wird ermöglicht, dass im DorfFunk dieselben Informationen abrufbar sind wie direkt vor Ort über den #sh_wlan Access Point am Parkplatz Strandweg. Es ist somit nicht erforderlich, sich in unmittelbarer Nähe des relevanten Access Points aufzuhalten, sondern die entsprechenden Informationen des Recommenders für den Parkplatz Strandweg werden über den DorfFunk angezeigt (vgl. Abb. 129).

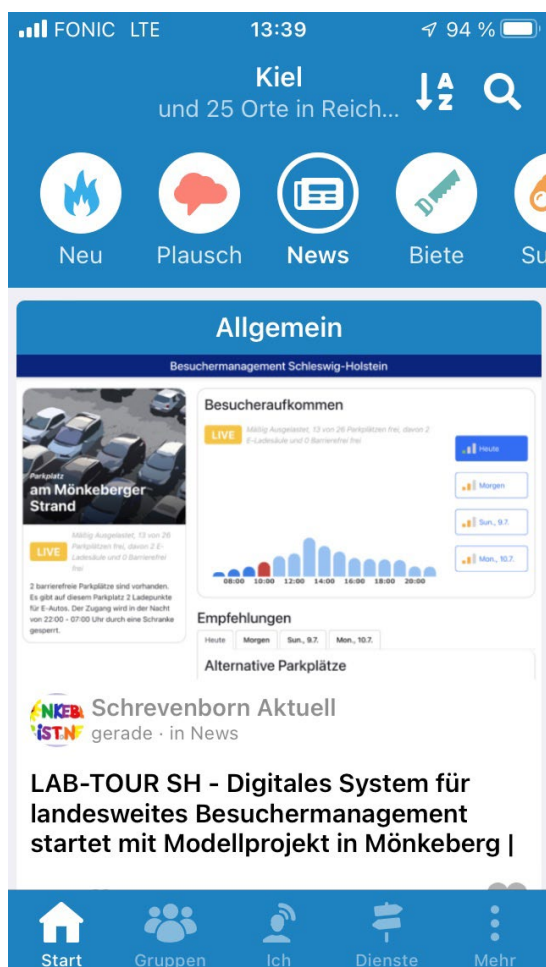


Abbildung 129: Ausspielung einer Kampagne im Rahmen von DorfFunk

6.6 Testszenarios

Neben der realen Ausspielung, die in den vorherigen Abschnitten erläutert wurde, können auch unterschiedliche Testszenarios umgesetzt werden. Diese sind nachfolgend dokumentiert und zeigen damit auf, wie die Ausspielung des Recommenders mit Hilfe der vier Use Cases getestet werden könnte. Da-

für werden unterschiedliche Arten von Ausflügen und Aktivitäten gezeigt, die Besucher*innen während ihres Aufenthaltes erleben können. Anhand der Grafiken werden mögliche Verläufe einer Customer Journey aufgezeigt, sofern der Recommender genutzt oder nicht genutzt wird.

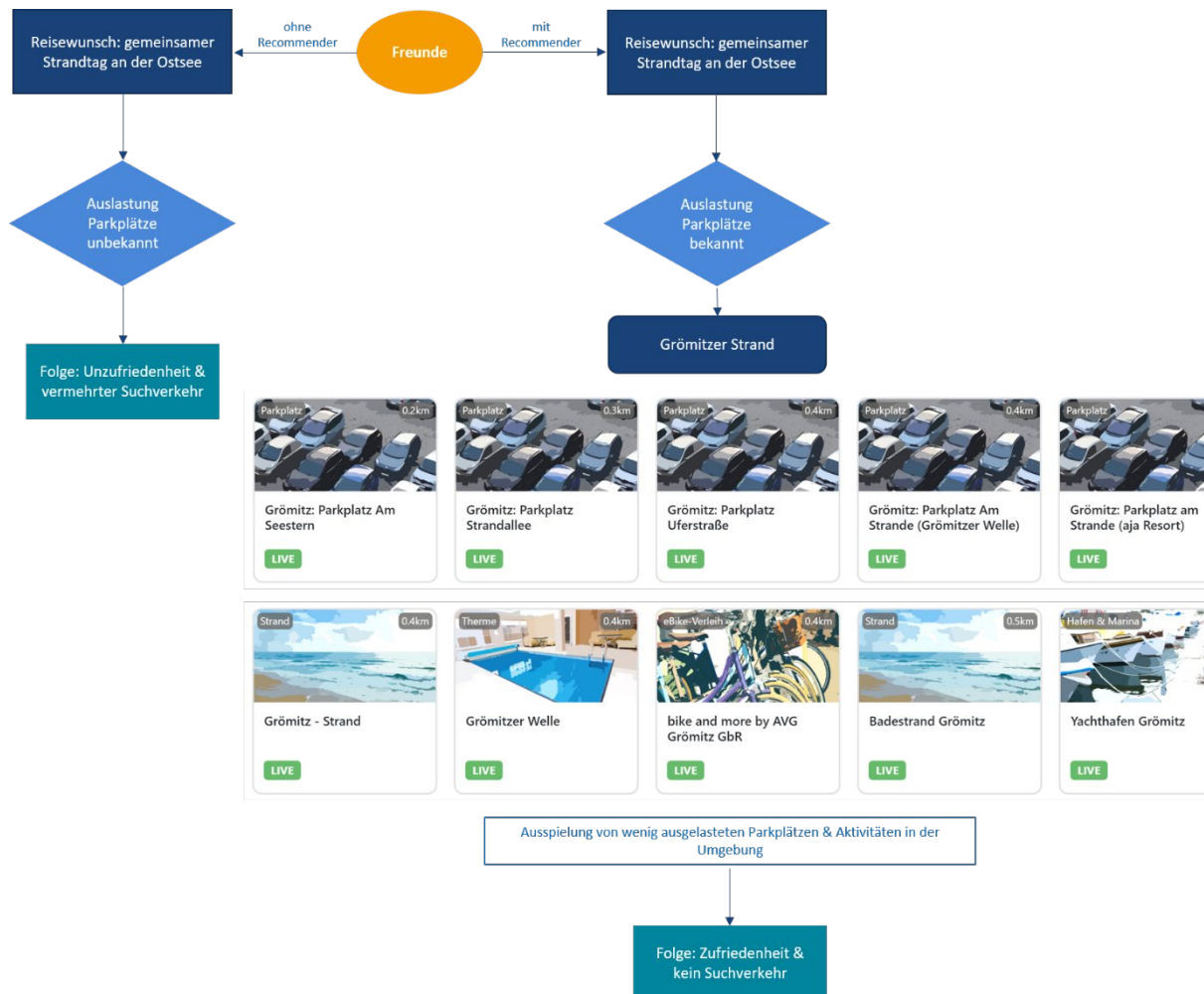


Abbildung 130: Testszenario bei Nutzung bzw. Nicht-Nutzung des Recommenders am Beispiel des Use Cases "Parkplätze Ostsee"

Bei einem gemeinsamen Strandtag an der Ostsee wird oft ein Parkplatz in Strandnähe benötigt, da viele Besucher*innen mit dem Auto anfahren. Sollte im Vorfeld kein Recommender genutzt werden, ist die Auslastung der Parkplätze in der Umgebung des Strandes unbekannt. In vielen Fällen müssen zunächst mehrere Parkplätze angefahren werden, um eine geeignete Parklücke zu finden. Dadurch sind die Besuchenden nicht nur aufgrund der langwierigen Suche häufig bei der Anreise bereits unzufrieden, sondern tragen zusätzlich auch zu einem erhöhten Verkehrsaufkommen bei. Um schon vor oder während der Anreise zum Strand nach Grömitz Informationen zum aktuellen Belegungsstand verschiedener geeigneter Parkplätze zu erhalten, werden den Nutzenden des Recommenders Parkplätze ausgespielt, die nach geringer Auslastung sowie Entfernung sortiert sind. Hierzu muss lediglich auf den POI Grömitzer Strand geklickt werden, damit entsprechende Parkplätze und alternative Aktivitäten in der Umgebung angezeigt werden. Der Suchverkehr nach freien Parklücken wird dadurch deutlich reduziert und Gäste sind durch die schnelle Ankunft am Ziel zufriedener (vgl. Abb. 130).

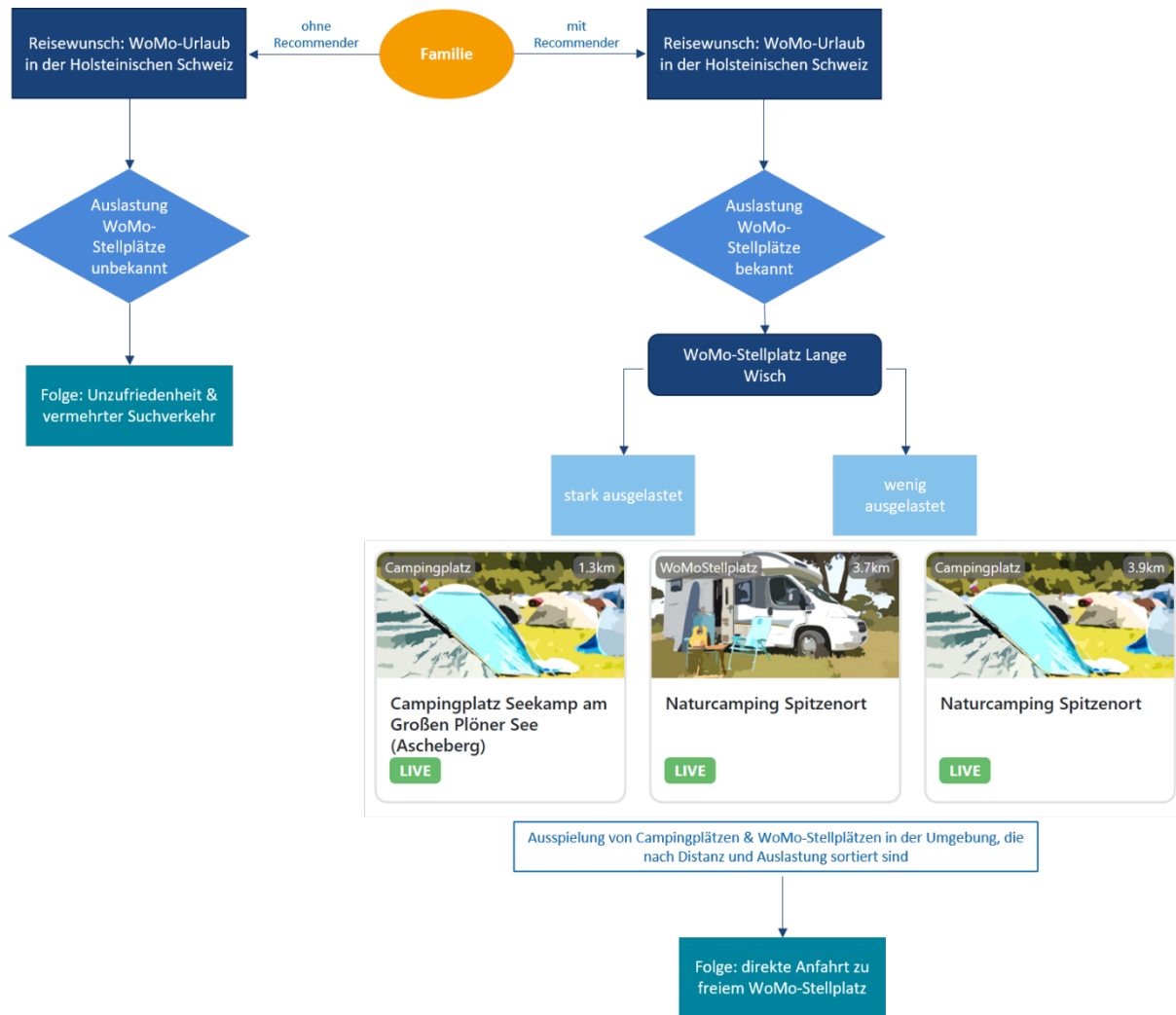


Abbildung 131: Testszenario bei Nutzung bzw. Nicht-Nutzung des Recommenders am Beispiel des Use Cases "Wohnmobilstellplätze Holsteinische Schweiz"

Einer ähnlichen Logik folgt der Recommender bei einer Reise mit einem Wohnmobil in der Holsteinischen Schweiz. Auch hier werden - ohne Nutzung des Recommenders - vermehrt unterschiedliche WoMo-Stellplätze aufgesucht, da die Auslastung und Belegung der einzelnen Stellplätze unbekannt ist. Die Suche nach einem geeigneten freien Stellplatz verursacht erneut ein erhöhtes Verkehrsaufkommen in der Umgebung sowie Unzufriedenheit bei den Besuchenden. Der Recommender hingegen zeigt eingangs schon mit jedem Klick auf einen Wohnmobilstellplatz einen Überblick, welche Wohnmobilstellplätze und Campingplätze in der Nähe vorhanden sind. Diese wurden bereits so sortiert, dass die näher liegenden und weniger ausgelasteten Stellplätze in der Rangfolge weiter vorne angezeigt werden. Die Auslastung der jeweiligen Orte ist dabei deutlich ersichtlich. Sollte ein Stellplatz ausgelastet sein, wird dieser dennoch angezeigt – allerdings eindeutig als voll markiert. So kann verhindert werden, dass dieser Ort angefahren und nicht notwendiger Suchverkehr verursacht wird. Zudem können sich Nutzer*innen des Recommenders ebenfalls über die Prognose der Auslastung in den nächsten Tagen informieren und dadurch bereits im Vorfeld der Reise einen Stellplatz auswählen, der zum Zeitpunkt der Ankunft voraussichtlich wenig ausgelastet sein wird. Eine direkte Anfahrt zu einem freien Wohnmobilstellplatz wird dadurch besser ermöglicht (vgl. Abb. 131).

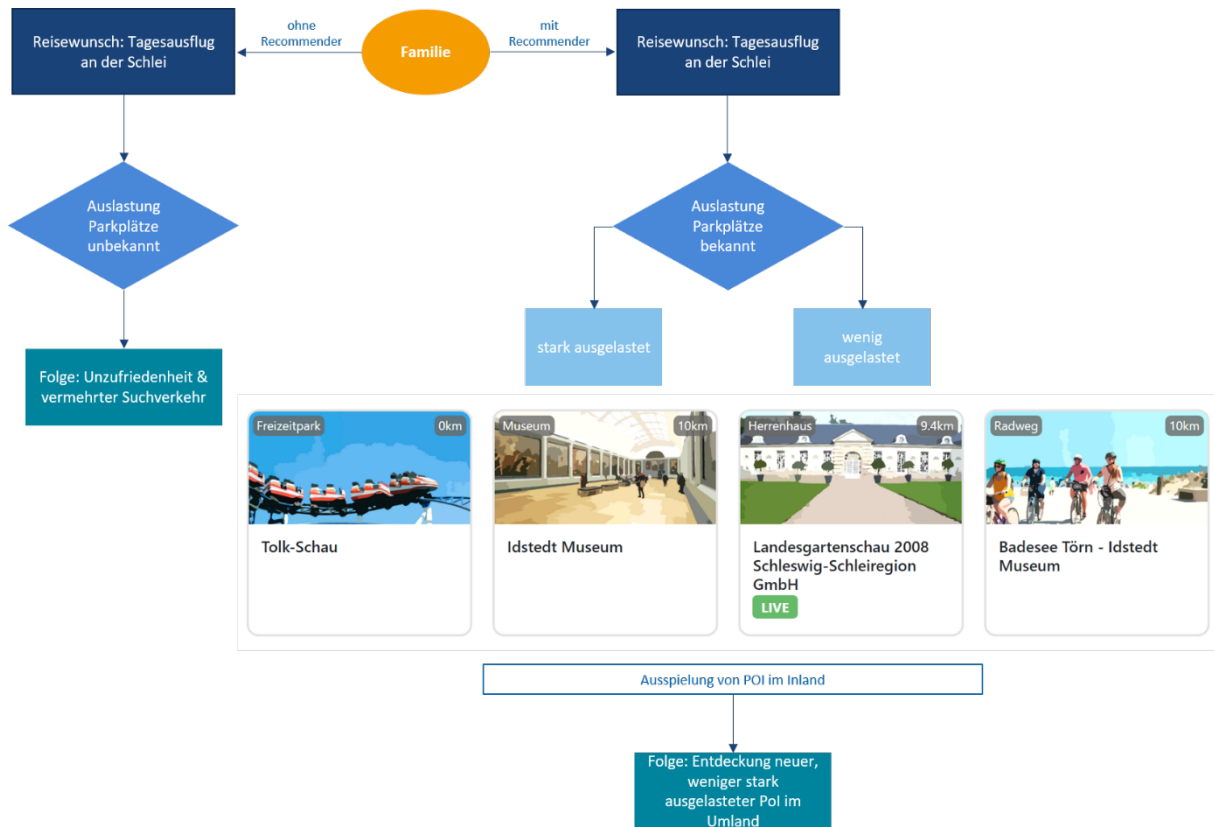


Abbildung 132: Testszenario bei Nutzung bzw. Nicht-Nutzung des Recommenders am Beispiel des Use Cases "Hotspots Schlei"

Eine Familie möchte einen Tagesausflug an der Schlei unternehmen. Ohne die Nutzung des Recommenders sind sowohl die Auslastungen der örtlichen Parkplätze als auch die, der damit verbundenen POI, unbekannt. Das führt zeitweilig dazu, dass wiederum verschiedene Parkplätze angefahren werden müssen und der Parkplatzsuchverkehr erneut eine große Belastung auf den Straßen darstellt. Weiterhin kann auch davon ausgegangen werden, dass das geplante Ausflugsziel bereits unter starkem Besucherdruck steht. Um solche negativen Begleiterscheinungen eines Ausfluges zu umgehen, ist der Recommender eine geeignete Lösung, da dieser nicht nur die Auslastung der Parkplätze ausspielt, sondern gleichermaßen auch alternative POI in weniger stark frequentierten Regionen im Inland ausgibt. Dadurch ergibt sich für Nutzer*innen des Recommenders die Möglichkeit, sowohl neue Orte kennenzulernen als auch stark ausgelastete Orte mit den damit einhergehenden Problematiken zu vermeiden (vgl. Abb. 132).

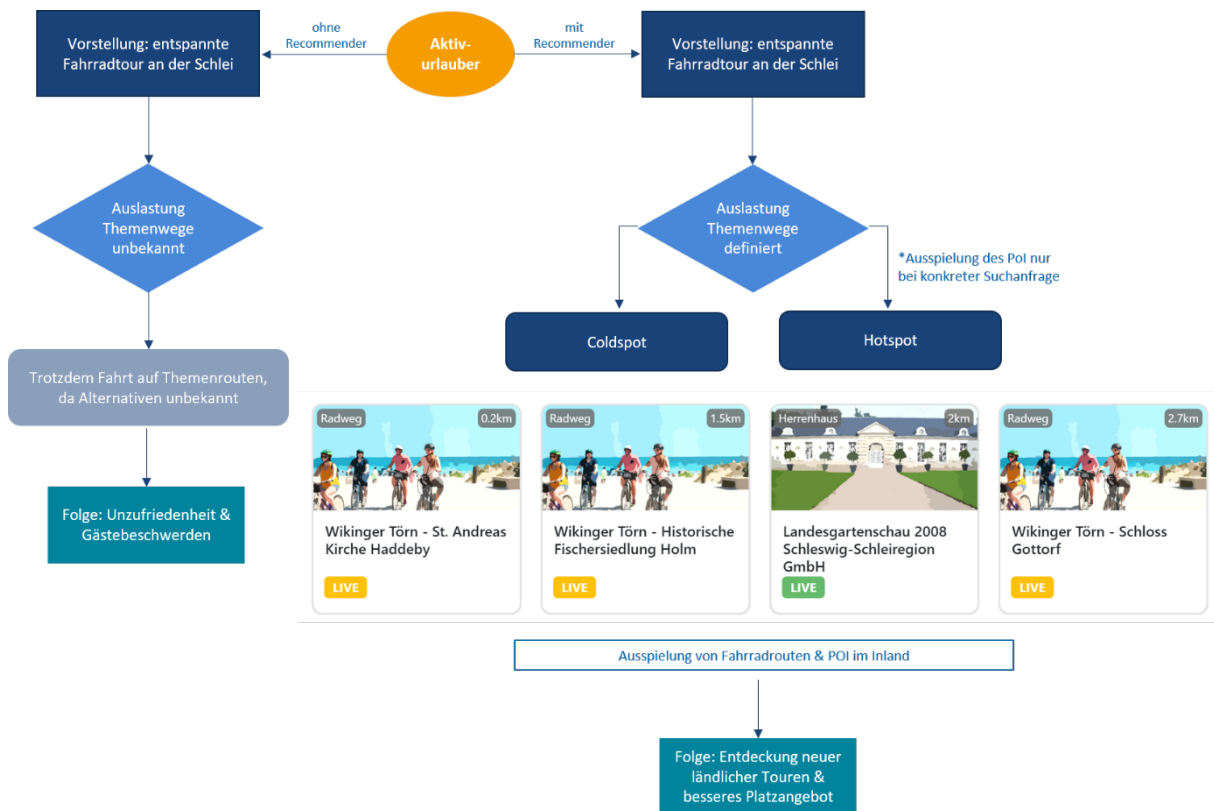


Abbildung 133: TestszENARIO bei Nutzung bzw. Nicht-Nutzung des Recommenders am Beispiel des Use Cases "Radrouen Schlei"

Eine entspannte Fahrradtour an der Schlei kann sich ohne den Recommender schnell ins Gegenteil wenden, da die Themenwege an der Schlei oft zu stark ausgelastet sind. Durch die schmalen Fahrradwege ist es mitunter nicht möglich, mit den Fahrrädern aneinander vorbei zu fahren, sodass sich Gäste bereits beschwerten. Da Fahrradfahrende ohne den Recommender keine Informationen zu der Auslastung oder Alternativen haben, bleibt dieses Problem bestehen. Um die im Inland gelegenen, i. d. R. nicht stark ausgelasteten Themenwege, in den Fokus der Radfahrenden zu stellen, wurden diese als Coldspots im System hinterlegt. Auf der anderen Seite wurden regelmäßig ausgelastete Fahrradwege als Hotspots markiert. Dadurch werden, unabhängig von der momentanen Auslastung, die stark besuchten Radwege nur bei der aktiven Suche als Alternativen ausgespielt. Wird innerhalb des Recommenders auf einen der oft ausgelasteten Radwege geklickt, werden dort präferiert die wenig befahrenen Fahrradwege empfohlen. Diese können außerdem auch empfohlen werden, wenn andere POI in den innerstädtischen Regionen von Arnis und Kappeln angeguckt werden. Insgesamt haben die Besucher*innen somit die Möglichkeit neue, ländliche Radtouren und Aktivitäten zu entdecken, die möglicherweise vorher noch nicht bekannt waren, und gleichermaßen bietet sich ihnen ein besseres Platzangebot durch die Verteilung des Besucher*innendrucks auf unterschiedliche Punkte in der Region (vgl. Abb. 133).

6.7 Zwischenfazit

Das AP 5 "Deployment" des Projektes LAB-TOUR SH diente der Konzeption eines möglichen Ausspielkanals sowie einer ersten "echten" Live-Ausspielung inklusive eines Use Interfaces. Hierbei waren verschiedene konzeptionelle Vorarbeiten wichtig. Zunächst ging es darum zu untersuchen, welche Art der Ausspielung vielversprechend sein könnte. Neben der Feststellung, dass das Smartphone als Endgerät

eine wichtige Rolle spielt konnte zudem konstatiert werden, dass ein “Andocken” an bestehende Anwendungen sinnvoll sein kann. An welche Websites und spezifische Anwendungen dies genau erfolgt, konnte in AP 5.2 ermittelt werden, in dem sowohl regionale Websites als auch Use Case spezifische Anwendungen analysiert wurden und deren Sichtbarkeit bei der Google-Suche untersucht wurden. Da das Verfahren der Relevanzbewertung auch übertragbar auf andere Projekte gemacht werden sollte, erfolgte in AP 5.3 die Dokumentation der Vorgehensweise in der Form, dass sie von anderen, ähnlich gelagerten Projekten mit aktuellen Studien übernommen werden könnten.

Da das Projekt LAB-TOUR SH jedoch nicht darin verharren wollte, sich auf eine konzeptionelle Ebene zu beschränken, wurde eine Live-Umgebung programmiert, in der die gesamte Architektur und Konzeption für die Endnutzenden sichtbar wurde. Der Recommender konnte so zum einen von der Datenerhebung bis zur Ausspielung von Empfehlungen erstmals bereits in diesem Projekt umgesetzt und getestet werden.

Dieser pilothafte Recommender wurde zudem über das #sh_wlan ausgespielt und konnte so in einer Live-Umgebung eingesetzt werden. Die Ausspielung erfolgte konkret am Parkplatz Strandweg Mönkeberg, wo zudem eine andere Art von Sensorik verbaut wurde (vom Hersteller Libelium), was eine zusätzliche Herausforderung darstellte. In diesem Live-Test konnte so dann auch geprüft werden, ob die Datenarchitektur, die in diesem Projekt entwickelt wurde, interoperabel und damit übertragbar auf anders geartete Sensorik sowie POI und weitere wichtige zu integrierende Parameter ist. Die Realisierung dieser Umgebung war somit auch ein PoC der erfolgreich verlief und bewies, dass die hier entwickelte Umgebung ausgebaut werden kann und weitere Sensorik und andere Datenquellen angebunden werden können, was für das Projekt elementar ist.

Abschließend wurden verschiedene weitere Testszenarien entwickelt, die Ansatzpunkte für eine weitere Implementierung bieten, wodurch die Ergebnisse dieses Projektes weiterentwickelt werden können und sogar schon definiert ist, in welcher Form dies erfolgen könnte.

7. Fazit

7.1 Kernergebnisse und Forschungsdesiderate

Im Rahmen von LAB-TOUR SH wurden in den verschiedenen Arbeitspaketen unterschiedliche Teilziele verfolgt und erreicht. Das hier vorliegende integrative Gesamtkonzept enthält alle wesentlichen Ergebnisse und Dokumentationen. Das Konzept dient als Grundlage für weitere Forschung zum digitalen Besuchermanagement im Tourismus sowie zur Umsetzung dieser Ergebnisse in der Praxis. Wesentliche Bestandteile können unmittelbar im Forschungsprojekt AIR weiterentwickelt werden. Im Arbeitspaket 1 “Messung und Sensorik” wurden vier verschiedene Use Cases, bei denen es Herausforderungen im Besuchermanagement gab, identifiziert und analysiert. Darüber hinaus konnte an diesen Standorten lokale Sensorik verbaut werden, die über den Projektzeitraum hinaus wichtige Daten über Besucherfrequenzen sendet. Mit Hilfe globaler Forschungsdaten konnten Zeitreihen analysiert und so sehr genau nachvollzogen werden, wie sich Tourist*innen an den identifizierten Hotspots in Schleswig-Holstein verhalten. Diese Arbeiten führen in Summe nicht nur zu einem besseren Verständnis der Besucherströme in Schleswig-Holstein, sondern stehen darüber hinaus öffentlich zur Verfügung und können so von der Praxis und für weitere Forschungen genutzt werden.

Statt gefühlter Wahrheiten konnten im Rahmen von LAB-TOUR SH erstmals für ausgewählte Hotspots in Schleswig-Holstein Saisonalitäten von Besucherfrequenzen nachgewiesen werden. Dabei besteht nun und auch für die Zukunft die Möglichkeit, die Saisonalität in Echtzeit zu messen, zu analysieren und daraus evidenzbasiert strategische Maßnahmen für das Destinationsmanagement abzuleiten. Es wurden signifikante saisonale Schwankungen und die Einflüsse von Feiertagen, Wochentagen und

Temperaturen auf die Nutzungsfrequenzen nachgewiesen. Bei den Parkplätzen an der Ostsee und Hotspots an der Schlei wurden höhere Frequenzen während der Sommermonate und an Feiertagen gemessen. Selbst wenn dies zu erwarten war, konnte dies nun erstmals validiert und mit konkreten Zahlen hinterlegt werden sowie hierzu Analysen der Wirkung möglicher Einflussfaktoren (Wetter, Feiertage usw.) durchgeführt werden. Wohnmobilstellplätze in der Holsteinischen Schweiz zeigten ebenfalls eine starke Saisonalität, und die Nutzung der Radrouten an der Schlei stieg an Feiertagen deutlich an. Dies ist keine überraschende Erkenntnis, gleichwohl konnte diese nun erstmals genau quantifiziert werden.

Trotz der gewonnenen Erkenntnisse sind bei der digitalen Messung von Besucherströmen weiterhin Herausforderungen zu bewältigen. Hierzu gehören ökonomische Herausforderungen wie Installations- und Genehmigungskosten sowie zeitliche und personelle Ressourcen, sozio-politische Herausforderungen wie die Abhängigkeit von Sensorik- bzw. Datenanbietern und die Notwendigkeit, wichtige Stakeholder wie politische Entscheidungsträger*innen von dem Mehrwert solcher Installationen zu überzeugen. Technische bzw. datenbezogene Herausforderungen, die die Zuverlässigkeit und Validität der Daten betreffen, sind ebenfalls nicht zu vernachlässigen. Für eine optimale Datenerfassung wird empfohlen, künftig unterschiedliche Sensortypen zu verwenden und diese durch manuelle Zählungen zu validieren. Insgesamt ermöglicht die Kombination von lokaler und globaler Sensorik also eine datengesteuerte Bewertung von Besucherströmen. Durch das Zusammenspiel dieser beiden Quellen konnte ein ganzheitliches Verständnis des Besucherverhaltens entwickelt werden. Weitere Forschungen sollten sich mit Verfeinerung der Definitionen und dem tiefergehenden Vergleich von lokalen Sensordaten und Big Data-Quellen beschäftigen. All diese Erkenntnisse bieten einen deutlichen Mehrwert für den Tourismus in Schleswig-Holstein, dem es nun möglich ist, auf Basis der Projektergebnisse weitere Hotspots in den Blick zu nehmen und die hier gewonnenen Erkenntnisse für weitere Use Cases zu berücksichtigen.

Der Datenfluss von der Erhebung bis zur Auspielung in einem Frontend (Recommender) im digitalen Besuchermanagement ist von zentraler Bedeutung. Daher wurden im dritten Arbeitspaket Plattformen, Schnittstellen und Webservices bereitgestellt, die einen Datenaustausch ermöglichen. Es wurden verschiedene Datenhubs eingebunden, Schnittstellen normiert und eine Erweiterbarkeit für mögliche neue Sensorikquellen sichergestellt. Die modulare Konfiguration der Konnektoren, die hierfür entwickelt worden ist, kann ebenfalls auf noch zukünftig hinzukommende Datenhubs ausgeweitet werden. Speicherpunkte und die Verbindungen der Komponenten wurden hergestellt und die essentiellen Schnittstellen für die Kommunikation zwischen den Komponenten beschrieben und getestet. Ebenfalls deutlich wurde die Wichtigkeit von Datenmodellen, wobei eine Strategie zur Identifizierung geeigneter Datenmodelle vorgestellt und ein Prozess für den Fall definiert wurde, dass kein passendes Datenmodell existiert. Die in LAB-TOUR SH neu definierten Datenmodelle wurden eingeführt und die mögliche Integration dieser Modelle in größere Ontologien diskutiert. Die zugrundeliegende Plattform, insbesondere die Integrationsplattform, die Daten aus verschiedenen Datenhubs zusammenbringt, wurde ebenfalls erklärt. Die Plattform bietet persistente und sichere Datenspeicherung durch zuverlässige Speicherstrukturen und dezentrale Datenhaltung. Die Implementierung in Kubernetes stellt sowohl Ausfallsicherheit als auch eine Skalierbarkeit sicher. Die ausgewählten Schnittstellen und Datenmodelle sowie das Microservice-Designprinzip tragen zusätzlich zur Übertragbarkeit und dem Ausrollen in großem Maßstab bei. Damit wurden wesentliche Ziele wie Skalierbarkeit, Persistenz und hohe Verfügbarkeit erreicht. Diese in LAB-TOUR SH prototypische Architektur bildet eine wichtige Grundlage für die Algorithmen des Recommenders, aber auch für weitere Forschung sowie Umsetzung in der Praxis in diesem Bereich.

Mit dem Aufsetzen eines Pilot-Recommendere und dem Frontend im Arbeitspaket 4 wurde in LAB-TOUR SH eine prototypische, aber vollständige End-to-End-Lösung für den Tourismus in Schleswig-Holstein implementiert. Der Recommender zielt darauf ab, die Attraktivität von Ausflügen zu erhöhen, Überfüllung zu reduzieren und den Verkehr bei der Parkplatzsuche zu minimieren. Dadurch wurde erstmals in Schleswig-Holstein eine Grundlage geschaffen, an der konkret die nächsten Schritte für ein effektives Gesamtsystem diskutiert und geplant werden können. Die hier implementierte Lösung ist modular aufgebaut und kann iterativ dort angepasst werden, wo Anpassungsbedarf besteht. Das entwickelte System ist durch eine Reihe von komplexen und innovativen Komponenten gekennzeichnet. Es beinhaltet eine automatisierte, kontinuierliche Abfrage von Auslastungsdaten aus verschiedenen Quellen, die für ein breites Spektrum an Anwendungen nützlich sind. Diese Auslastungsdaten werden weiter mit ortsabhängigen Kontextinformationen (bspw. Wetterdaten) und zeitlichen Spezifikationen angereichert, um eine präzise Analyse zu ermöglichen. Zu den Hauptmerkmalen des Recommenders gehört eine qualitativ hochwertige, automatisierte Auslastungsprognose. Diese Prognose bietet einen umfassenden Überblick über die zu erwartende Auslastung und ermöglicht so eine bessere Planung und Steuerung. Zusätzlich ist das System mit einem Meta-Knowledge-Graph ausgestattet, der Informationen über POI enthält und aus diesen Ähnlichkeiten und Unterschiede darstellt, die für einen Empfehlungsdienst elementar sind - diese Basis kann ebenfalls flexibel erweitert werden. Ein weiteres innovatives Feature ist ein Algorithmus zur Empfehlung alternativer POI. Dieser berücksichtigt die Live-Auslastung der POI sowie das Verhältnis der POI zueinander im Meta-Knowledge-Graph und bietet so eine wertvolle Entscheidungshilfe für Nutzer*innen. Zur weiteren Verbesserung der Funktionalität und Effizienz werden die POI-Metadaten, Prognosen und Empfehlungen auf einem FIWARE Context Broker zwischengespeichert. Eine Schnittstelle für den FIWARE Context Broker ist ebenfalls Teil des Systems und gewährleistet einen reibungslosen Informationsaustausch. Im Rahmen eines PoC wurde ein ansprechendes Frontend entwickelt, das die Daten sichtbar macht und das Nutzungspotential des Recommenders demonstriert.

Für dieses Frontend wurden in Arbeitspaket 5 konzeptionelle Vorarbeiten durchgeführt. Hierzu wurden zunächst Studien recherchiert, die Aufschluss darüber geben, über welche Endgeräte sich Nutzende wann informieren und welche Informationen sie dabei noch abrufen. Im Anschluss daran wurden diverse mögliche Ausspielkanäle identifiziert und deren Potenzial für einen Recommender abgeschätzt. Zusätzlich wurde der Funktionsumfang ähnlicher Anwendungen analysiert und deren Potenzial als erweiterte Funktion im Rahmen der Recommenderentwicklung eingeschätzt. Dieses Verfahren wurde so abstrahiert, dass es auf andere Projekte übertragbar ist. Somit konnte die konzeptionelle Grundlage für das Design des Frontends gelegt und darüber hinaus Ableitungen getroffen werden, wie eine Einbindung eines Recommenders in andere bestehende Websites und andere digitale Anwendungen (Apps etc.) aussehen könnte.

Die Pilotierung des Recommenders erfolgte sodann auf drei Ebenen: Erstens konnte ein exemplarisches Frontend entwickelt werden, das einen Eindruck darüber gibt, wie ein Recommender in der Praxis aussehen könnte. In die Gestaltung wurden die konzeptionellen Vorüberlegungen einbezogen, aber es bleiben Erweiterungen offen, die bei einer möglichen Weiterentwicklung des Frontends noch integriert werden können. Zweitens wurde dieses Frontend über das #sh_wlan ausgespielt und war damit bereits in einer Live-Umgebung im Einsatz, sodass ein realer Praxistest erfolgen konnte. Schließlich wurden Szenarien entwickelt, die bei einem weiteren Roll-Out als Ansatzpunkte dienen können, in welcher Form dies erfolgen kann. Aufgrund der kurzen Projektlaufzeit von anderthalb Jahren war das Forschungsprojekt LAB-TOUR SH so ausgelegt, dass es einen konzeptionellen Charakter hat. Gleichwohl sind hinter den Konzepten reale Umgebungen hinterlegt, die hier ausführlich dokumentiert werden konnten. Die Umsetzung erfolgte neben dem hier vorliegenden Endbericht in Form von weiteren Forschungsdesideraten:

- 5 Leitfäden zum Digitalen Besuchermanagement für die Destinationspraxis, in denen die wesentlichen Schritte, Herausforderungen und Umsetzungsmöglichkeiten als Erfahrungsberichte aus dem Projekt in einfacher Sprache, aber wissenschaftlich fundiert, der Öffentlichkeit zur Verfügung gestellt werden. Die Ratgeber sind auf der Forschungsplattform Zenodo einsehbar: <https://zenodo.org/communities/labtour>
- Frequenzmessungen in den Use Cases inkl. Zugang zu Eco Visio für Verbund- und Destinationspartner*innen sowie weitere Forschungsprojekte (www.eco-visio.net). Die Daten stehen darüber hinaus auch auf dem Open Data Portal zur Verfügung (<https://opendata.schleswig-holstein.de>). Bei ausreichender Batterielaufzeit laufen die Frequenzmessungen bis Mitte des Jahres 2024 weiter.
- Anonymisierte Location Events aus App-Daten für 38 Polygone in Schleswig-Holstein zum Download auf der Website des Projektes
- Datenschema für digitales Besuchermanagement im Tourismus: <https://github.com/LAB-TOUR-SH/dvm>
- Darüber hinaus sind im Rahmen von LAB-TOUR SH zwei wissenschaftliche Beiträge in der internationalen Forschungslandschaft erschienen (Schmücker, Reif, Horster et al., 2023; Reif et al., im Druck) sowie der Wissenstransfer im Rahmen von nationalen und internationalen Forschungskongressen sowie Praxisveranstaltungen sichergestellt.

7.2 Implikationen für die Forschung: Ausgewählte weitere Ansätze

Die Effektivität der hier vorgestellten Möglichkeiten der Prognose, Alternativengenerierung und Auspielungsmöglichkeiten konnte im Rahmen des Projektes nicht überprüft werden. Hierfür ist es notwendig, die hier aufgeführten Daten und Modelle längerfristig in ein reichweitenstarkes Live-System zu überführen. Weitere Forschungsprojekte mit experimentellem Design sind hier gefordert, um die tatsächliche Lenkungswirkung solcher digitalen Recommendersysteme zu evaluieren. Denkbar sind konkrete Interventionen auf Destinationsseiten, die es erlauben, eine Lenkungswirkung zu analysieren. Beispielsweise könnten mit Hilfe digitaler Nudges (Pop-Up Fenster, besondere Markierung etc.) Parkplätze digital als zu voll markiert werden. Zugriffszahlen sowie Besuchermessungen vor und nach den Interventionen könnten Aufschluss über die Effektivität geben. Solche Nudges sind jedoch auch nur ein Teilbereich möglicher Eskalationsstufen im Rahmen eines digitalen Besuchermanagements (Schmücker et al., 2023). Weitere Interventionen, beispielsweise die Online-Reservierung von Parkplätzen, die (künstliche) Verknappung von Tickets oder Stellplätzen, die Schließung von Straßen oder auch die Installation von ansprechenden Design-Elementen (Bänke, Aussichtsplattformen etc.) können Aufschluss über eine mögliche Beeinflussung der Aktionsräume von Tourist*innen in der Destination geben. Hier stehen Forschung und Praxis noch am Anfang. Die Prinzipien der Besucherlenkung sind dabei nicht neu, wie die aufgeführten möglichen Interventionen aufzeigen. Gleichwohl ist es mit dem hier entwickelten konzeptionellen Rahmen möglich, eine Übertragung in eine moderne und konstant messende digitale Umgebung zu erreichen und so die Grundlage für ein langfristig angelegtes digitales Besuchermanagement zu legen. Gewissermaßen stellen die Ergebnisse des Projektes LAB-TOUR SH damit eine digitale Infrastrukturmaßnahme dar, die es dem Tourismus in Schleswig-Holstein in Zukunft ermöglicht, eine moderne Form der Marktforschung mit Echtzeitdaten zu realisieren und so wertvolle und vor allem verifizierbare Implikationen für das Besuchermanagement zu generieren.

Neben eher praktischen Möglichkeiten der Analyse von Maßnahmen zur Besucherlenkung, am Beispiel von konkreten Interventionen in den Destinationen, können somit auch Simulationen durchgeführt werden, die helfen, das komplexe aktionsräumliche Verhalten von Gästen vor Ort zu verstehen

und entsprechend zu beeinflussen. Zukünftige Projekte können hier den Fokus auf bestimmte Szenarien legen und untersuchen, wie sich Tourist*innen bei der Annahme von bestimmten Prämissen verhalten. Gäste, aber auch Anbieter von Dienstleistungen in der Destination, nehmen in diesem Beispiel die Rolle von Agenten ein, die bestimmte Verhaltensmerkmale zugeschrieben bekommen und entsprechend dieser agieren. So könnten im Rahmen einer derartigen Agentenbasierten Modellierung (ABM) die Besucherströme entlang der Schlei simuliert werden, wenn Fähren ausfallen, oder wie sich die Verkehrsflüsse in Scharbeutz ändern, wenn Parkplatzkapazitäten an den Stränden erweitert oder verknappt werden. Die im Rahmen von LAB-TOUR SH generierten Forschungsdaten können dabei helfen, die verschiedenen Prämissen für solche Modelle zu verbessern. In der internationalen Literatur gibt es bereits erste Ansätze hierfür (Li et al., 2020). Für deutsche Reiseziele, aber auch für Mega-Events, beispielsweise die Kieler Woche oder die Bundesgartenschauen, wird hier ein großes Potenzial gesehen.

Weitere Forschungsansätze ergeben sich bei der Ähnlichkeitsbestimmung von POI. Die im Rahmen des Forschungsprojektes genutzte Bestimmung von Ähnlichkeiten bildet eine wichtige Grundlage. Zukünftige Forschungsprojekte sollten hierauf aufbauen, indem beispielsweise mit Hilfe des Place2Vec-Ansatzes die Ähnlichkeiten zwischen POI bestimmt werden. Dabei handelt es sich um ein auf Basis von Word2Vec basierendes neuronales Netz, welches die Wörter in Vektoren abbildet und die semantischen Beziehungen erfasst (Egger, 2022). Voraussetzungen hierfür sind eine qualitativ hochwertige Datenbeschreibung der POI sowie eine ausreichend große Menge an Daten. Mit der Verortung der POI in diesem Projekt im Meta-Knowledge-Graph besteht hier schon eine Basis, die weiter ausgebaut werden kann - auch und insbesondere innerhalb von Schleswig-Holstein.

7.3 Implikationen für die Praxis

Für die Praxis ergeben sich weiterhin vielfältige Herausforderungen in der Einführung von Mess- und Managementsystemen (Reif et al., 2023). Dabei zeigen sich neben Herausforderungen auf der Daten-seite (bspw. Validität, Reliabilität), ökonomische (bspw. Kosten für Sensorik) oder auch sozio-politische (bspw. Datenschutz) Problembereiche. Trotz dieser Schwierigkeiten können die gesammelten Daten über das Verhalten von Gästen in Raum und Zeit einen unschätzbaren Mehrwert für Reiseziele liefern. Dies gilt insbesondere für interne Prozesse wie die Personalplanung, Saisonplanung, Marketingaktivitäten etc. Dies erfordert jedoch einen Perspektivwechsel: Während der bislang in diesem Projekt skizzierte Ansatz eines digitalen Besuchermanagements die Ausspielung von Prognosen und Empfehlungen an den Endkunden im Fokus hatte (B2C), können die aufgebauten Strukturen natürlich auch im B2B-Bereich genutzt werden. Ein gutes Beispiel, wie derartige Prozesse funktionieren können, zeigt ein Projekt auf der dänischen Insel Rømø (Jacobsen et al., 2020). So können beispielsweise Daten zu Eisverkäufen, Daten aus Buchungssystemen der Ferienwohnungen sowie lokale und globale Sensorik in einem Destinations-Dashboard so kombiniert werden, dass konkrete Maßnahmen ableitbar sind. Ansätze wären hier die Anpassung der Öffnungszeiten an aktuelle Besucherströme oder das bessere Handling von Schlüsselübergaben zur Vermeidung von Verkehrsstaus. Durch derartige Projekte wird deutlich, dass die Art der Nutzung von Sensorikdaten sowie deren Verarbeitung noch stark in der Entwicklung befindlich sind. Gleichwohl scheint mehr denn je auch deutlich zu werden, dass die Ergebnisse der konzeptionellen Entwicklungen im Projekt LAB-TOUR SH die infrastrukturelle Basis für alle weiteren Entwicklungen in diesem Bereich darstellen.

LAB-TOUR SH hat deutlich gemacht, dass eine einheitliche Benennung von POI von essenzieller Bedeutung ist. Um eine optimierte Datenstrukturierung zu gewährleisten, wird dringend empfohlen, einheitliche Standards zu verwenden, die die Interoperabilität zwischen Systemen und Diensten fördern. Dies ist nicht nur relevant für die Auffindbarkeit und Sichtbarkeit im Internet, sondern auch essentiell für

die Beschreibung und Klassifizierung von POI im Rahmen des Recommenders. LAB-TOUR SH hat sich hier an bestehenden Datenmodellen (schema.org, ODTA, Smart Data Models) orientiert und diese gleichzeitig für die Use Cases weiterentwickelt, die mit den bestehenden Datenmodellen kompatibel sind. An diesen Bestand sollte angeknüpft und dieser erweitert werden. Zudem ist es wichtig, dass die Basis touristischer Daten einheitlich ausgezeichnet wird (entlang der Datenmodelle der ODTA).

Eine flächendeckende Installation von lokaler Sensorik stellt mit Blick auf die Skalierbarkeit und Effizienz große Herausforderungen dar. Der ökonomische Aufwand wird als sehr groß erachtet. Ein gezielter Einsatz von Sensoren in Modellregionen erscheint als eine tragfähigere Strategie. Die Erkenntnisse, die aus diesen Modellregionen gewonnen werden, können anschließend auf das gesamte Gebiet von Schleswig-Holstein übertragen werden – auch wenn hier noch weitere Forschung, beispielsweise zu typischen Tages- und Wochenverläufen, und die Möglichkeit der Skalierung betrieben werden muss. Zudem sollte in diesem Zusammenhang – gerade bei der Bestimmung von Parkplatzauslastungen – mit Transaktionsdaten wie beispielsweise von digitalen Paymentanbietern (bspw. Easy Park etc.) gearbeitet werden. Auch hier ist noch Forschung über die Nutzbarkeit der Daten erforderlich, gleichwohl können die Daten einen Mehrwert bieten, da sie bereits vorhanden sind. Erste Modellprojekte hierzu laufen beispielsweise mit dem Unternehmen Outdooractive in der Stadt Dortmund. Ein gemischter Ansatz mit lokalen Modellregionen und dem Einsatz von Transaktionsdaten sollte in Betracht gezogen werden.

Auch beim Deployment gibt es Herausforderungen, die Destinationen beachten sollten. Beispielsweise erfordert die direkte Anbindung der Auslastungsdaten in Progressive Web Apps (PWA) von Reisezielen eine Schnittstelle. Die Integration hängt von der erfolgreichen Authentifizierung der Clients am Context Broker ab und der Fähigkeit, Daten von diesem zu interpretieren und zu nutzen. Wenn der PWA ein "Nur Lesen" Zugang zum Context Broker gewährt wird, kann sie eigenständig die Auslastungsdaten auf Basis der POI-ID ermitteln und nutzen. Alternativ können diese Funktionen aus der PWA ausgelagert und in die Backend-Systeme der Destinationen integriert werden, um die Datenverwaltung zu vereinfachen und eine einheitliche Informationsquelle zu gewährleisten.

7.4 Ausblick

Die in LAB-TOUR SH generierten Daten, Infrastrukturen, Modelle und Dokumentationen stehen öffentlich zur Verfügung und dienen als Grundlage für weitere angewandte Forschungsprojekte im Tourismus. Im Kontext des bundesweiten Forschungsprojekts „AI-basierter Recommender für nachhaltigen Tourismus (AIR)“ werden die in LAB-TOUR SH entstandenen Daten durch das DI Tourismusforschung weiter analysiert und ausgewertet. Die Grundlagen des Prognosemoduls und des Alternativengenerierers werden von der Fachhochschule Kiel weiterentwickelt. Nach Beendigung des Projektes AIR am 31.12.2024 steht der weiterentwickelte Recommender zur Verfügung.

Im Zuge einer Förderung der Digitalisierung des Tourismus könnte der Recommender potenziell als Teil eines mittelfristigen Entwicklungsprojektes zusammen mit den Akteuren der schleswig-holsteinischen Tourismuspraxis in einen produktiven Einsatz überführt werden. Mit der TA.SH als Partner könnte die POI-Datengrundlage weiter aufgearbeitet werden, um so ein umfangreiches digitales Abbild der Tourismuslandschaft in Schleswig-Holstein zu generieren. Dieses könnte dann in einen produktionsreifen Knowledge-Graphen geladen werden, aus dem sich ein Alternativengenerierer speist, der ähnlich funktioniert wie hier skizziert. Die produktive Entwicklung eines solchen Alternativengenerierers scheint auf Grundlage der umfangreichen Vorarbeit ein realistisches Ziel einer kurz- oder mittelfristigen Entwicklungsphase.

Durch das in LAB-TOUR SH aufgesetzte Konzept für ein digitales Besuchermanagement im Tourismus sowie die Implementierung der Datengrundlage und der technischen Infrastruktur lässt sich der Weg hin zu einer öffentlichen Nutzung nun realistisch einschätzen und detailliert planen. Damit könnte Schleswig-Holstein bundesweit Vorreiter beim Einsatz von künstlicher Intelligenz im Tourismus werden und die Nachhaltigkeit und Attraktivität des Tourismus stärken.

8. Literaturverzeichnis

- Allende, H., & Valle, C. (2017). Ensemble Methods for Time Series Forecasting. In R. Seising & H. Allende-Cid (Hrsg.), *Claudio Moraga: A Passion for Multi-Valued Logic and Soft Computing. Studies in Fuzziness and Soft Computing* (Bd. 349). Springer. https://doi.org/10.1007/978-3-319-48317-7_13
- Allgemeiner Deutscher Fahrrad-Club e.V. (2022). ADFC Radreiseanalyse für das Radreisejahr 2021. https://www.adfc.de/fileadmin/user_upload/Handout_ADFC-Radreiseanalyse_2022.pdf
- ARD/ZDF-Forschungskommission. (2021). ARD/ZDF-Onlinestudie 2021. <https://www.ard-zdf-online-studie.de/archiv-1997-2021/>
- Bergler, M., Lorenz, V., Niemeijer, C., & Keller, R. (2023a). Ergebnisbericht zur Gästebefragung im Use Case Nordsee Schleswig-Holstein - Speicherkoog. <https://doi.org/10.5281/zenodo.8005520>
- Bergler, M., Lorenz, V., Niemeijer, C., & Keller, R. (2023b). Ergebnisbericht zur Gästebefragung im Use Case Nordsee Schleswig-Holstein - Büsum. <https://doi.org/10.5281/zenodo.8142615>
- Bergler, M., Lorenz, V., Niemeijer, C., & Keller, R. (2023c). Ergebnisbericht zur Gästebefragung im Use Case Nordsee Schleswig-Holstein - St. Peter-Ording. <https://doi.org/10.5281/zenodo.7957385>
- Bergzeit GmbH. (2022). Die Bergzeit Wanderstudie 2022. <https://www.bergzeit.de/wanderstudie-2022-ergebnisse/>
- Bitkom Research. (2020). Digitaler Tourismus 2020: So smart reisen die Deutschen. https://www.bitkom.org/sites/default/files/2020-03/bitkompräsentation_tourismus2020.pdf
- Bitkom Research. (2022). Digitaler Tourismus 2022: So smart reisen die Deutschen. https://www.bitkom.org/sites/main/files/2022-07/Bitkom-Charts_Digitaler_Tourismus_05.07.2022_0.pdf
- Bitkom Research. (2019, 27. Juni). 8 digitale Technologien, die den Urlaub noch besser machen [Pressemitteilung]. <https://www.bitkom.org/Presse/Presseinformation/8-digitale-Technologien-die-den-Urlaub-noch-besser-machen>
- Bitkom e.V. (2020a, 23. Juli). Das Radio ist für Autofahrer die wichtigste Informationsquelle [Pressemitteilung]. <https://www.bitkom.org/Presse/Presseinformation/Das-Radio-ist-fuer-Autofahrer-die-wichtigste-Informationsquelle>
- Bitkom e.V. (2020b, 2. Juli). So nutzen Urlauber ihr Smartphone sorglos im Ausland [Pressemitteilung]. https://www.bitkom.org/Presse/Presseinformation/So-nutzten-Urlauber-ihr-Smartphone-sorglos-im-Ausland#_
- Bundesverband Digitale Wirtschaft (BVDW) e.V. (2018). Digital Trends: Umfrage zur Nutzung des Smartphones im Urlaub. https://www.bvdw.org/fileadmin/user_upload/180711_BVDW_Digital_Trends_SmartphonenutzungUrlaubsreise.pdf

- Chawla, N.V., Bowyer, K.W., Hall, L.O., & Kegelmeyer, W.P. (2017). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–35.
<https://doi.org/10.1613/jair.953>
- DataReportal. (2022). Digital 2022 Germany. <https://datareportal.com/reports/digital-2022-germany>
- die medienanstalten - ALM GbR. (2021). Intermediäre und Meinungsbildung: Gewichtungsstudie zur Relevanz der Medien für die Meinungsbildung in Deutschland, 2021-I. https://www.die-medienanstalten.de/fileadmin/user_upload/die_medienanstalten/Forschung/Intermediaere_und_Meinungsbildung/Intermediaere_Meinungsbildung_2021-I.pdf
- DI Tourismusforschung (2023). Campingstudie 2022: Betrachtung ausgewählter Aspekte der Campingplatzwahl deutscher Camper. https://www.ditf-fhw.de/fileadmin/content/downloads/aktuelle_projekte/Dateien_ohne_Abkuerzung/DI_Tourismusforschung_Campingstudie_2022_Ergebnisbericht_final_2023.pdf
- die medienanstalten - ALM GbR. (2022). Studie 2022-I Intermediäre und Meinungsbildung. <https://www.die-medienanstalten.de/forschung/intermediaere-und-meinungsbildung>
- Egger, R. (2022). Text Representations and Word Embeddings. In R. Egger (Hrsg.), *Applied Data Science in Tourism. Tourism on the Verge*, (S. 335-361).
- Eisenstein, B., & Reif, J. (2017). Saisonalität im Deutschlandtourismus: Problemzone Küste. In B. Eisenstein, R. Schmutte, J. Reif & C. Eilzer (Hrsg.), *Tourismusatlas Deutschland* (S. 112–113). UVK.
- Eisenstein, B., Köchling, A., Reif, J., Schmücker, D., & Seeler, S. (2021). Tourismus in Zeiten der Corona-Pandemie. Ein Virus erschüttert das System Tourismus. In B. Eisenstein, J. Kampen, R. Weis, J. Reif & C. Eilzer (Hrsg.), *Tourismusatlas Deutschland*, 2. Auflage (S. 14-22).
- GfK SE. (2023). GfK DestinationMonitor Deutschland für Schleswig-Holstein. Datenauszug für das Jahr 2019; Aktivitäten bei mehrtägigen Urlaubs- und Privatreisen in Schleswig-Holstein. Per Mail am 21.02.2023 (bislang unveröffentlicht).
- Hardy, A. (2020). *Tracking tourists: Movement and mobility*. Goodfellow Publishers Ltd.
- Initiative D21 e. V. (2022). D21-Digital-Index 2021/2022: Wie digital ist Deutschland? Jährliches Lagebild zur Digitalen Gesellschaft (1. Aufl.). D21-Digital-Index: Bd. 9. Initiative D21.
- Jacobsen, P. Ø. et al. (2020). Destination:Digital. ET DATAPROJEKT PÅ RØMØ 2020. https://www.kystognaturturisme.com/sites/kystognaturturisme.com/files/2021-05/destination%20digital_0.pdf
- Krüger, M., Mecke, M., Goerge, M., Schulze Düding, J., & Dylla, C. (2021). Tourismus-Satellitenkonto (TSA) für Schleswig-Holstein 2019. Ergebnisbericht. Heide. Deutsches Institut für Tourismusforschung. https://www.ditf-fhw.de/fileadmin/content/downloads/aktuelle_projekte/DITF_TSA_Schleswig-Holstein_2019_Website.pdf
- Li, S., Yang, Y., Zhong, Z., & Tang, X. (2021). Agent-Based Modeling of Spatial Spillover Effects in Visitor Flows. *Journal of Travel Research*, 60(3), 546–563.
<https://doi.org/10.1177/0047287520930105>
- Mitas, O., Badal, R., Verhoeven, M., Verstraten, K., Graaf, L. de, Mitsova, H., Weijdema, W., & Klijs, J. (2023). Tell Me Where to Go: An Experiment in Spreading Visitor Flows in The Netherlands. *International journal of environmental research and public health*, 20(8).
<https://doi.org/10.3390/ijerph20085441>
- Moon, J., Kim, J., Kang, P., & Hwang, E. (2020). Solving the Cold-Start Problem in Short-Term Load Forecasting Using Tree-Based Methods. *Energies*, 13(4),
<http://dx.doi.org/10.3390/en13040886>

- O.trend GmbH. (2022). Kontinuierliche Marktforschung zum öffentlichen Verkehr in Schleswig-Holstein: Jahresbericht 2021 für den Nahverkehrsverbund Schleswig-Holstein GmbH. https://unternehmen.nah.sh/assets/Subsite2/Files/2022-08-23-LandesweiteMafo2021_Jahresbericht_20220822.pdf
- Postbank - eine Niederlassung der Deutsche Bank AG. (2022). Postbank Digitalstudie 2022: Internetnutzung in Corona-Zeiten. <https://www.postbank.de/unternehmen/medien/meldungen/2022/april/das-smartphone-wird-immer-mehr-zum-internet-tool-der-deutschen.html>
- Reif, J., Schmücker, D., Naschert, L. & Horster, E. (im Druck). Visitor management in tourism destinations: Current challenges in measuring and managing visitors' spatio-temporal behaviour. In M. Pillmayer, M. Hansen & M. Karl (Hrsg.), *De Gruyter Studies in Tourism. Tourism destination development: A geographic perspective on destination management and tourist demand*. De Gruyter.
- Reif, J., Schmücker, D., Horster, E. & Staubert, T. (2023). PoI – Points of Interest in Destinations. AIR Brief Reports. <https://doi.org/10.5281/zenodo.7736823>
- Schmücker, D. (2023). Datensatz: Ferientage und Feriendichte in Deutschland 2016–2023 (Version 2) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7716281>
- Schmücker, D., & Reif, J. (2021). Die zweifache digitale Transformation: Ausgewählte Datenquellen für ein digitales Besuchermanagement:: Vortragspräsentation zur DGT-Jahrestagung 2021 mit dem Thema „Tourismus und Transformation – Perspektiven einer Erlebnisökonomie im Blickwinkel von Krisen und Resilienz“.
- Schmücker, D., & Reif, J. (2022a). Digitale Besuchermessung im Tourismus: Ziele, Methoden, Bewertungen. UVK Verlag; Narr Francke Attempto Verlag GmbH + Co. KG. <https://doi.org/10.24053/9783739882079>
- Schmücker, D., & Reif, J. (2022b). Measuring tourism with big data? Empirical insights from comparing passive GPS data and passive mobile data. *Annals of Tourism Research Empirical Insights*, 3(2), 100061. <https://doi.org/10.1016/j.annale.2022.100061>
- Schmücker, D., Keller, R., Reif, J., Schubert, J., & Sommer, G. (2023). Digitales Besuchermanagement im Tourismus - Konzeptioneller Rahmen und Gestaltungsmöglichkeiten. In: Gardini, M. & Sommer, G. (Hrsg): *Digital Leadership im Tourismus*. Wiesbaden. S. 239-315.
- Schmücker, D., Reif, J., & Bergler, M. (2023). Wahrnehmung von Crowding (AIR Kurzberichte). <https://doi.org/10.5281/zenodo.7874758>
- Schmücker, D., Reif, J., & Staubert, T. (2023, in Arbeit). Drift in local sensor data. (AIR Kurzberichte). (bislang unveröffentlicht).
- Schmücker, D., Reif, J., Horster, E., Engelhardt, D., Höftmann, N., Naschert, L., & Radlmayr, C. (2023). The INPReS Intervention Escalation Framework for Avoiding Overcrowding in Tourism Destinations. *Tourism and Hospitality*, 4(2), 282–292. <https://doi.org/10.3390/tourhosp4020017>
- Seeler, S. (2023). Tourismusakzeptanz in der ortsansässigen Wohnbevölkerung 2022.: Zentrale Ergebnisse für die Regionen Schleswig-Holstein 2022. Heide. Deutsches Institut für Tourismusforschung.
- StatCounter. (2022). Desktop vs Mobile vs Tablet vs Console Market Share Germany. <https://gs.statcounter.com/platform-market-share/all/germany/#monthly-202201-202212-bar>
- StatCounter. (2023). Desktop, Mobile & Tablet Search Engine Market Share Germany: Jan 2022 - May 2023. <https://gs.statcounter.com/search-engine-market-share/desktop-mobile-tablet/germany/#monthly-202201-202305-bar>

- Statista. (2019). Statista Umfrage ÖPNV 2019: Umfrage zur Nutzung von Apps für Fahrplanauskunft des ÖPNV in Deutschland 2019. <https://de.statista.com/prognosen/1036304/umfrage-zur-nutzung-von-apps-fuer-fahrplanauskunft-des-oepnv-in-deutschland>
- Statista Consumer Insights. (2023). Global Consumer Survey 2022. <https://de.statista.com/prognosen/999785/deutschland-beliebte-arten-von-smartphone-apps?locale=de>
- Zeng, J., Li, Y., Li, J. W., & Hirokawa, S. (2017). A Point-of-Interest Recommendation Method Using Location Similarity. 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), Hamamatsu, Japan, 2017, pp. 436-440, doi: 10.1109/IIAI-AAI.2017.122.
- Zimmermann-Janschitz, S. (2014). Statistik in der Geographie. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-8274-2612-3>

8. Anhang

Anhang 1: Überblick der lokalen Sensorik

Parkplätze an der Ostsee

Use Case	Parkplätze Ostsee - Fehmarn
Name	Parkplatz Südstrand Burgtiefe
ID Polygon globale Sensorik	2
Sensor	Flex SENSEBIT-72063; Flex SENSEBIT-72064
Site Eco Counter	#27; #28
Koordinaten	54.412966, 11.209753; 54.413454, 11.212696
Anzahl Stellplätze	
Overnight-Parken	
Schranke	
Kostenpflichtig	
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	05.08.2022



Use Case	Parkplätze Ostsee - Fehmarn
Name	Parkplatz Meeschendorf
ID Polygon globale Sensorik	4
Sensor	Flex SENSEBIT-72068
Site Eco Counter	#29
Koordinaten	54.413912, 11.24317

Anzahl Stellplätze

ca. 145

Overnight-Parken

Schranke

Kostenpflichtig

Mess-Art

Richtungserkennung, Nutzungsart

Installationsdatum

05.08.2022



Use Case	Parkplätze Ostsee - Fehmarn
Name	Parkplatz Südstrand Mitte
ID Polygon globale Sensorik	3
Sensor	Flex SENSEBIT-72078; Flex SENSEBIT-72070
Site Eco Counter	#30; #31
Koordinaten	54.411685, 11.202906; 54.411691, 11.20291
Anzahl Stellplätze	
Overnight-Parken	
Schranke	
Kostenpflichtig	
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	05.08.2022



Use Case	Parkplätze Ostsee - Fehmarn
Name	Parkplatz Orth
ID Polygon globale Sensorik	6
Sensor	Flex SENSEBIT-72072; Flex SENSEBIT-72073
Site Eco Counter	#33; #34
Koordinaten	54.449155, 11.050856; 54.44859, 11.050763
Anzahl Stellplätze	ca. 180
Overnight-Parken	
Schranke	
Kostenpflichtig	
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	27.08.2022; 05.08.2022



Use Case	Parkplätze Ostsee - Fehmarn
Name	Parkplatz Grüner Brink
ID Polygon globale Sensorik	7
Sensor	Flex SENSEBIT-72080
Site Eco Counter	#36
Koordinaten	54.51186, 11.182638
Anzahl Stellplätze	ca. 54
Overnight-Parken	
Schranke	
Kostenpflichtig	
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	11.08.2022



Use Case	Parkplätze Ostsee - Fehmarn
Name	Parkplatz Gold
ID Polygon globale Sensorik	5
Sensor	Flex SENSEBIT-72076
Site Eco Counter	#37
Koordinaten	54.424145, 11.101173
Anzahl Stellplätze	ca. 300
Overnight-Parken	
Schranke	

Kostenpflichtig

Mess-Art

Installationsdatum

Richtungserkennung, Nutzungsart

05.08.2022



Use Case	Parkplätze Ostsee - Grömitz
Name	Parkplatz Seebrückenvorplatz Seestraße 31
ID Polygon globale Sensorik	9
Sensor	Flex SENSEBIT-72079
Site Eco Counter	#39
Koordinaten	54.145299, 10.962139
Anzahl Stellplätze	ca. 156
Overnight-Parken	ja
Schranke	nein
Kostenpflichtig	ja
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	09.10.2022



Use Case	Parkplätze Ostsee - Grömitz
Name	Parkplatz Königsredder Straße
ID Polygon globale Sensorik	8
Sensor	Flex SENSEBIT-72074; Flex SENSEBIT-72081
Site Eco Counter	#40; #41
Koordinaten	54.138895, 10.943519; 54.138824, 10.945025
Anzahl Stellplätze	ca. 430
Overnight-Parken	ja
Schranke	nein
Kostenpflichtig	ja
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	05.08.2022; 11.08.2022



Use Case	Parkplätze Ostsee - Grömitz
Name	Großraumparkplatz
ID Polygon globale Sensorik	11
Sensor	Flex SENSEBIT-72082; Flex SENSEBIT-72083; Flex SENSEBIT-72084; Flex SENSEBIT-72085; Flex SENSEBIT-72086
Site Eco Counter	#42a; #42b; #43; #44; #45

Koordinaten	54.145264, 10.952132; 54.145234, 10.952088; 54.143736, 10.952992; 54.143223, 10.953733; 54.144916, 10.956096
Anzahl Stellplätze	ca. 1080
Overnight-Parken	ja
Schranke	nein
Kostenpflichtig	ja
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	11.08.2022



Use Case	Parkplätze Ostsee - Grömitz
Name	Parkplatz Deichweg
ID Polygon globale Sensorik	10
Sensor	Flex SENSEBIT-72090; Flex SENSEBIT-72091; Flex SENSEBIT-72092; Flex SENSEBIT-72093;
Site Eco Counter	#46; #47; #48; #49
Koordinaten	54.153374, 10.977295; 54.153406, 10.977269;

	54.152503, 10.97888;
	54.152515, 10.978946
Anzahl Stellplätze	ca. 250
Overnight-Parken	ja
Schranke	nein
Kostenpflichtig	nein
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	09.10.2022



Hotspots an der Schlei

Use Case	Hotspots an der Schlei
Name	Parkplatz Ellenberg
ID Polygon globale Sensorik	16
Sensor	Flex SENSEBIT-73140
Site Eco Counter	#56
Koordinaten	54.6599937, 9.9417884
Anzahl Stellplätze	ca. 60
Overnight-Parken	ja
Schranke	nein
Kostenpflichtig	nein
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	06.11.2022



Use Case	Hotspots an der Schlei
Name	Großparkplatz Kappeln
ID Polygon globale Sensorik	12
Sensor	Flex SENSEBIT-73141; Flex SENSEBIT-73142; Flex SENSEBIT-73143; Flex SENSEBIT-73144
Site Eco Counter	#52; #53; #54; #55
Koordinaten	54.659427, 9.932983; 54.659381, 9.931433; 54.659509, 9.931191; 54.659514, 9.929319
Anzahl Stellplätze	ca. 190
Overnight-Parken	ja
Schranke	nein
Kostenpflichtig	Der obere nein/untere ja
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	06.11.2022



Use Case	Hotspots an der Schlei
Name	Arnis Parkplatz Neuer Damm
ID Polygon globale Sensorik	15
Sensor	Flex SENSEBIT-73145; Flex SENSEBIT-73146;
Site Eco Counter	#58a; #58b
Koordinaten	54.631327, 9.93047; 54.631327; 9,9305
Anzahl Stellplätze	ca. 80
Overnight-Parken	ja
Schranke	nein
Kostenpflichtig	Ja, von 9-18 Uhr
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	06.11.2022



Wohnmobilstellplätze in der Holsteinischen Schweiz

Use Case	Wohnmobilstellplätze
Name	WoMo Stop Malente
ID Polygon globale Sensorik	17
Sensor	Flex SENSEBIT-73883; Flex SENSEBIT-73882; Flex SENSEBIT-73876; Flex SENSEBIT-73884;

Site Eco Counter	Flex SENSEBIT-73886; Flex SENSEBIT-73885
Koordinaten	#1 54.171207, 10.550127; 54.172278, 10.549735; 54.172259, 10.549628; 54.171327, 10.550054; 54.171149, 10.549463; 54.171174, 10.549537
Anzahl Stellplätze	5
Overnight-Parken	
Schranke	ja
Kostenpflichtig	
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	31.01.2023



Use Case	Wohnmobilstellplätze
Name	Reisemobilpark Eutin
ID Polygon globale Sensorik	1
Sensor	Flex SENSEBIT-73154
Site Eco Counter	#3
Koordinaten	54.133837, 10.629551
Anzahl Stellplätze	24

Overnight-Parken	ja
Schranke	ja
Kostenpflichtig	ja
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	06.11.2022



Use Case	Wohnmobilstellplätze
Name	Spitzenort Naturcamping Plön
ID Polygon globale Sensorik	19
Sensor	Flex SENSEBIT-73148; Flex SENSEBIT-73147
Site Eco Counter	#4a; #4b
Koordinaten	54.147286, 10.398096; 54.147212, 10.397426
Anzahl Stellplätze	12 vor der Schranke 17 hinter der Schranke
Overnight-Parken	ja
Schranke	ja
Kostenpflichtig	ja
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	13.11.2022; 07.11.2022



Use Case	Wohnmobilstellplätze
Name	Naturcamping Prinzenholz Eutin
ID Polygon globale Sensorik	13
Sensor	Flex SENSEBIT-73149; Flex SENSEBIT-73150
Site Eco Counter	#5; #6
Koordinaten	54.159993, 10.60219; 54.16107, 10.601992
Anzahl Stellplätze	2 vor der Schranke 128 hinter der Schranke (inkl. Mietobjekte)
Overnight-Parken	ja
Schranke	ja
Kostenpflichtig	ja
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	07.11.2022; 08.11.2022



Use Case	Wohnmobilstellplätze
Name	Campingplatz Lange Wisch Ascheberg
ID Polygon globale Sensorik	18
Sensor	Flex SENSEBIT-73151; Flex SENSEBIT-73152
Site Eco Counter	#7a; #7b
Koordinaten	54.147123, 10.342858; 54.14711, 10.342751
Anzahl Stellplätze	5 vor der Schranke 141 hinter der Schranke
Overnight-Parken	ja
Schranke	ja
Kostenpflichtig	ja
Mess-Art	Richtungserkennung, Nutzungsart
Installationsdatum	06.11.2022



Radrouten Schlei

Use Case	Radrouten Schlei
Name	Wikinger Törn/Weseby
ID Polygon globale Sensorik	
Sensor	ZLT22080002
Site Eco Counter	#11
Koordinaten	54,500531; 9,715287
Mess-Art	Richtungserkennung, Nutzungsart Fahrräder
Installationsdatum	17.05.2023

Use Case	Radrouten Schlei
Name	Fähre Arnis/Sundsacker
ID Polygon globale Sensorik	
Sensor	ZLT22080009
Site Eco Counter	#14
Koordinaten	54,62876; 9,936704
Mess-Art	Richtungserkennung, Nutzungsart Fahrräder
Installationsdatum	21.02.2023

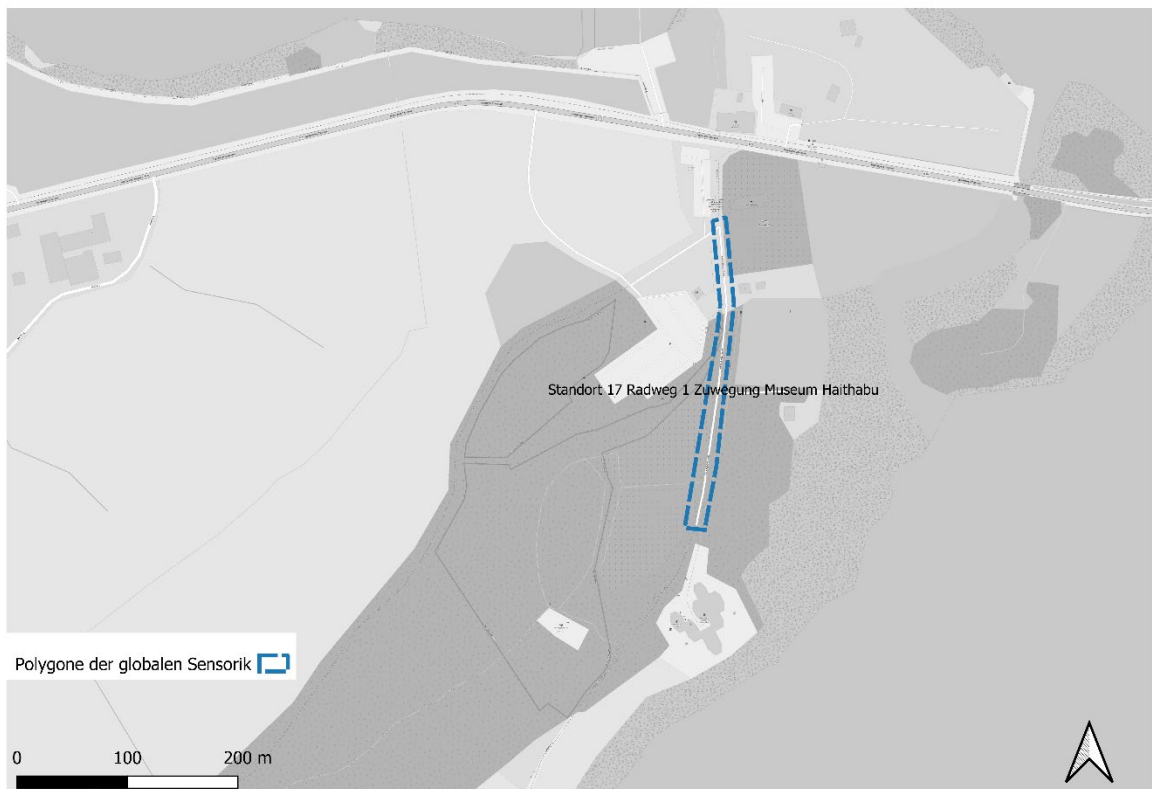
Use Case	Radrouten Schlei
Name	Am Haddebyer Noor
ID Polygon globale Sensorik	

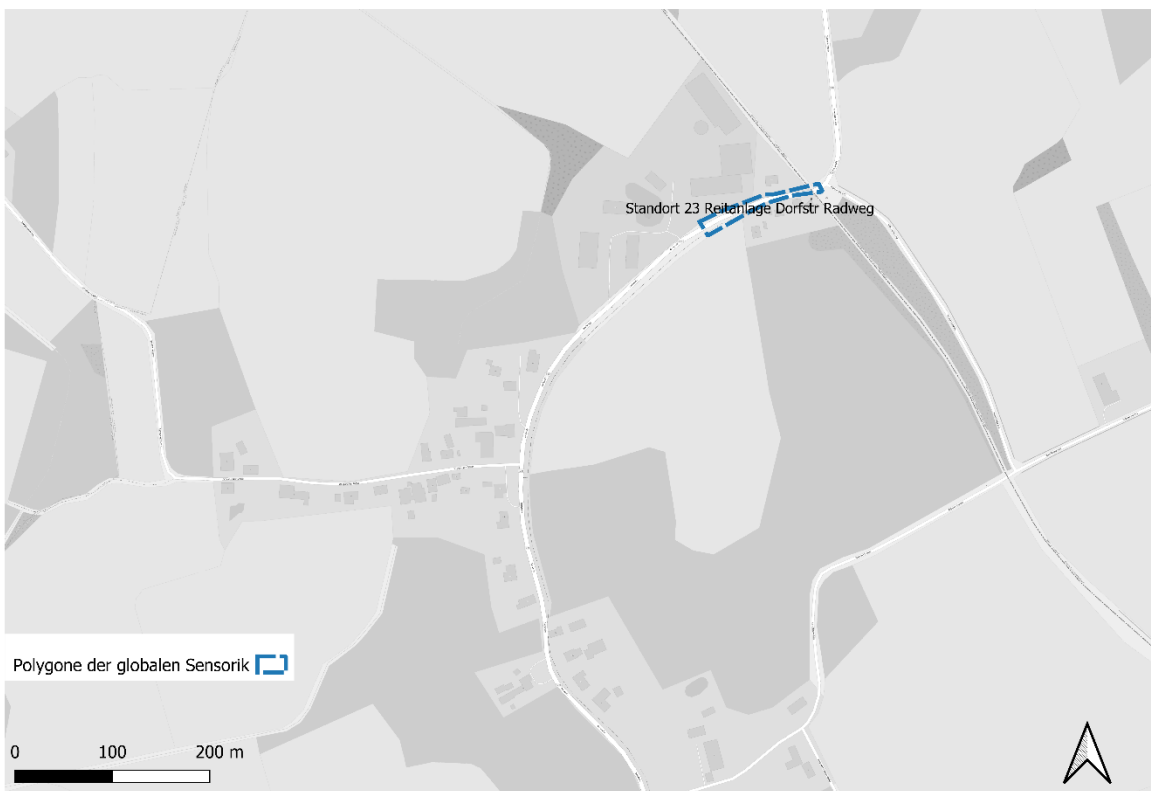
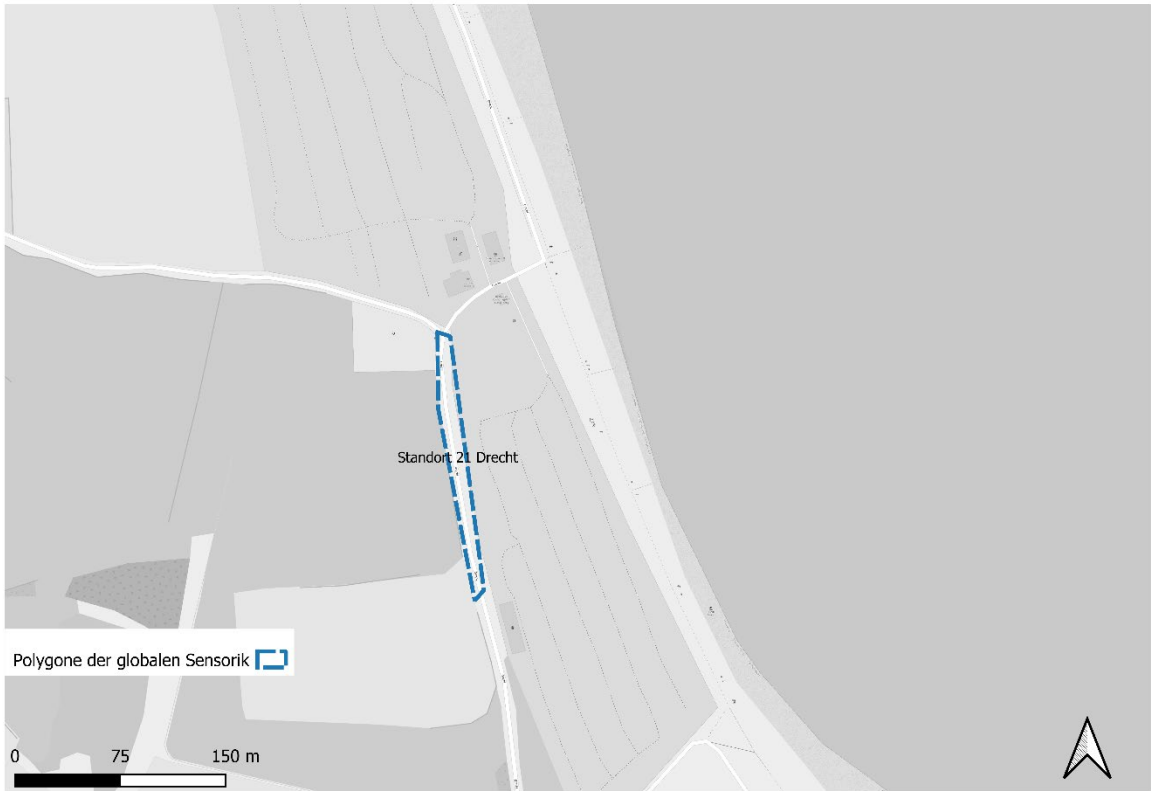
Sensor	ZLT22080011
Site Eco Counter	#17
Koordinaten	54,499086; 9,569664
Mess-Art	Richtungserkennung, Nutzungsart Fahrräder
Installationsdatum	27.04.2023

Use Case	Radrouten Schleie
Name	Hasselberg
ID Polygon globale Sensorik	
Sensor	ZLT22080010
Site Eco Counter	#21
Koordinaten	54,7191797623998; 9,98619675636292
Mess-Art	Richtungserkennung, Nutzungsart Fahrräder
Installationsdatum	27.04.2023

Use Case	Radrouten Schleie
Name	Gulde
ID Polygon globale Sensorik	
Sensor	ZLT22080015
Site Eco Counter	#22
Koordinaten	54,689105; 9,837617
Mess-Art	Richtungserkennung, Nutzungsart Fahrräder
Installationsdatum	01.05.2023

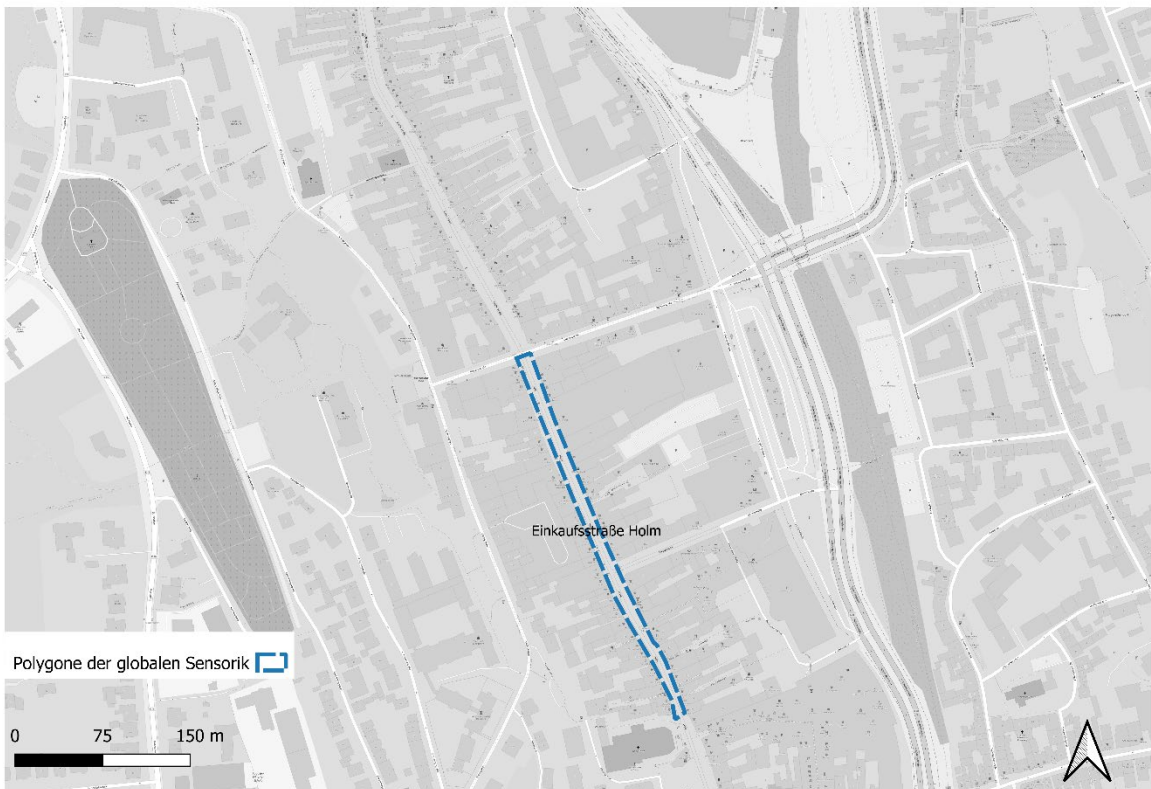
Anhang 2: Überblick der Polygone der globalen Sensorik

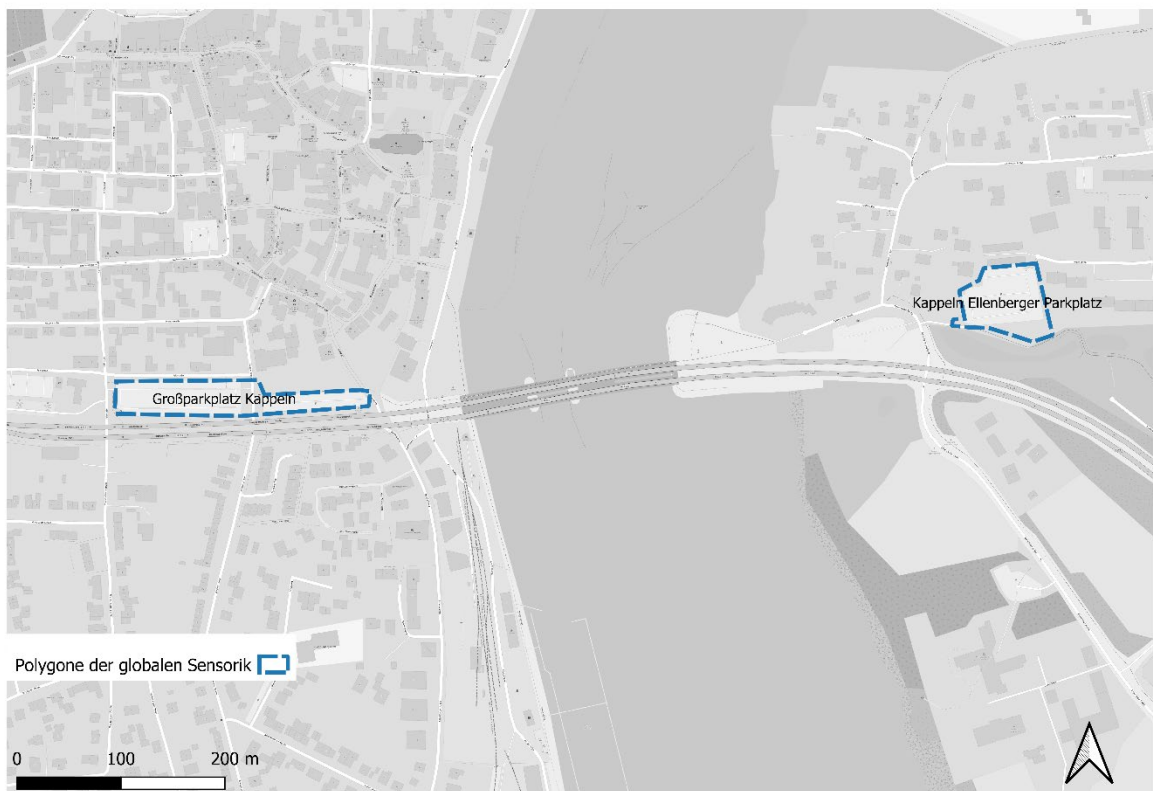




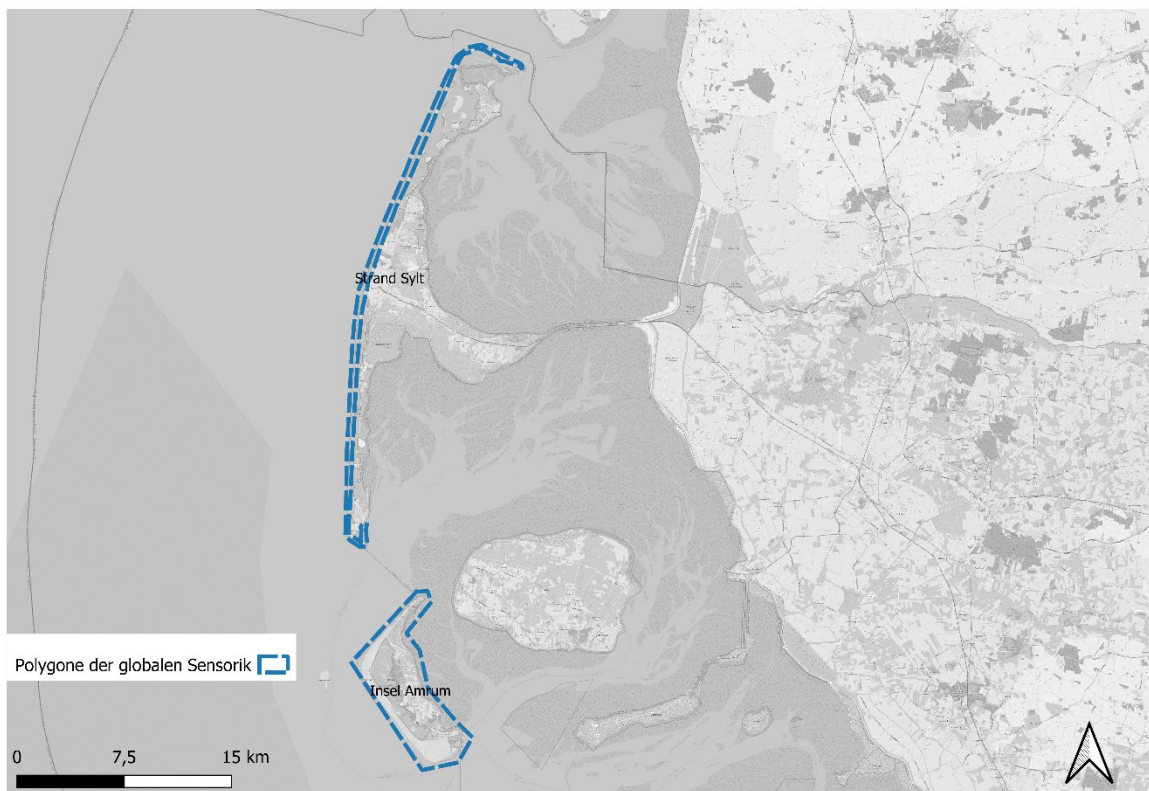


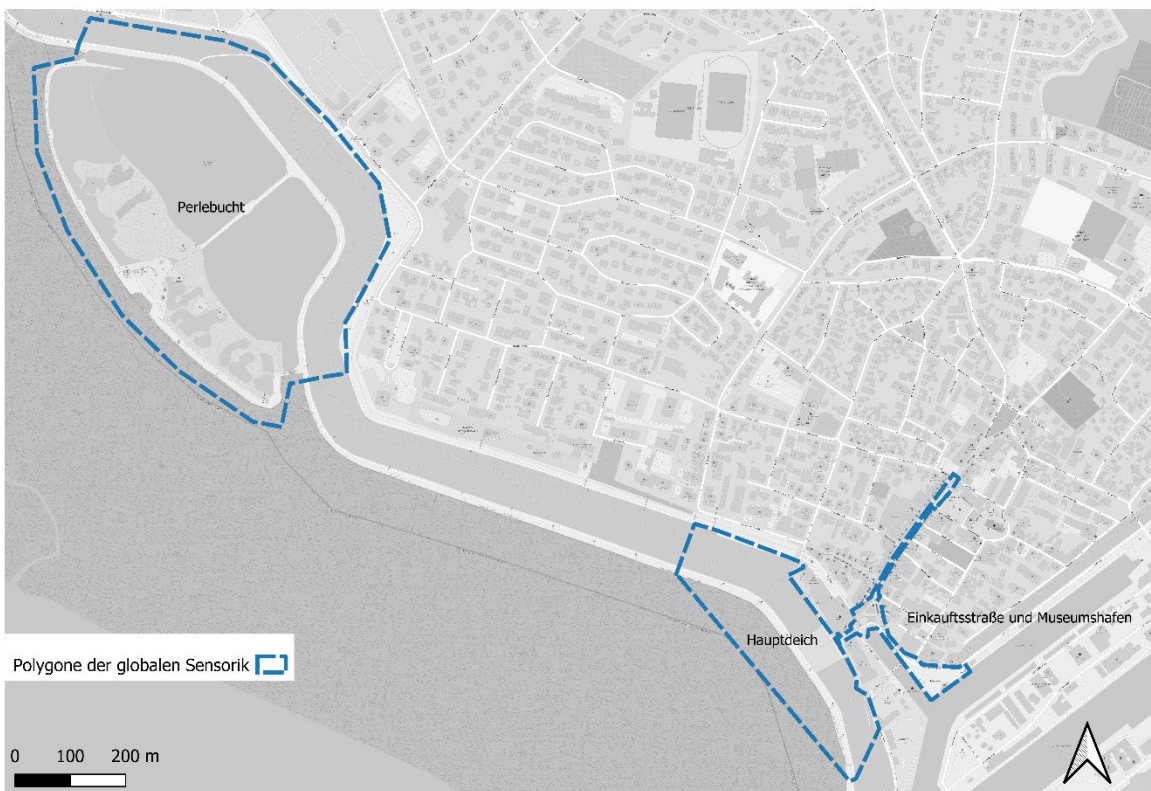




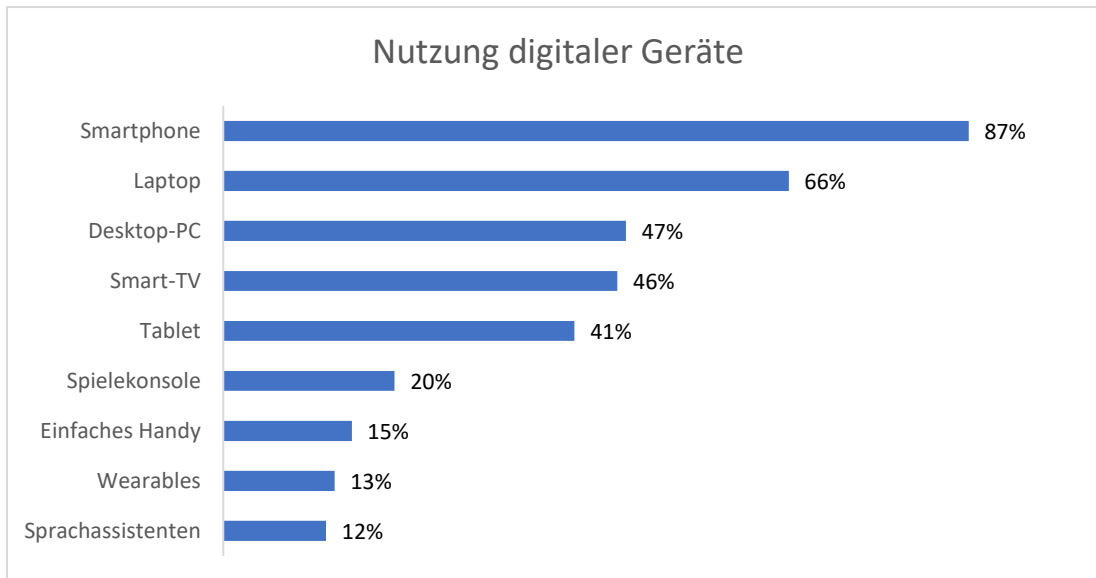






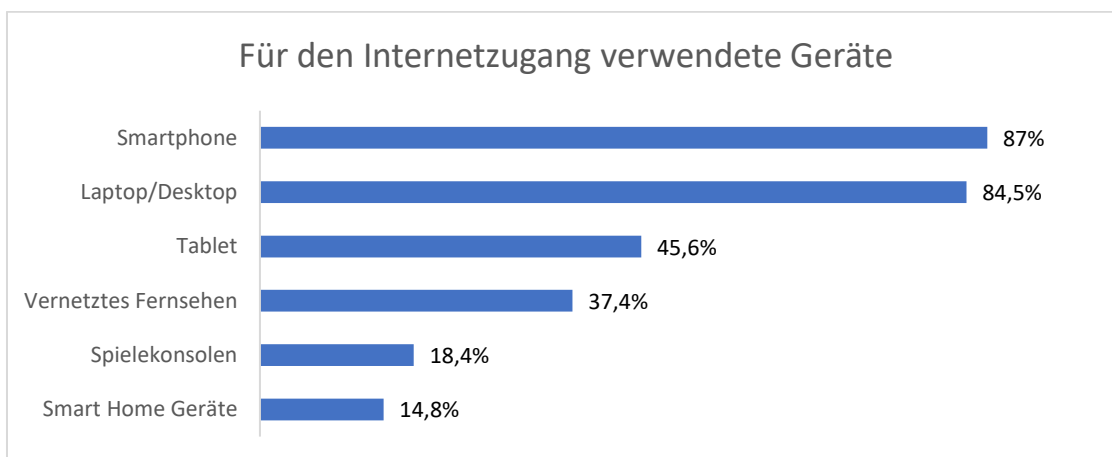


Anhang 3: Ergänzende Recherchen für das Relevanzbewertungsverfahren



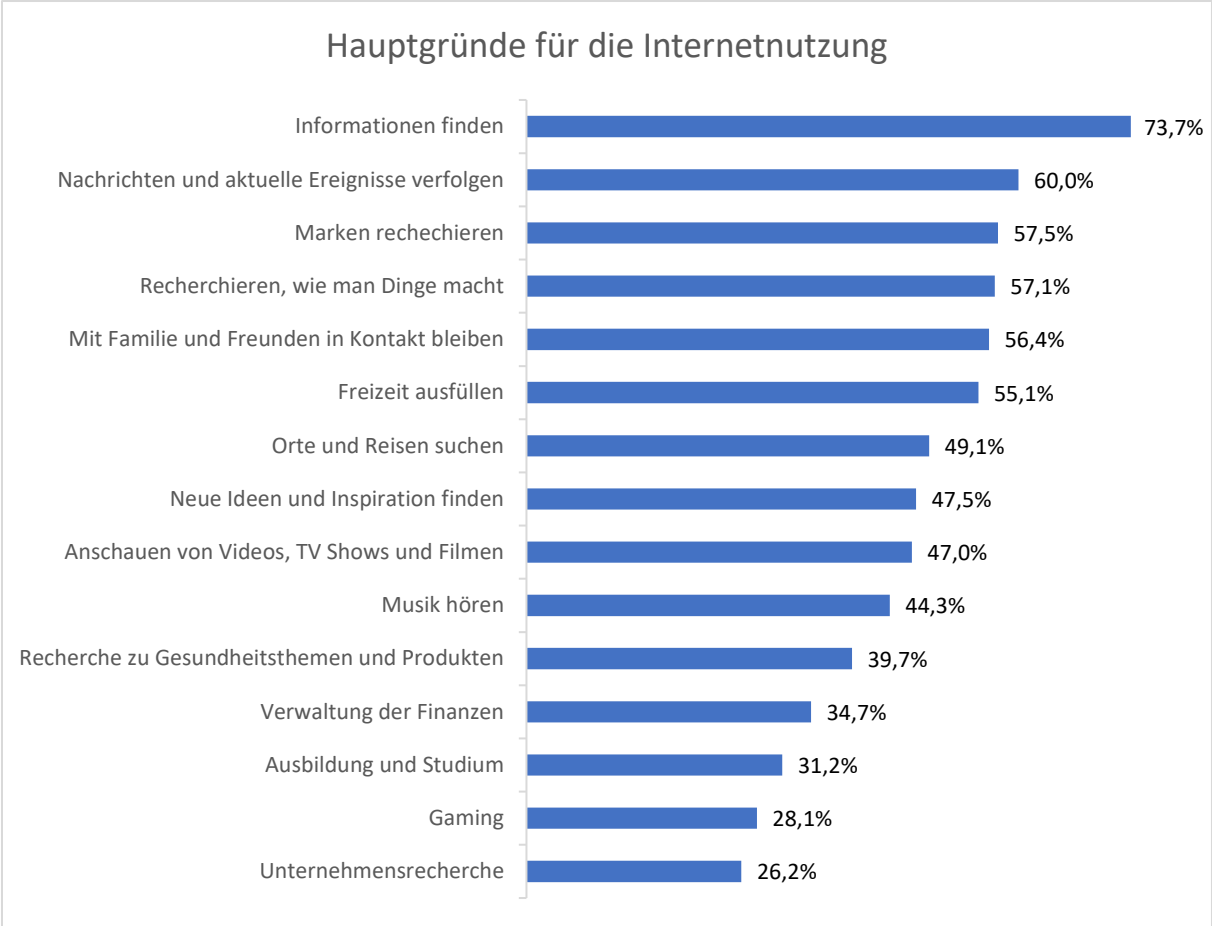
Nutzung digitaler Geräte

Quelle: Eigene Darstellung basierend auf Initiative D21 e. V., 2022



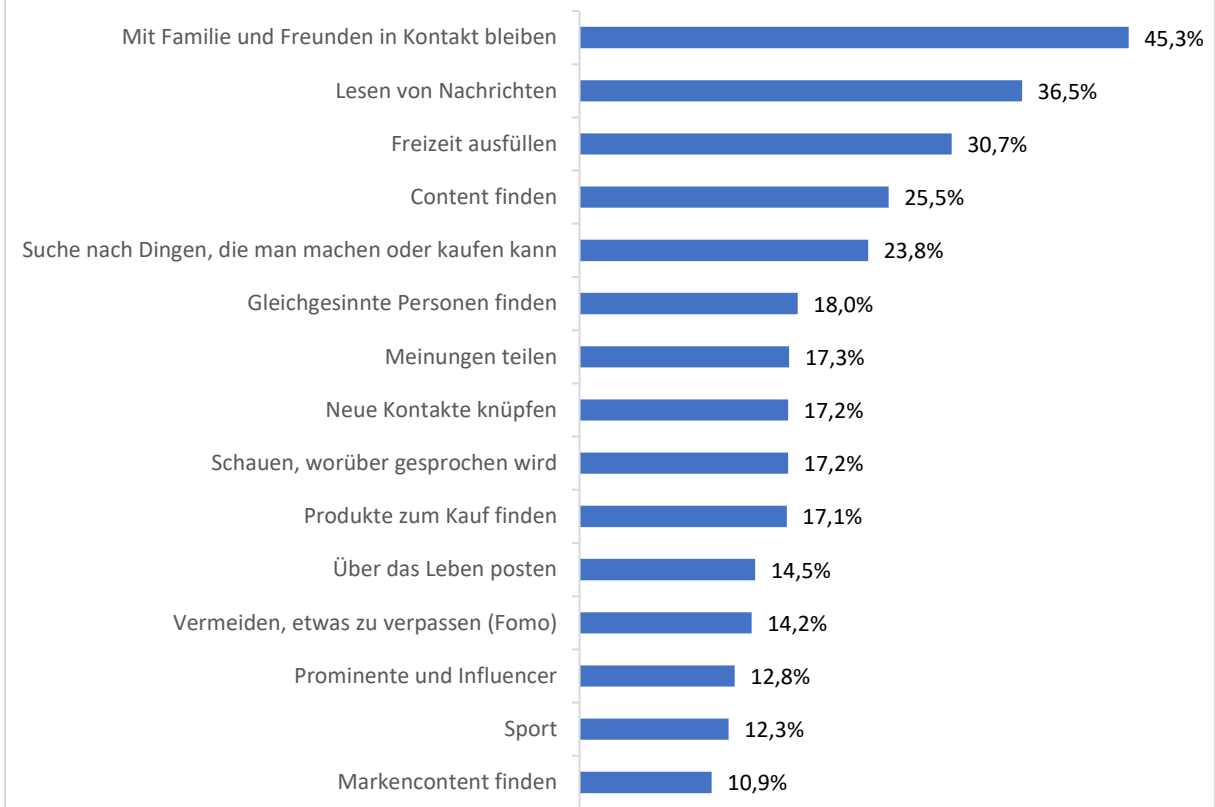
Für den Internetzugang verwendete Geräte

Quelle: Eigene Darstellung basierend auf DataReportal, 2022



Hauptgründe für die Internetnutzung
Quelle: Eigene Darstellung basierend auf DataReportal, 2022

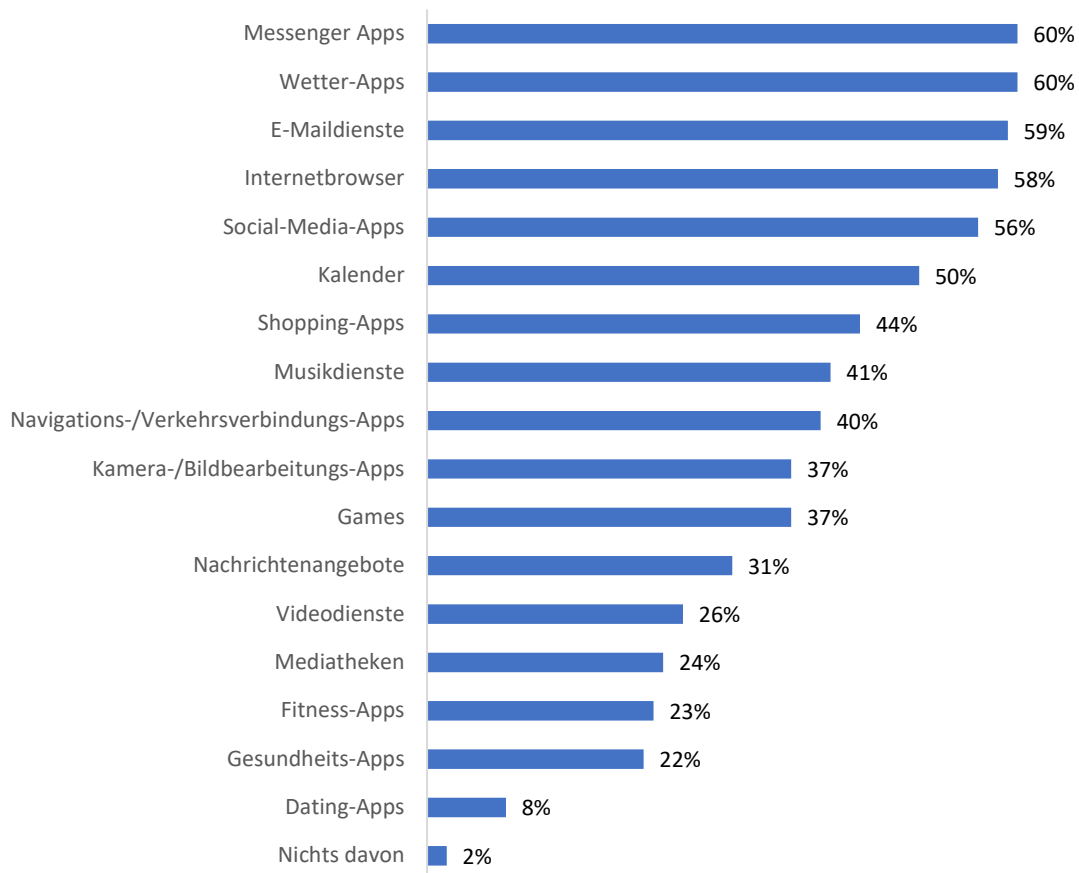
Hauptgründe für die Nutzung von Social Media



Hauptgründe für die Nutzung von Social Media

Quelle: Eigene Darstellung basierend auf DataReportal, 2022

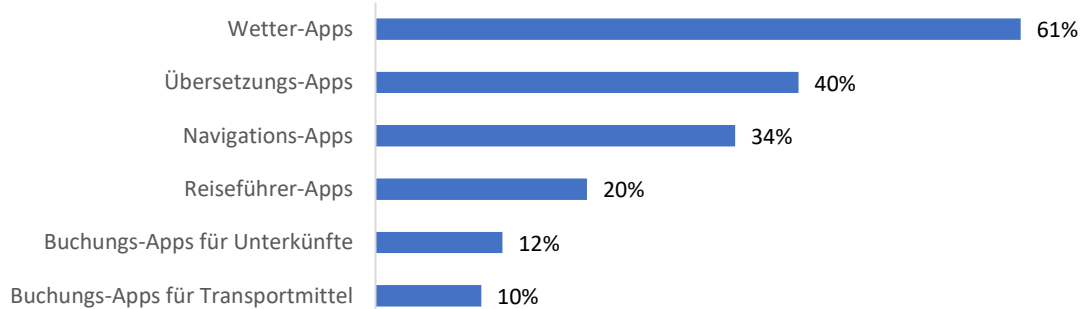
Beliebte Arten von Smartphone-Apps in Deutschland im Jahr 2022



Beliebte Arten von Smartphone-Apps in Deutschland im Jahr 2022

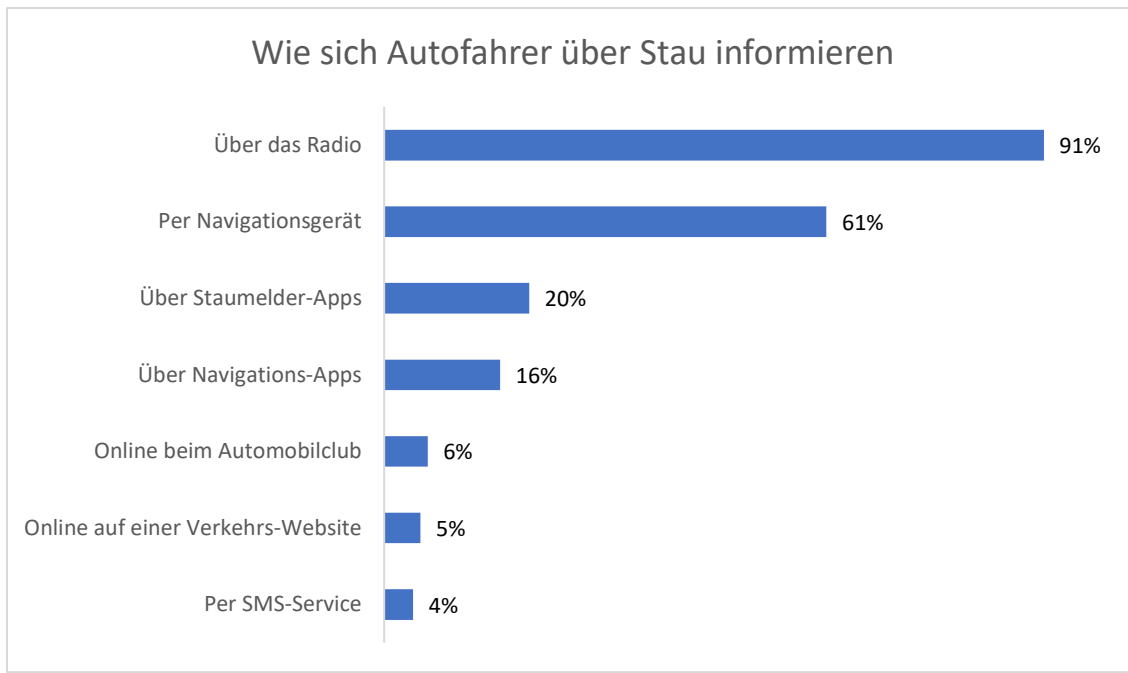
Quelle: Eigene Darstellung basierend auf Statista Consumer Insights, 2023

Nutzung von Apps im Urlaub 2020



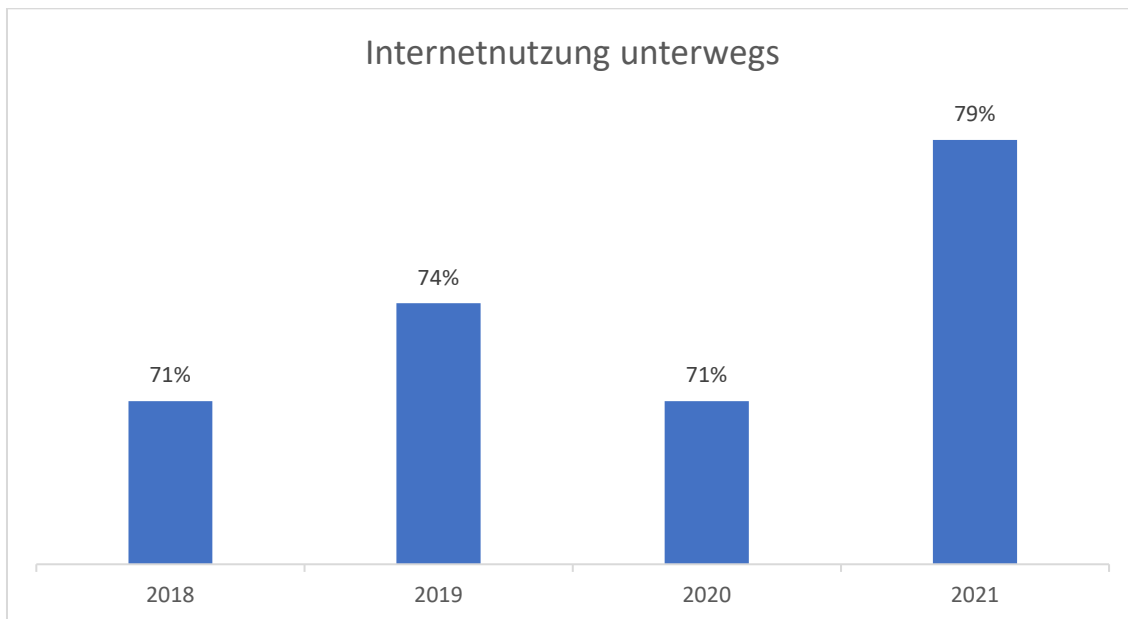
Nutzung von Apps im Urlaub

Quelle: Eigene Darstellung basierend auf Bitkom e.V., 2020b



Wie sich Autofahrer*innen über Stau informieren

Quelle: Eigene Darstellung basierend auf Bitkom Research, 2020



Internetnutzung unterwegs – zumindest gelegentlich

Quelle: Eigene Darstellung basierend auf ARD/ZDF-Forschungskommission, 2021

Anhang 4: Datenbankensysteme/Technische Dienstleister

Zusammenfassung ausgewählter Systeme

Datenbankensystem/ Technische Dienstleister	Aufgabenbereiche	Referenzen/Produkte/ Auspielungskanäle
Datacycle.info	<ul style="list-style-type: none"> - Datenmanagement-System - Zusammenfassung unterschiedlicher Datenbestände in einem zentralen System 	<ul style="list-style-type: none"> - Open!dataCycle.cloud: zentraler Daten-Hub für alle Arten von frei verfügbaren touristischen Daten im DACH-Raum
Destination.one	<ul style="list-style-type: none"> - Abgestimmte Soft- und Hardware-Lösungen für Destinationen, Touristinformationen, Gastgebern und Leistungsträgern für eine individuelle Gäste-Information auf allen Touchpoints - Hardware zur Messung und Ausspielung von Daten 	<ul style="list-style-type: none"> - Open.destination.one: Open Data Plattform für die DACH-Region - Destination.data: Multi-Channel-Datenbank - Destination.cms: Content-Management-System mit TYPO3 - PWA: Smarte Datenbereitstellung - Destination.city: Webseiten für Destinationen
Discover swiss	<ul style="list-style-type: none"> - Betrieb einer offenen digitalen Backend Service-Plattform für den Schweizer Tourismus - alle touristischen Service-Lieferanten können ihre Informationen, Angebote und Services den Service-Nutzer*innen wie Destinationen oder Touring-anbietern zur Verfügung stellen 	<ul style="list-style-type: none"> - discover swiss Plattform
Feratel.de	<ul style="list-style-type: none"> - Entwickler und Anbieter touristischer Informationssysteme - Aufteilung der Geschäftstätigkeit in: Telekommunikation, Informations- und Reservierungssysteme und Medien 	<p>Lösungen in der Praxis u.a. in den Bereichen:</p> <ul style="list-style-type: none"> - Destinationsmanagement - Gästekarte - Digitaler Urlaubsbegleiter - digitale Gästeregistrierung "myVisitPass"
General-solutions.at	<ul style="list-style-type: none"> - maßgeschneiderte Softwarelösungen im Bereich Tourismus-, Destination- sowie Krisen- und Einsatzmanagement - Unterstützung im digitalen Organisationsmanagement 	<ul style="list-style-type: none"> - Contwise Maps: interaktive Karte zur übersichtlichen Darstellung touristischer Angebote - Contwise Infra: digitale Erfassung, Verwaltung und Erhaltung von Infrastruktur - Contwise Apps: individuelle Softwarelösungen
Infomax-it.de	<ul style="list-style-type: none"> - Entwicklung von digitalen Lösungen und Services für Tourismusmarken, Destinationen und Mobilitätsunternehmen im deutschsprachigen Raum 	<ul style="list-style-type: none"> - Imx.Platform: Datenmanagementsystem, offizielle Landesplattform von Schleswig-Hol-

	<ul style="list-style-type: none"> - Frontend- oder Backend-Programmierung, Anbindung von Content Management Systemen oder individueller Lösungen, Systemintegration und Schnittstellenprogrammierung 	<p>stein, Mecklenburg-Vorpommern, Hamburg, Tirol, Niederösterreich, Luxemburg</p> <ul style="list-style-type: none"> - Destinations-Datenhub - Whitelabel-Eventportal
Join.de	<ul style="list-style-type: none"> - Unterstützung bei Digitalisierungsprojekten - konzeptionelle Beratung, kundenindividuelle Anwendungsentwicklung, Implementierung und Betrieb der Anwendungen 	<ul style="list-style-type: none"> - Entwicklung und technische Realisierung der Tourismusdatenbank ThüCAT (Thüringer Content Architektur Tourismus) - ThüCAT: digitale Präsentationsmöglichkeiten für Unternehmen, Regionen, Städte und Freizeiteinrichtungen - Gäste erhalten aktuelle Informationen (z.B. Übernachtungsmöglichkeiten, Sehenswürdigkeiten, Reiseinspirationen, Öffnungszeiten)
Land-in-sicht.de	<ul style="list-style-type: none"> - Touristisches Datenmanagement - Tägliche Verwaltung, Vernetzung und Veröffentlichung von tausenden Datensätze 	<ul style="list-style-type: none"> - Tool: mein.toubiz <p>Im Einsatz als:</p> <ul style="list-style-type: none"> - Landessystem zur Digitalisierung der touristischen Infrastruktur in BW - Zentrale Datenbank für die Digitalisierungsstrategie vom Schwarzwald - Grundlage des Datenmanagements beim Wanderportal "Wanderbares Deutschland"
Maqnify	<ul style="list-style-type: none"> - Entwicklung von digitalen Stadtrundgängen - Entwicklung von Besucherleitsystemen 	<ul style="list-style-type: none"> - Maqnify-Plattform
Outdooractive.com	<ul style="list-style-type: none"> - Outdoor-Plattform fürs Web und als App - Tourenplaner und Reiseführer - Datenmanagement, Marketing, Web- und App-Technologie 	<ul style="list-style-type: none"> - Website und App von outdooractive
Pixelpoint.at	<ul style="list-style-type: none"> - Entwicklung: Web- und App-Development, Software- & Datenmanagement, eCommerce, AR & VR - Design, Marketing, Performance Analyse, Consulting - Hosting & Wartung 	<ul style="list-style-type: none"> - Webdevelopment & -design für Destinationen und Unternehmen wie Donau Niederösterreich, Reiseland Brandenburg, Burgenland Tourismus, feratel - Entwicklung und Konzeptionierung der BayernCloud Tourismus, in Zusammenarbeit mit feratel
Ttg.at	<ul style="list-style-type: none"> - IT-Infrastrukturanbieter - Innovative, digitale Tourismuslösungen 	<ul style="list-style-type: none"> - TOURDATA: Tourismusdatenbank für Österreich, multimediales Datenmanagementsystem

		<p>zur Erfassung, Verarbeitung und Nutzung touristischer Daten</p> <ul style="list-style-type: none">- TOURismus CRM: maßgeschneidertes Customer-Relationship-Management Tool- TOURismus Intelligence: Dashboard mit Webstatistiken- Upperguide: online Reiseführer, für Oberösterreich, liefert aktuelle Infos aus der Umgebung
--	--	--