Geokodierung von Befragungsergebnissen im Rahmen von „Mobilität in Deutschland“

Präsentation am 27. Juni 2002

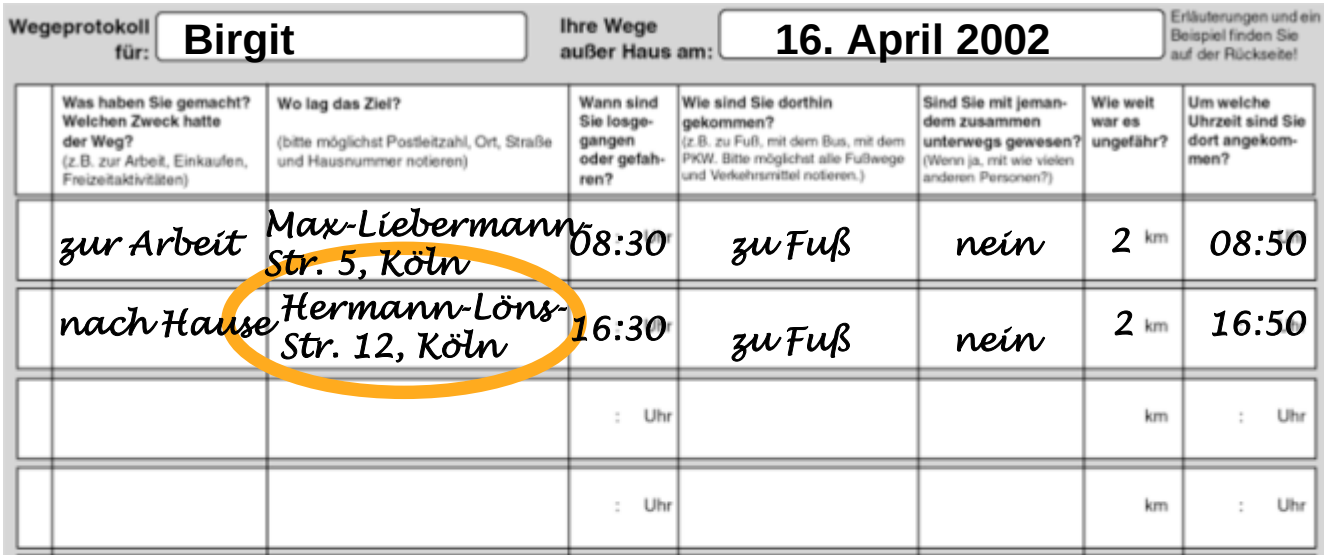
Inhalte der Präsentation

- Was ist eine Geokodierung?
- Qualitätsstufen der Geokoordinaten
- mögliche Zusatzdaten auf Basis von Geokoordinaten
- Ergebnisse der Machbarkeitsstudie „Geokodierung“
- Geokodierung: Leistungen und Kosten

Basisdaten zur Geokodierung:

Erhebung der Start- und Zieladressen im telefonischen Interview

Bei „Mobilität in Deutschland“ werden die Zielpersonen nach der genauen Start- und Zieladresse eines jeden Weges gefragt:



Wegeprotokoll für: **Birgit** — Ihre Wege außer Haus am: **16. April 2002** — Erläuterungen und ein Beispiel finden Sie auf der Rückseite!

	Was haben Sie gemacht? Welchen Zweck hatte der Weg? (z.B. zur Arbeit, Einkaufen, Freizeitaktivitäten)	Wo lag das Ziel? (bitte möglichst Postleitzahl, Ort, Straße und Hausnummer notieren)	Wann sind Sie losgegangen oder gefahren?	Wie sind Sie dorthin gekommen? (z.B. zu Fuß, mit dem Bus, mit dem PKW. Bitte möglichst alle Fußwege und Verkehrsmittel notieren.)	Sind Sie mit jemandem zusammen unterwegs gewesen? (Wenn ja, mit wie vielen anderen Personen?)	Wie weit war es ungefähr?	Um welche Uhrzeit sind Sie dort angekommen?
	zur Arbeit	Max-Liebermann-Str. 5, Köln	08:30 Uhr	zu Fuß	nein	2 km	08:50 Uhr
	nach Hause	Hermann-Löns-Str. 12, Köln	16:30 Uhr	zu Fuß	nein	2 km	16:50 Uhr
			: Uhr			km	: Uhr
			: Uhr			km	: Uhr

In diesem Beispiel ist die Zielperson am Stichtag um 16:30 von ihrem Arbeitsplatz in der Max-Liebermann-Straße nach Hause in die Hermann-Löns-Straße gegangen. Diese Adressen werden aufgenommen und gespeichert.

Qualitätsstufen der Geokoordinaten

Gewinnung von Koordinaten für Start- und Zieladressen

Die im Interview gewonnenen Adressdaten können - je nach ihrer Qualität - einer geographischen Koordinate* zugeordnet werden.

Insgesamt existieren fünf unterschiedliche Qualitätsstufen für die Geokodierung:

- hausgenau
- straßenabschnittsgenau
- stimmbezirksgenau
- postleitzahlgenau
- nicht geokodierbar

Je nach Genauigkeit der Geokodierung können Start- und Zielpunkt dieses Weges könnten auf einer Karte exakt dargestellt werden.

* Lambert-konforme Projektion, Bessel-Ellipsoid

Übersicht:
Genauigkeit der Qualitätsstufen (Mittelwerte)

postleitzahlgenau:	gesamt:	1.140 m
stimmbezirksgenau:	gesamt:	430 m
	über 20.000 Einwohner	235 m
	über 50.000 Einwohner	180 m
	über 100.000 Einwohner	155 m
	über 500.000 Einwohner	130 m
straßenabschnittsgenau:	gesamt:	70 m
	über 20.000 Einwohner	61 m
	über 50.000 Einwohner	57 m
	über 100.000 Einwohner	54 m
	über 500.000 Einwohner	52 m
hausnummerngenau:	gesamt:	30 m
	über 20.000 Einwohner	30 m
	über 50.000 Einwohner	30 m
	über 100.000 Einwohner	30 m
	über 500.000 Einwohner	28 m

Übersicht:
Maximal erreichbare Kodierqualität

Qualitätsstufe	Prozent
hausgenau	21,9%
straßenabschnittsgenau	56,4%
stimmbezirksgenau	15,2%
postleitzahlgenau	6,4%
nicht geokodierbar	0,0%

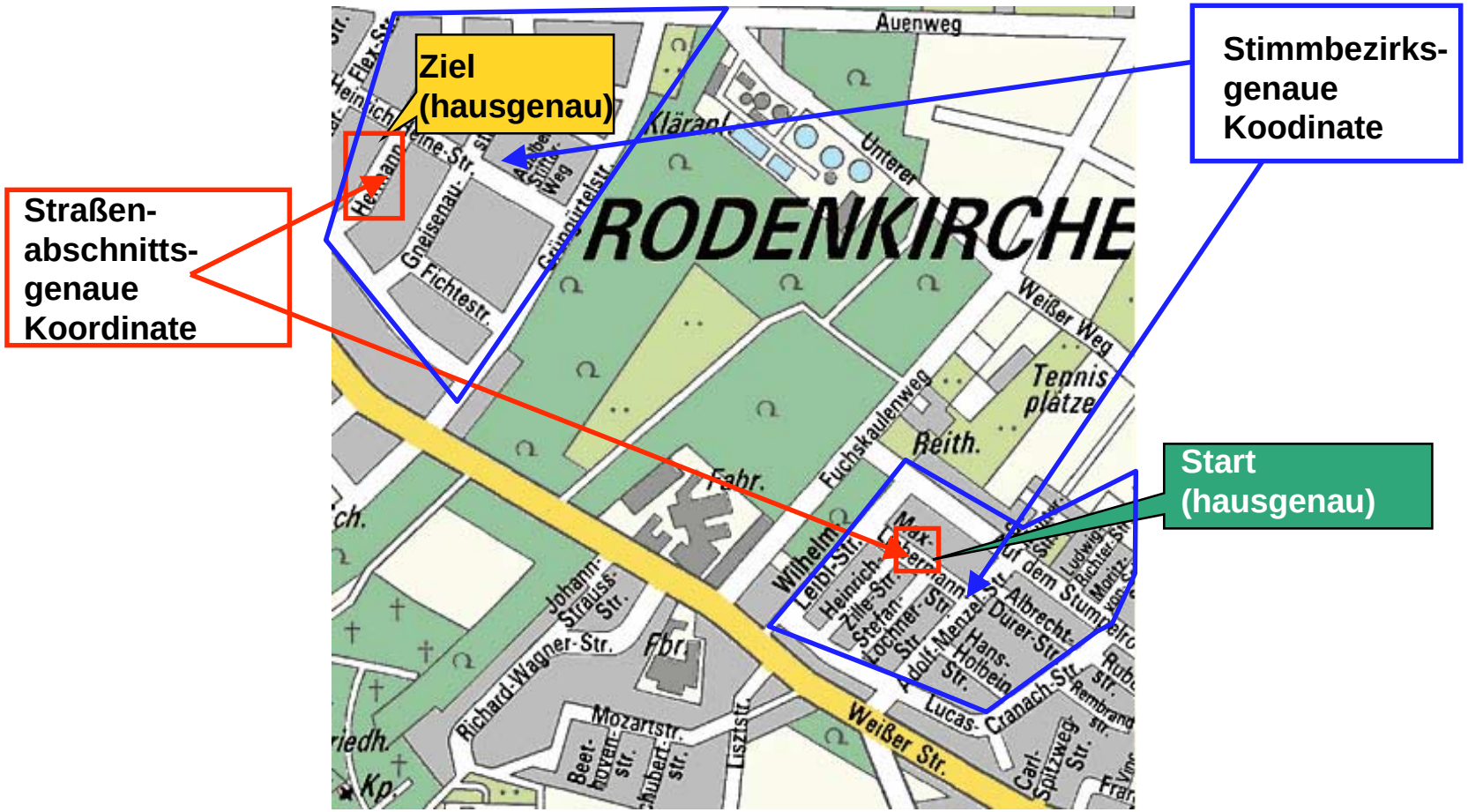
gegenwärtige bundesweite Verteilung

entspricht

maximal erreichbarer Kodierqualität

Beispiel:

Genauigkeit ausgewählter Qualitätsstufen



Zusatzdaten auf Basis einer Geokodierung

Beispiel 1: mögliche Zusatzdaten aus einer Geokodierung *LOCAL-Wohnumfeld*

Diesen geokodierten Adressen können Daten zum LOCAL-Wohnumfeld zugespielt werden, die gegenwärtig geokodiert vorliegen.

Insgesamt existieren 53 unterschiedliche LOCAL-Typen, die das Wohnumfeld stimmbezirksscharf beschreiben und so weitere Analysen ermöglichen.

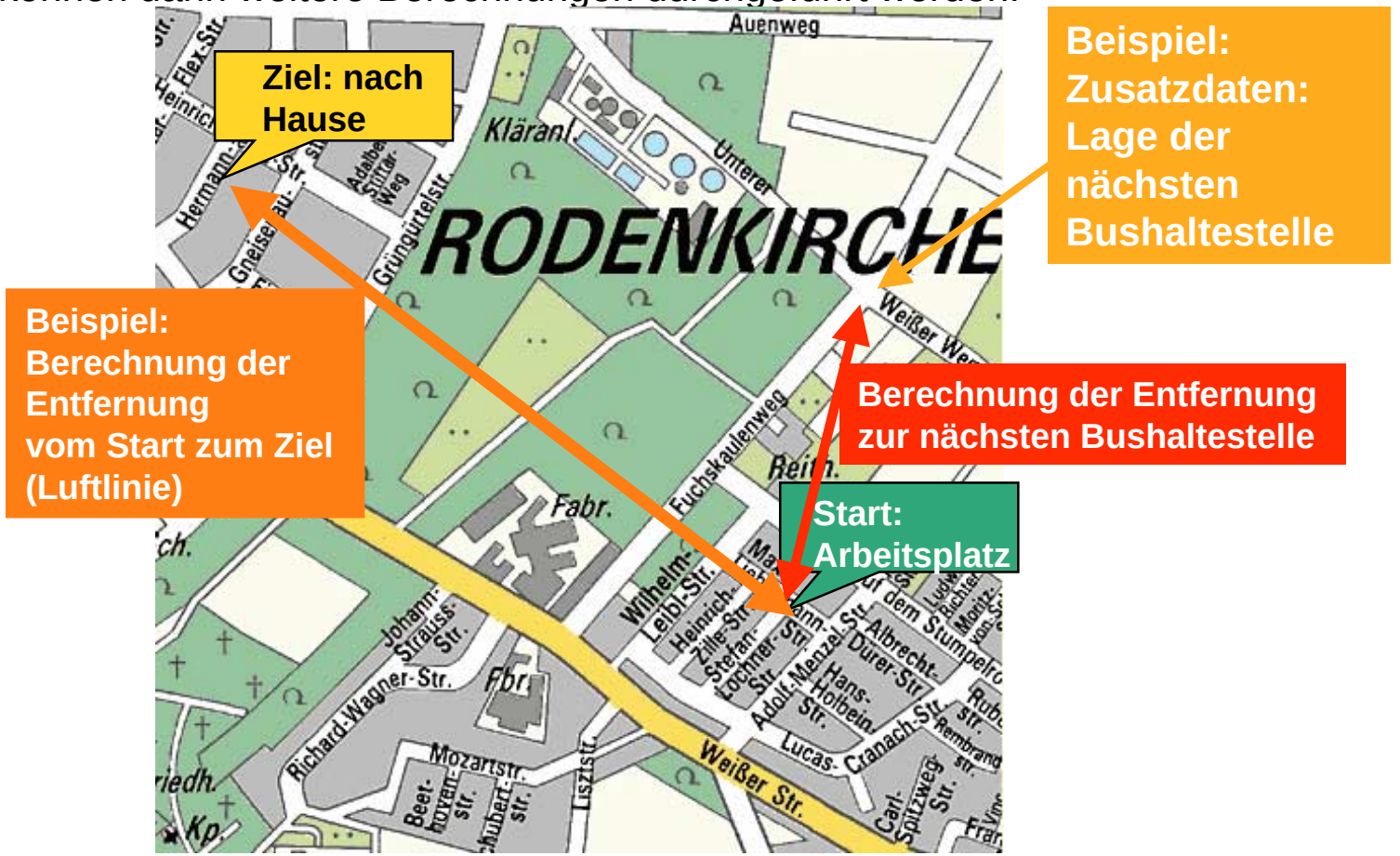
Zusatzdaten: LOCAL-Wohnumfeld der Zielperson

Beispiel: *LOCAL-Wohnumfeldtyp*

	Aggr. Stimmbezirks-Typologie	Beschreibung Aggregierte Stimmbezirks-Typologie	Stimmbezirks-Typologie	Beschreibung Stimmbezirks-Typologie
M1: Ältere Mehrfamilienhäuser				
	M2	Ältere bessere Mehrfamilienhausgebiete	M11	**Gehobene Mietwohnungen** Überwiegend 1945 bis 1970 gebaute Mehrfamilienhäuser von mittlerem und gehobenem Status; eher ältere Haushalte, überdurchschnittlicher Anteil von Angestellten und Beamten in mittlerer Position, durchschnittliche Einkommen
	M2	Ältere bessere Mehrfamilienhausgebiete	M12	**Gute 20-30er Jahre Mietwohnungen** 1918 bis 1938 gebaute Mehrfamilienhäuser von mittlerem und gehobenem Status; gemischte Haushalts- und Altersstruktur, eher überdurchschnittlicher Anteil von Angestellten und Beamten, knapp überdurchschnittliche Einkommen
	M5	Ältere sehr einfache Mehrfamilienhausgebiete	M13	**Gemischter Mietwohnungs-Altbau** Vor 1918 erbaute Mehrfamilienhäuser; teilweise modernisiert, durchschnittlicher Erhaltungszustand; überdurchschnittlicher Anteil älterer und 20- bis 40jähriger Personen; häufig mit Geschäften, Kleingewerbe/mittlerem Gewerbe, Gaststätten durchsetzt; bevorzugtes Studentenwohngebiet (soweit Hochschulstandort)
	M4	Ältere einfache Mehrfamilienhausgebiete	M14	**Ältere, einfache Mehrfamilienhäuser** 1945 bis 1970 erbaute Mehrfamilienhäuser; kleinere ältere Haushalte, durchschnittlicher Ausländeranteil, überdurchschnittlicher Arbeiteranteil, unterdurchschnittliche Einkommen
	M4	Ältere einfache Mehrfamilienhausgebiete	M15	**Einfache 20-30er Jahre Mietwohnungen** 1918 bis 1945 gebaute Mehrfamilienhäuser; gemischte Haushalts- und Altersstruktur, überdurchschnittlicher Ausländeranteil, überdurchschnittlicher Arbeiteranteil, unterdurchschnittliche Einkommen
	M5	Ältere sehr einfache Mehrfamilienhausgebiete	M16	**Sehr einfacher Altbaubestand** Vor 1918 erbaute Mehrfamilienhäuser, sehr einfacher Erhaltungszustand; ältere Haushalte, überdurchschnittlicher Ausländeranteil, weit überdurchschnittlicher Arbeiteranteil, teilweise mit Geschäften und Kleingewerbe durchsetzt, häufig Problemgebiete, unterdurchschnittliche Einkommen

Beispiel 2: mögliche Zusatzdaten aus einer Geokodierung
Berechnung von Entfernungen

Diesen geokodierten Adressen könnten in einem weiteren Schritt Daten aus anderen Quellen zugespielt werden, die gegenwärtig geokodiert vorliegen. Auf dieser Basis können dann weitere Berechnungen durchgeführt werden.



Beispiel 3: mögliche Zusatzdaten aus einer Geokodierung

Wegeanalyse

Mithilfe eines Routenplaners kann der schnellste oder der kürzeste Weg vom Start zum Ziel berechnet und darauf aufbauend bspw. die benutzten Straßentypen identifiziert werden.

Ergebnisse
der Machbarkeitsstudie

Machbarkeitsstudie Geokodierung

Im Rahmen der Machbarkeitsstudie „Geokodierung“ wurden im Auftrag der Bundesanstalt für Straßenwesen (BASt) unter anderem folgende Aspekte einer Geokodierung untersucht:

- **Qualität der erhaltenen Geokoordinaten**

- **Nutzen und Aufwand einer manuellen Nachbereitung**
  (für nicht geokodierbare oder nicht mindestens stimmbezirksgenauer Adressen)

Ergebnisse der automatischen Geokodierung (Machbarkeitsstudie): maximal jede zweite Adresse und jeder dritte Weg exakt bestimmbar

	CATI *			schriftlich **	
Geocodierung Adressen*)	absolut	Prozent		absolut	Prozent
gesamt	19004	100,0%		18500	100,0%
hausgenau	**4089**	**21,5%**		**3264**	**17,6%**
straßenabschnittsgenau	**6374**	**33,5%**		**4348**	**23,5%**
stimmbezirksgenau	1816	9,6%		1307	7,1%
Postleitzahlgenau	**3038**	**16,0%**		**1740**	**9,4%**
Gemeindegenau	**1954**	**10,3%**		**51**	**0,3%**
nicht geocodiert	**1733**	**9,1%**		**7790**	**42,1%**

	CATI			schriftlich	
Geocodierung Wege	absolut	Prozent		absolut	Prozent
Gesamt	9502	100,0%		9250	100,0%
Start und Ziel mind. Stimmbezirksgenau	**3652**	**38,4%**		**2625**	**28,4%**
- darunter Start und Ziel hausgenau	386	4,1%		441	4,8%
- darunter Start und Ziel mind. Straßenabschnittsgenau	3139	33,0%		2105	22,8%
Start und/oder Ziel weniger als Stimmbezirksgenau	**5850**	**61,6%**		**6625**	**71,6%**

*) maschinelle Geocodierung mit manueller Korrektur/Ergänzung der PLZ in 3125 Fällen

**) maschinelle Geocodierung

Selektivitätsbetrachtung der zugeordneten Wege

Qualität der Adressen nach Wegezweck
Beispiel: telefonische Befragung

		Qualität der Zieladresse		
Zweck / Ziel des Weges		weniger als stimmbezirks-genau	mindestens stimmbezirks-genau	Total
Erreichen des Arbeitsplatzes	n	294	498	792
	%	**37,1%**	**62,9%**	**100,0%**
dienstlich oder geschäftlich	n	96	53	149
	%	**64,4%**	**35,6%**	**100,0%**
Erreichen der Ausbildungsstätte oder Schule	n	93	101	194
	%	**47,9%**	**52,1%**	**100,0%**
Einkaufsfahrt	n	595	280	875
	%	**68,0%**	**32,0%**	**100,0%**
private Erledigungen	n	345	256	601
	%	**57,4%**	**42,6%**	**100,0%**
Holen oder Bringen anderer Personen	n	194	155	349
	%	**55,6%**	**44,4%**	**100,0%**
Besuch von Verwandten oder Bekannten	n	178	170	348
	%	**51,1%**	**48,9%**	**100,0%**
Restaurantaufenthalt oder Besuch einer Gaststätte	n	85	40	125
	%	**68,0%**	**32,0%**	**100,0%**
Kirche, Kultur, Bildung, Sport	n	196	122	318
	%	**61,6%**	**38,4%**	**100,0%**
Spaziergang, Ausflug	n	324	61	385
	%	**84,2%**	**15,8%**	**100,0%**
nach Hause	n	78	2561	2639
	%	**3,0%**	**97,0%**	**100,0%**
andere Aktivität	n	305	180	485
	%	**62,9%**	**37,1%**	**100,0%**
Rückweg vom vorherigen Weg	n	165	1269	1434
	%	**11,5%**	**88,5%**	**100,0%**
(Vor-)Schule	n	127	98	225
	%	**56,4%**	**43,6%**	**100,0%**
Kindertagesstätte oder Kindergarten	n	57	45	102
	%	**55,9%**	**44,1%**	**100,0%**

Ergebnis:

manuelle Nachbearbeitung **erforderlich**:

- wg. der **Repräsentativität** der Wege-Stichprobe
- wg. **Vergrößerung** der Analysebasis

Geokodierung im Rahmen von „Mobilität in Deutschland“: Leistungen und Kostenschätzung

Geokodierung im Rahmen von „Mobilität in Deutschland“

Leistungen (Basisgeokodierung)

- Automatische Geokodierung aller Wegeadressen (Lambert-konforme Projektion, Bessel-Ellipsoid)
- Durchführung einer Selektivitätsanalyse
- Manuelle Nachrecherche zur „Komplettierung“ unterrepräsentierter Wege
- Zuspielen von LOCAL-Wohnumfeld Informationen an alle Heimadressen
- Auslieferung eines Datensatzes mit den Zusatz-informationen:
    - Koordinate der Start- und Zieladresse (aus Datenschutzgründen maximal straßenabschnittsgenau)
    - Gemeindekennziffer (GKZ)
    - Stimmbezirkskennung

Geokodierung im Rahmen von „Mobilität in Deutschland“

Leistungen (Zusatzmodule)

Neben der Basisgeokodierung können weitere Daten oder Analysen individuell vereinbart werden. Diese Zusatzmodule werden auf die individuellen Erfordernisse der Auftraggeber zugeschnitten und gesondert kalkuliert.

Beispiele für Zusatzmodule:

- Wegeanalyse (Berechnung des längsten/kürzesten Wegs vom Start zum Ziel; Fahrleistung nach Straßentypen)
- Bestimmung der Fahrleistung nach Ortslagen (innerorts, außerorts)
- Zuspielen von weiteren geokodierten Datenbeständen (bspw. GIS-Daten) an den Datensatz
- Analysen und Berechnungen auf Basis der nicht-öffentlichen hausgenauen Geokoordinaten