

2 Stand der Forschung

Die Erhebung des Standes der Forschung, Entwicklung und Technik raumanalytischer und geovisualer Methoden und Techniken bei Immobilienportalen ist Inhalt des zweiten Kapitels. Diese Bestandsaufnahme zeigt die Breite und die Vernetzungen eines möglichen neuen Forschungsfeldes, das bisher allein durch die geovisuellen und geoinformatischen Methoden der Portalbetreiber abgesteckt wird und zu dessen forschungsbasierter Erweiterung diese Arbeit beitragen möchte. Zur Bestandsaufnahme des State of the Art wird vom Autor eine technologische Recherche zur Analyse der raumanalytischen und geovisuellen Methoden und Techniken von Immobilienportalen sowie eine diesbezügliche Literaturrecherche durchgeführt. Ergebnis dieser Recherchen ist die Darstellung des Status quo der vorherrschenden raumanalytischen und geovisuellen Methoden und Techniken bei Immobilienportalen.

2.1 Immobilienportale

Seit Immobilienportale die Printmedien als primäre Anlaufstelle zur Immobiliensuche verdrängt haben, ist für die Mehrzahl der Wohnungsuchenden eine Suche ohne die Nutzung eines Immobilienportals undenkbar geworden. Das Maklergeschäft wird zunehmend abhängig von Immobilienportalen, deshalb pflegen Makler ihre Angebotsdaten in die Portale ein. Immobilienportale geben einen Überblick über die einzelnen angebotenen Objekte, stellen Immobilienmerkmale in Form von Portfolios dar, visualisieren die Objektlage in Kartenform und analysieren die Angebotspreise. Nutzern werden die Analyseergebnisse in Kartenform bereitgestellt.

Die aktuelle Relevanz von Immobilienportalen lässt sich anhand zahlreicher Statistiken belegen. Im Jahr 2011 publizierte das Hamburger Marktforschungsinstitut *Fittkau und Maaß* Ergebnisse einer im November 2010 durchgeführten Studie von 15.300 Haushalten, die innerhalb eines Zeitraums von zwölf Monaten umgezogen sind. 70.9 % dieser Haushalte gaben an, ihre Immobilie online gesucht zu haben (*Rebhan*, 2011). Wird diese Erhebung der in Kapitel 1.1 vom Autor geschätzten Zahl von jährlich 5,6 Millionen Immobiliengeboten gegenübergestellt und werden zusätzlich die Besucherzahlenauswertungen von *Alexa.com*³ betrachtet, unterstreicht dies die Relevanz von Immobilienportalen

3 *Alexa.com* ist eine von Amazon.com betriebene Webseite, die Nutzerzahlen von Webseiten analysiert. Das deutsche Immobilienportal *ImmobilienScout 24* liegt bei *Alexa* seit Jahren in

für die Immobiliensuche und ganz allgemein für Immobilienmärkte sehr deutlich.

Immobilienportale speichern alle eingestellten Angebotsdaten und haben sich zu großen Datensammlern entwickelt. Allein das 1999 gegründete deutsche Immobilienportal *ImmobilienScout 24* speichert derzeit 1,5 Mio. Immobilienangebote in seiner Datenbank (*Immobilienportale*, 2012). Die Tatsache, dass alle aufgenommenen Daten geokodiert sind, machen Immobilienportale zu einer großen Quelle für raumanalytische Ansätze der Immobilienmarktanalyse. Die enorme verfügbare Datenfülle erlaubt es den Betreibern von Immobilienportalen, umfangreiche, auf Methoden der Statistik basierende Marktanalysen durchzuführen. Maklern und Vertreibern von Immobilien werden diverse kostenpflichtige Reporting-Dienste zur Analyse der Nachfrage ihres Angebots offeriert. Neben Reporting-Diensten wurde zudem das Thema Transparenz als zusätzliches Geschäftsfeld erkannt. So werden die vorhandenen Datenbestände genutzt, um zusätzliche – nicht nur monetäre – Mehrwerte über die reine Immobiliensuche hinaus zu generieren. Beispielsweise führten die *Gesellschaft für Konsumforschung (GfK)* bzw. das *Forschungsinstitut Empirica* diverse Marktanalysen basierend auf Daten von Portalbetreibern durch (*Enderle*, 2009).

2.2 Literatur

Trotz der unbestreitbaren Relevanz von Immobilienportalen gibt es bisher kaum Studien hierüber. Es existieren in der Literatur allenfalls Untersuchungen, die Teilaspekte des Forschungsgegenstandes behandeln. Ein Artikel von *Borchert* (2006) beispielsweise beschäftigt sich mit der Anwendung von Geoinformationssystemen (GIS) für die Immobilienwirtschaft. Im Jahr 2009 analysierten *Schernthanner* und *Tyrallova* die Rolle freier quelloffener Geoinformationssysteme (FOS-GIS) als Instrument der Wohnungsmarktanalyse. Aus den Feldern der Wirtschaftswissenschaft und der Immobilienmarktforschung gibt es unterschiedliche Beiträge zur Thematik „Immobilienportale“. *Enderle* (2009) analysiert die durch Immobilienportale veränderte Markttransparenz aus dem Blickwinkel der Wirtschaftswissenschaften. *Rohmert* und *Böhm* (2013) gaben eine Aufsatzsammlung zu den Einflüssen des Internets auf die Immobilienwirtschaft heraus.

den Top 50 der meistbesuchten Internetseiten Deutschlands (Platz 21 am 20.11.2014). Das US-amerikanische Portal *Zillow* liegt auf Platz 37 (Stand 07.05.2015) der meistbesuchten Seiten in den USA (*Alexa*, 2014).

2.2.1 Literatur aus dem Bereich der räumlichen Analyse⁴

Dubin gilt als der Pionier in der Anwendung von Kriging-Verfahren zur Schätzung von Immobilienpreisen. Er unternahm bereits 1988 einen Versuch zur Erweiterung der hedonischen Regression um ein Maximum-Likelihood-Verfahren unter Berücksichtigung räumlicher Autokorrelation. Im Jahr 1998 stellte er dann erstmals eine Anwendung eines Kriging-Interpolators als Schätzverfahren für Immobilienpreise vor. *Basu* und *Thibodeau* (1998) verbinden zur Schätzung der Preise von Wohnimmobilien in Dallas (USA) Kriging mit einem Kleinste-Quadrate-Schätzer (Ordinary least squares, OLS). *Chica Olmo* (1995, 2007, 2013) widmet sich ausführlich der Thematik der Schätzung von Immobilienpreisen.

In *Chica Olmos* (2007) vielfach zitiertem Artikel "Prediction of housing location price by a multivariate spatial method" werden zur Schätzung der Preise von Wohnimmobilien ein Ordinary Kriging (OK) und ein Ordinary Co-Kriging Ansatz (OCK) genutzt. Im Jahr 2009 schätzen *Larraz* und *Monterro* wiederum in Toledo (Spanien) die Preise für Gewerbeimmobilien unter der Anwendung von Kriging Interpolatoren. *Páez* (2007) vergleicht Methoden des Moving Window Kriging, der GWR und der Moving Window Regression zum Messen von Effekten der räumlichen Abhängigkeit. *Bourrousa et al.* (2010) stellen einen umfassenden Modellvergleich von OLS-, geostatistischen und Trendmodellen an und nutzen dabei u. a. ein hedonisches Modell in Kombination mit einem Ordinary Kriging Modell. Eine Anwendung eines Regression-Krigings stellt *Tsutsumi* (2011) zur Berechnung von Grundstückspreisen in Tokio (*Japan*) vor. *Kuntz* und *Helbich* (2014) stellen einen empirischen Vergleich an, mit dem Schwerpunkt auf der Validierung von Kriging und OCK Interpolatoren zum "Kartographieren" von Einfamilien-Wohneigentum in Wien, Österreich. Die nachfolgende Tabelle 2-1 gibt eine Übersicht über Preisschätzverfahren in der Immobilienmarktforschung.

4 Alle nachfolgenden, in den erwähnten Studien vorkommenden Verfahren werden im im Kapitel 4 (Methoden) ausführlich erläutert.

Autoren	Anwendungsfall Preisschätzung	Verwendete Methoden
<i>Chica Olmo</i> (2007)	Eigentumswohnungen in Granada/ Spanien	OK/OCK
<i>Bourassa et al.</i> (2010)	Einfamilienhäuser in Kentucky/ USA	HED/OK
<i>Montarro und Larraz</i> (2011)	Gewerbeimmobilien Toledo / Spanien	HED/OCK/OK
<i>Wong et al.</i>	Eigentumswohnungen Apartmenthochhäuser in Hong Kong/China	Hedonische Regression
<i>Páez et al.</i> (2008)	Einzelhäuser in Toronto/Kanada	Moving Window Regression/GWR
<i>Tsutsumi et al.</i> (2011)	Grundpreise in Tokio/Japan	Regression Kriging
<i>Kuntz und Helbich</i> (2014)	Wohnungseigentum Wien/ Österreich	Detrend Kriging/Universal Kriging (KED)/Universal Co – Kriging

Tabelle 2-1: Übersicht zu Studien über Preisschätzverfahren in der Immobilienmarktforschung.

Es finden sich in der Literatur zur räumlichen Schätzung von Immobilienpreisen Erweiterungen herkömmlicher Regressionsanalysen um geographische Komponenten. *Brundson et al.* (1998) diskutieren in ihrem Artikel "Geographically weighted regression modelling spatial non-stationarity" erstmals das Prinzip der geographisch gewichteten Regressionsanalyse (GWR) und demonstrieren damit das erste Mal die Möglichkeit, Regressionskoeffizienten auf Karten darzustellen. Seither sind zahlreiche Anwendungen von GWR zum Schätzen von Immobilienpreisen aufgetaucht. *Mc Clusky et al.* (2011) versuchen Immobiliensubmärkte über geographisch gewichtete Regression zu identifizieren. Eine weitere Anwendung von GWR zur exemplarischen Analyse des Belfast Mietmarktes stammt von *McCor et al.* (2014). Weitere Lösungsansätze in Kombination mit GWR finden sich in Studien zu Lagemodellierungen. Beispiele für Lagemodellierungen geben u. a. *Prinz und Reithofer* (2005), wobei Wohnstandorte in der Stadt Salzburg betrachtet und hierfür ein indikatorbasiertes multikriterielles Modell entwickelt wurden.

2.2.2 *Literatur aus dem Bereich des Maschinellen Lernens*

Neben Lösungsmöglichkeiten der räumlichen Statistik lohnt sich ein Blick auf Verfahren des Maschinellen Lernens, wobei vorausszuschicken ist, dass es sich bei diesen Verfahren um nicht räumliche Verfahren handelt. Es gibt jedoch erste Ansätze in der Forschung, die Methoden des Maschinellen Lernens auch für räumliche Analysen zu nutzen. Große Beliebtheit erlangten die Methoden beispielsweise in der Fernerkundung.

Antipov und Pokryshevskaya (2010) stellen einen Ansatz der Preisschätzung von Wohnungen mittels der Methode Random Forest (RF) dar. *Gu et al.* (2010) verbinden einen genetischen Algorithmus mit Random Forest zur Vorhersage von Immobilienpreisen. In ihrem Artikel "Machine Learning and the Spatial Structure of House Prices and Housing Returns" wenden *Caplin et al.* (2008) Methoden des Maschinellen Lernens zur Ermittlung von Immobilienpreisen an und untersuchen mit ihrem Modell die räumliche Struktur von Transaktionsdaten. *Bark und Bae* (2014) vergleichen mehrere Algorithmen aus dem Maschinellen Lernen zur Schätzung von Immobilienpreisen. Ein weiteres Anwendungspotential haben (hierarchische) Clusterverfahren. *Hepsen und Vatansever* (2012) stellen eine Anwendung verschiedener hierarchischer Clusterverfahren für die Clusterung der Mietmärkte von 71 türkischen Großstädten vor.

2.2.3 *Literatur aus dem Bereich der Geovisualisierung*

Studien aus dem Feld der Geovisualisierung behandeln meist Teilaspekte, die für die Überprüfung der zweiten formulierten These in Frage kommen. Für diverse in der Forschung im Diskurs stehende Lösungsansätze sei auf *Huang und Gartner* (2012) verwiesen, die einen Ansatz der kollaborativen Filterung besprechen. Die Forschung über die Thematiken cloudbasiertes Mapping, Webmapping und Map APIs findet sich bei *Peterson* (2012, 2014). Eine relevante Studie zur Thematik „Gridmapping“ veröffentlichte *Trainor* (2010).

2.3 **Technologische Recherche**

Zur Erhebung des Status quo wurde vom Autor, komplementär zur Literaturrecherche, eine technologische Recherche durchgeführt. Nachfolgend werden die Ergebnisse dieser technologischen Recherche über die in Immobilienporta-

len zum Einsatz kommenden Methoden der räumlichen Analyse und der Geovisualisierung präsentiert. Zur standardisierten Untersuchung verfügbarer Immobilienportale wurde vom Verfasser eine Auswertematrix konzipiert. Insgesamt wurden anhand dieser Matrix 32 Immobilienportale nach jeweils 14 Kriterien untersucht. Einzelne Methoden, die für die räumliche Sicht von Immobilienportalen als besonders relevant einzustufen sind, werden im Kapitel 2.3 ausführlicher erläutert.

Bei der Auswahl der Portale wurden verschiedene Gesichtspunkte berücksichtigt. Es wurden die Immobilienportale mit den nach *Alexa.com*⁵ höchsten Besucherzahlen ausgewählt. Danach erfolgte eine Selektion der nach *Alexa.com* drei bis fünf größten nationalen Portale in den USA, Großbritannien, Deutschland, Österreich und der Schweiz. Hierbei zeigen die Besucherzahlen, dass im angelsächsischen und im deutschen Sprachraum die meisten Immobilienportale zu finden sind. Die Liste der Portale wurde anschließend noch um einige wenige unterschiedliche innovative Ansätze verfolgende Immobilienportale erweitert. Es fand eine Bereinigung aller ausgewählten Portale um Portale ohne Angebotskarten statt. Übrig blieben 32 Immobilienportale, die anhand von 14 Kriterien untersucht wurden. Für alle spezifizierten Portale wurde jeweils festgestellt, ob sie ein Kriterium aufweisen, teilweise aufweisen oder nicht aufweisen. Im Anhang 2 findet sich die Auswertematrix, mittels derer die gewählten Portale analysiert wurden. Die überprüften Kriterien wurden in Anlehnung an die Ergebnisse der Expertenbefragung gewählt und umfassen zwei Kategorien, eine geoinformatische, raumanalytische Kategorie und eine Geovisualisierungskategorie (siehe Tabelle 2-2).

5 *Alexa.com* zählt über eine Toolbar im Browser die Seitenaufrufe eines Nutzers und erstellt mit diesen Daten eine quartalsweise Hochrechnung. Auf Grundlage dieser Hochrechnung wird im Dreimonatsintervall eine Liste mit den 100 000 meist aufgerufen Webseiten erstellt (*Alexa*, 2014).

Geoinformatik	Geovisualisierung
Immobilienpreiskarte	Immobilienpreiskarte
Künstliche Administrative Grenzen	Dynamische Legende
Statistische Auswertungen	Grundkarte/Basemap
Preismodellierung mittels räumlicher Statistik	API zu Angebotsdaten
Immobilienindex	Zusatzinformation in Marker/Layer
Hedonische Regression	Pin/Marker Problematik
Indikatorenbasierte distanzabhängige Suche	Mobile Angebotskarten

Tabelle 2-2: Kategorien technologische Recherche. Der Punkt „Immobilienpreiskarte“ fällt in beide Kategorien. Beim Punkt Basemap/Grundkarte erfolgte eine Auswertung nach dem Grundkarten-Anbieter.

Ergebnisse der Recherche

Ein Ergebnis der Recherche ist eine Unterscheidung von Immobilienportalen in drei Kategorien.

- **Kategorie 1:** Immobilienportale ohne Angebots- und Immobilienpreiskarten. Portale dieser Kategorie wurden nicht weiter berücksichtigt.
- **Kategorie 2:** Immobilienportale mit Angebotskarten und ohne Immobilienpreiskarten.
- **Kategorie 3:** Immobilienportale mit Angebotskarten und auf Basis der hedonischen Regression erstellten Immobilienpreiskarten.

Tageszeitungen spielen bei der webbasierten Immobiliensuche praktisch keine Rolle mehr (Enderle, 2009). Diese Beobachtung von Enderle (2009) kann durch die durchgeführte Analyse bestätigt werden. Zwei Ausnahmen davon sind im deutschen Sprachraum zu nennen. Das ist einerseits die österreichische Tageszeitung *Der Standard*, die in Österreich das nach Nutzerzahlen und Bekanntheit zweitwichtigste Immobilienportal betreibt, sowie in Deutschland die *Bildzeitung*, die über den *Axel Springer Verlag* das Portal *Immonet.de* führt (Wer zu Wem Firmenverzeichnis, 2014). Abbildung 2-2 zeigt die Häufigkeiten, mit denen bestimmte Kriterien in Portalen festgestellt wurden. Die Portale sind hierbei, ihrer Herkunft entsprechend, nach Nationen zusammengefasst.

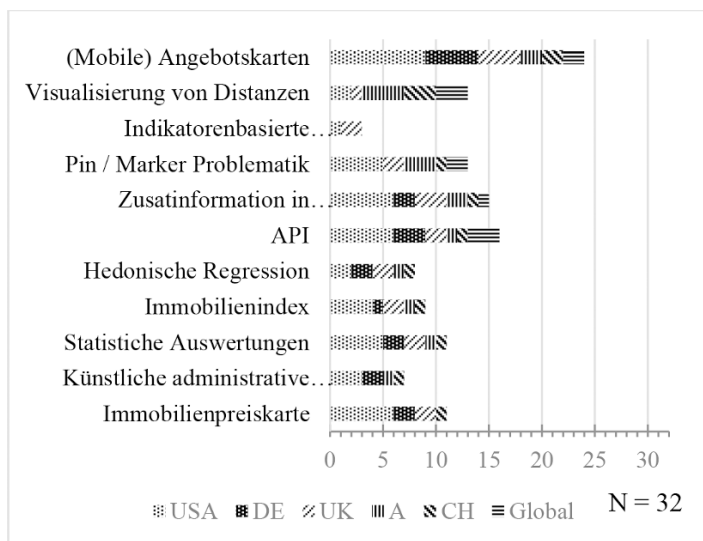


Abb. 2-2: Häufigkeit der Anwendung von Methoden der Geoinformatik und Visualisierung in 32 untersuchten Immobilienportalen nach Nationen zusammengefasst. (Eigene Darstellung).

Eine der auffälligsten Beobachtungen ist, dass 24 der 32 Portale mobile Angebotskarten veröffentlichen. Dies geschieht sowohl in Form mobiler Applikationen für die gängigen mobilen Betriebssysteme (Android/iOS/Windows Phone) als auch plattformübergreifend im Browser. Mobile Angebotskarten und responsives Webdesign⁶ werden mittlerweile bei der Mehrheit der beobachteten Portale umgesetzt.

Alle Immobilienportale setzen auf die Verwendung von Grundkarten, die über Map APIs bereitgestellt werden. Als Grundlage (Basemap) wird hierbei der in das Medium des Word Wide Web (WWW) portierte Ansatz aus der thematischen Kartographie verstanden, eine topographische Grundlage als geometrische Bezugs- und Orientierungsmöglichkeit bereitzustellen (*Arnberger, 1997*). Basemaps werden in der Regel als „gekachelte“ Webservices (TMS, Tiled

6 Responsives Webdesign meint das erstmals von Marcotte (2010) erläuterte Paradigma, das Webseiten auf unterschiedliche Bildschirmgrößen hin reagieren und eine auf das jeweilige Endgerät optimierte Darstellung erfolgt.

Mapservices) mittels Raster oder Vector Tiles über Map-APIs bereitgestellt. Die Webanwendung *Housingmaps.com* (*Housingmaps*, 2014) ist beispielsweise eine der frühesten bekannten Karten-Mashups (*Peterson*, 2014)⁷. *Housingmaps* stellte Immobilienangebote einer Webkarte von *Google maps* dar. Die verwendeten Daten stammten dabei aus dem in den USA verbreiteten Online-Kleinanzeigen dienst *Craigslist* (*Craigslist*, 2005).

Die Dominanz von *Google maps* als Grundkarte fiel bei der Untersuchung besonders auf (siehe Abbildung 2-3). So nutzen 21 von 32 Portalen die Google API als Schnittstelle zu *Google maps*. Alle 31 betrachteten Portale nutzen Grundkarten proprietärer Anbieter.

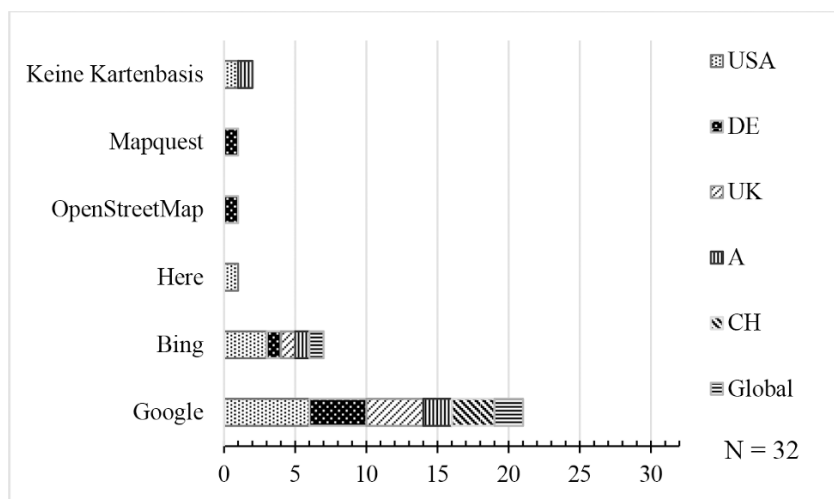


Abb. 2-3: Diagramm der in Immobilienportalen verwendeten Basemaps unterschiedlicher Map-API-Anbieter. (Eigene Darstellung).

Angebote werden bei allen Portalen in ihrer Verortung als Positionssignatur bzw. als „Pin map“ dargestellt. Hierbei wird häufig das Symbol eines Hauses gewählt. Signaturen werden meist nicht maßstabsabhängig skaliert. Eine weite-

7 Der aus der Popmusik stammende Begriff „Mashup“ wird in dieser Arbeit wie von Batty et al. (2010) charakterisiert verwendet. In diesem Verständnis bezieht sich der Begriff Mashup auf Webseiten, die Daten aus verschiedenen Quellen zu neuen Online-Diensten – im Falle von Karten-Mashups zu integrierten Kartendiensten – verflechten.

re, oft gewählte Darstellung erfolgt mittels Angabe des Preises als absoluter Zahlenwert (Zahlenwertsignatur) innerhalb geometrischer Formen, wie z. B. innerhalb eines Rechtecks. Auch hierbei treten häufig Probleme mit überlappenden Signaturen auf. Zwar wird eine jeweils von den Map-APIs bereitgestellte maßstabsabhängige Generalisierung der Grundkarte angeboten, bei den Markern selbst fehlt jedoch jegliche maßstabsabhängige Generalisierung. Eine Änderung der Zoomstufe⁸ bewirkt bei 13 Immobilienportalen keinerlei und bei neun untersuchten Portalen teilweise starke Überlappungen der Positionsmarker. Ein Fehlen von maßstabsabhängigen Generalisierungen im Zusammenhang mit Map-APIs bzw. Webkarten ist ein bestehendes Forschungsproblem, bekannt als „Icon Cluttering“, ausführlich diskutiert u. a. von *Huang und Gartner* (2012). In Kapitel 2.3.4 kommt es zu einer weiteren Diskussion der Problematik. Immerhin sechs Portale bieten Lösungsansätze für diese Problematik, meist in Form einer Aggregation durch Angabe einer Ziffer innerhalb des Positionsmarkers oder über eine Clusterung von Angeboten bei kleinerem Maßstab. Es wird teils beim Herauszoomen in Zoomstufen kleinerer Maßstabsebenen die Anzahl der aggregierten Angebote in eine geometrische Signatur geschrieben. Über eine Farbkodierung der Positionsmarker und das Einbinden von Popups integrieren viele Portale Zusatzinformation in die Marker. Einige Portale stellen zusätzliche Datenschichten bereit. Eine distanzbasierte Angebotssuche ist in wenigen Portalen möglich. Teils erfolgt eine Visualisierung von euklidischen Distanzen ähnlich einem Puffer, basierend auf denen von Suchenden festgelegten Umkreisen. Ob hierbei netzwerkbasiert oder über euklidische Distanzen gesucht wird, kann nur vermutet werden. Eine Suche über zusätzliche Indikatoren, z. B. über eine räumliche Selektion von Angeboten nach topologischen Regeln, ist in keinem Portal möglich.

Fast alle großen Portale bieten eine eigene Programmierschnittstelle zu ihren Angeboten. Hierdurch wird es möglich, diese Angebote dann geokodiert oder revers gekodiert für eigene Anwendungen zu beziehen. In den meisten Fällen handelt es sich dabei um APIs die dem Paradigma der RESTful Webservices (REST) unterliegen⁹. Alle größeren Portale stellen ihren Nutzern Immobilien-

8 Auf der Wiki des *Openstreetmap* Projektes (Openstreetmap, 2014) findet sich eine Übersicht über den Maßstab der jeweiligen Zoomstufen. Beispielsweise bedeutet in der Openstreetmap eine Zoomstufe von 13 einen Kartenmaßstab von 1 : 19 093.

9 Einem von Fielding (2000) vorgeschlagenen Paradigma für Webservices, dass gewünschte Informationen im Falle von Angebotsdaten der Immobilienportale über Formate wie XML

preiskarten bereit. Die errechneten Preise werden dabei in der Mehrzahl der Fälle auf willkürlich gewählte räumliche Bezugseinheiten umgerechnet und als Choroplethenkarten dargestellt. Die Preise sind dabei reine Durchschnittspreise oder ein auf z. B. einen Stadtteil umgelegter Median, ein Durchschnittspreis oder das auf eine Bezugsgeometrie umgerechnete Ergebnis einer hedonischen Regression. *Homegate* (2014) weist ganz klar aus, dass Preise mittels einer hedonischen Regression gerechnet werden. Die räumlichen Bezugseinheiten werden nicht angegeben. Neben einfachen deskriptiven Maßen zur Preisberechnung ist die Methode der hedonischen Regression die am weitesten verbreitete Methode der Preisberechnung. Acht Portale geben ganz klar an, dass diese Form der Immobilienpreisschätzung eingesetzt wird. Es ist jedoch anzunehmen, dass noch mehr Portale auf diese Methode setzen. Räumliche Zusammenhänge finden in der Methode der hedonischen Regressionsgleichungen keine Beachtung. Sie zählen demgemäß zu den nicht räumlichen Verfahren. Die Ergebnisse der hedonischen Regression werden dabei ebenfalls auf fachfremd verwendete Raumeinheiten umgelegt. Die Ergebnisse hedonischer Regressionsgleichungen sind vielfach Basis von Immobilienpreisindizes, die von Portalen veröffentlicht werden. Insgesamt neun Portale bieten Immobilienpreisindizes an, wovon wiederum acht Fällen auf einer hedonischen Regression – teils gekoppelt mit einem, teils ohne ein automatisiertes Immobilienbewertungsverfahren – beruhen.

Keines der Portale nutzt in den veröffentlichten Preiskarten eine dynamische Legende, weder bei den veröffentlichten Immobilienpreiskarten noch bei den Angebotskarten.

2.4 Status quo der Methoden

Nachfolgend wird der *Status quo* der Methoden der Immobilienpreisanalyse und -darstellung von Immobilienportalen aus Forschungssicht der Geoinformatik adressiert. Im Einzelnen sind das: die Methode der hedonischen Regression (siehe Kapitel 2.4.1) und die damit im Zusammenhang stehenden fachfremd

als JavaScript Object Notation (JSON) oder in der Hypertext Markup Language (HTML) ausgegeben werden.

gewählten räumlichen Bezugseinheiten (siehe Kapitel 2.4.2), die nicht dynamischen Legenden (siehe Kapitel 2.4.3), die Verortung von Angeboten über Punktsignaturen und das damit entstehende Problem der Signaturenüberlagerung (Icon Cluttering) (siehe Kapitel 2.4.4) sowie die Integration von Distanz und räumlichen Indikatoren in Immobilienportale (siehe Kapitel 2.4.5).

2.4.1 Hedonische Modellverfahren

Hedonische Regressionsgleichungen, als methodische Vergleichswertverfahren, sind ein Standardverfahren der Miet- und Kaufpreisschätzung bei Immobilienportalen. Es erfolgt ganz allgemein eine Preisschätzung von Produkten mit laufender Veränderung des Preis-Leistungs-Verhältnisses. Der Preis wird über eine Qualitätsanpassung ständig neu berechnet (Gordon, 1990). Der Preis einer Immobilie wird über ihre Qualität, ausgedrückt in den intrinsischen Werten (Ausstattungsmerkmale wie z. B. Fläche, Zimmeranzahl, Etage etc.), und über die Verfügbarkeit geschätzt. Die räumliche Autokorrelation¹⁰ von Immobilien wird hierbei vernachlässigt (eine ausführliche Diskussion des Konzeptes der räumlichen Autokorrelation folgt in Kapitel 4). In Bezug auf die Anwendung von Immobilienportalen ist demnach eine Vernachlässigung von Toblers (1970) postuliertem Ersten Gesetz der Geographie: „Everything is related to everything else, but near things are more related than distant things“ festzustellen. Das Feld der hedonischen Regression geht auf Court (1939) und Griliches (1961) zurück. Angewendet wird die hedonische Regression zur Preisschätzung eines Produktes bei laufender zeitlicher Veränderung des Preis-Leistungs-Verhältnisses. Gordon (1990) spricht deshalb von einer Preisberechnung mit Qualitätsanpassung.

Die Grundlage der hedonischen Regression ist die hedonische Funktion, in der verschiedene Varianten eines Produkts k einem Preis zugeordnet werden (Triplet, 2004). Laut Diewert (2003) werden einzelne Varianten eines Produktes mittels verschiedener Merkmalsausprägungen der dem Produkt zugesprochenen m Eigenschaften unterschieden, das ergibt einen Tupel in der Form $K = (X_{k1}, \dots, X_{km})$. Als Preis wird laut Diewert (2003) immer der Preis

10 Bei etlichen räumlichen Prozessen sind die Ähnlichkeiten oder die Unterschiede der von einer gewissen räumlichen Distanz getrennten Umstände größer, als bei zufälligen Prozessen zu erwarten wäre (Legendre, 1993). In Kapitel 4.1.2 erfolgt eine umfassende Erklärung der räumlichen Autokorrelation.

gesehen, den ein möglicher Käufer bereit ist zu zahlen. Der Preis wird in verschiedene Komponenten aufgeteilt, d. h., das Verfahren beruht auf der Annahme, dass ein Gesamtpreis eines Produktes auf Grundlage unterschiedlicher Einzelkomponenten eines Produktes berechnet werden kann. Eine Preisschätzung erfolgt nur in Abhängigkeit von einer betrachteten Zeitperiode (t). Der Faktor Zeit muss also in die Gleichung integriert werden. Das kann auf zwei Arten geschehen. Man findet für t entweder eine Dummy-Variable¹¹ oder berechnet die hedonische Regression für einzelne Zeitintervalle. Die betrachteten Immobilienportale schätzten ihre Preise dabei ausschließlich für das Zeitintervall "Quartal", also in vierteljährlichen Zeitschritten.

Exemplarisch wird nachfolgend die Semilog-Variante von *Gordon* (1990) beschrieben: Für ein Objekt k wird der Preis P_k ermittelt.

$$\ln P_k = \beta_0 + \sum_{j=1}^m \beta_j X_{kj} + \epsilon_k$$

Formel 2-1: Semilog-Variante der hedonischen Regression.

Der zu ermittelnde Preis ist die Primärvariable, die Sekundärvariable m wird zur Preisermittlung herangezogen (*Li und Heap*, 2008). Es werden Anteile am Gesamtpreis berechnet, die auf einzelne Eigenschaften bzw. Ausprägungen eines Produktes zurückgehen. Es lassen sich mit dem Regressionskoeffizienten der hedonischen Regression Marktpreise berechnen, die noch nicht bekannt sind und unterschiedliche Produktvarianten (z. B. Produktvariante Reihenhaus oder Dachgeschosswohnung) schätzen. Die Marktausprägungen m sind Ausprägungen des Einzelobjektes. Räumliche Autokorrelation wird bei der hedonischen Regression außer Acht gelassen. Immobilienportale legen die Ergebnisse der hedonischen Regression auf fachfremd verwendete geometrische Bezugseinheiten um.

11 Binäre „Scheinvariable“, die ausdrückt, ob eine Eigenschaft zutrifft oder nicht.

2.4.2 Räumliche Bezugseinheiten

Ergebnisse der hedonischen Regression werden auf fachfremd verwendete räumliche Bezugseinheiten umgelegt. Das können Stadtteile, Postleitzahlenbereiche, statistische Blöcke etc. sein. Im Rahmen der Dissertation vom Autor untersuchte US-amerikanische Portale wiesen mehrheitlich von den Portalen selbst definierte Geometrien, genannt Nachbarschaften („Neighborhoods“), Postleitzahlenbezirke („Zip Codes“), Stadtteilgrenzen oder statistische Einheiten des US-Zensus („Core based statistical areas“) auf. Core based statistical areas der US amerikanischen nationalen Statistik bestehen beispielsweise aus Bezirken und Landkreisen die mindestens einen urbanen Kern aufweisen, mindestens 10. 000 Einwohner haben und einen hohen Grad an sozioökonomischer Verbundenheit mit dem urbanen Kern zeigen (*United States Census*, 2010).

Das Immobilienportal *Zillow.com* veröffentlicht alle „Nachbarschaften“ unter der freien „Gnu Public License“. Es handelt sich um insgesamt 7.000 Nachbarschaften, die für größere Städte gebildet wurden (*Zillow*, 2014). Abbildung 2-4 zeigt beispielhaft drei Zip-Code-Bereiche in New York Manhattan im Vergleich von der von *Zillow.com* gebildeten „Neighborhood“ namens Gramercy. Teils orientieren sich diese „Neighborhoods“ an den Zip Codes, teils sind sie größer, teils kleiner gewählt. Es ist keine Methodik für die Auswahl der Neighborhoods erkennbar. *Zillow* berechnet jeweils für jede Nachbarschaft den Median der Miet- und Kaufpreise.

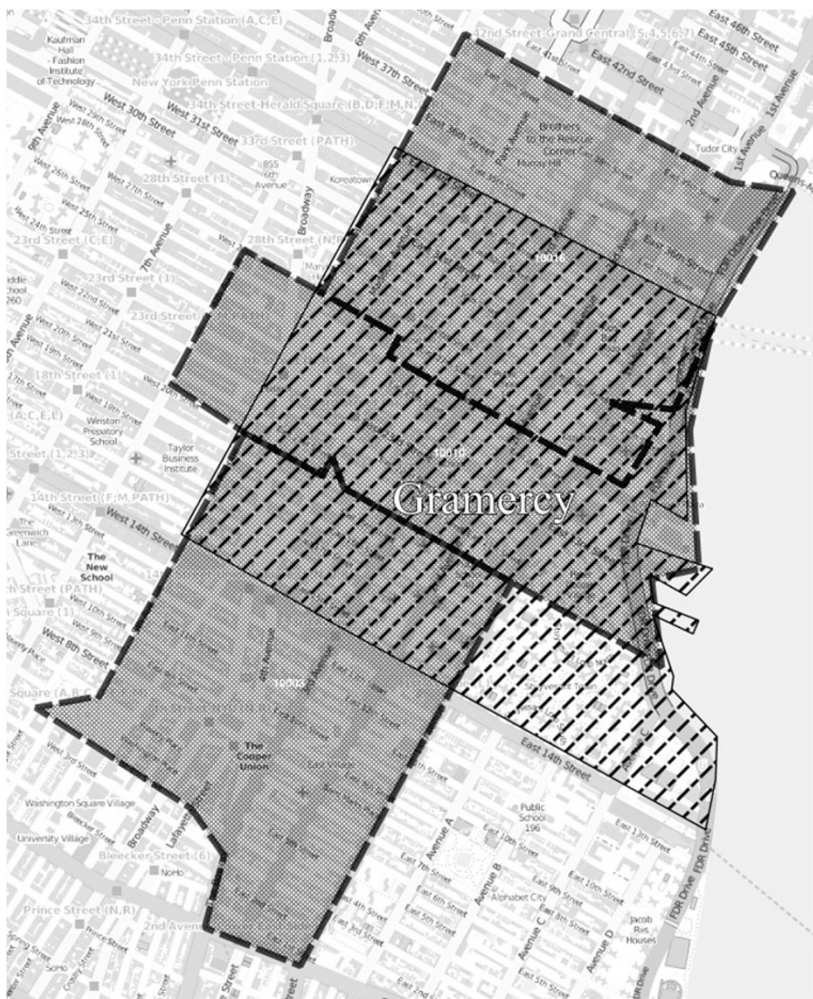


Abb. 2-4: Links: Vergleich von drei Zip-Code-Bereichen (graue Fläche) mit einer Zillow-Neighborhood in New York (schwarz diagonal gestrichelte Fläche). Maßstab: 1 : 15.000. (Eigene Darstellung, Grundkarte: © OpenStreetMap contributors).

Beim deutschen Portal *Immobilienscout 24 (IS24)* wurden als Wohnquartiere bezeichnete Bezugseinheiten der Firma *Infas Geodaten*¹² gewählt. Laut *Infas Geodaten* umfasst ein Wohnquartier im Durchschnitt 400 Haushalte und weist eine möglichst homogene Struktur auf (*Nexiga*, 2014). Die erste Erfassung der Wohnquartiere erfolgte 1994 auf Grundlage der Stimmbezirke der Bundestagswahl 1987 (alte Bundesländer) und auf Basis der Stimmbezirke der Bundestagswahl 1994 (Neue Bundesländer). 2006 waren 75.000 Wohnquartiere verfügbar (*Infas Geodaten*, 2006). Wohnquartiere beachten keine räumliche Autokorrelation. Die Größe der Wohnquartiere variiert sehr stark und ist im ruralen Raum viel größer als im dicht besiedelten urbanen Raum. Abbildung 2-5 zeigt einen Vergleich der Größe eines Wohnquartiers in Potsdam-Babelsberg mit einem Postleitzahlenbezirk.



Abb. 2-5: Geometrie eines Postleitzahlenbezirks (graue Fläche) im Vergleich zu einem Wohnquartier der Firma *Infas/Nexiga*. 1 : 30.000 (schwarz gestrichelte Fläche). (Eigene Darstellung, Grundkarte: © OpenStreetMap contributors).

12 Infas Geodaten wurde 2014 in Nexiga GmbH umbenannt.

2.4.3 Nicht dynamische Legenden

Bei den veröffentlichten Immobilienpreiskarten dominieren nicht dynamische Legenden. Diese Legenden passen sich nicht an die im Kartenausschnitt vorhandenen Immobilienpreise an, dieselben Klassen werden uniform für ganz Deutschland verwendet. Das führt dazu, dass in Immobilienpreiskarten unabhängig vom tatsächlichen Preisverhältnis des Untersuchungsgebietes dieselben Preisklassen bei Immobilienpreisen angezeigt werden. Das Histogramm der Mietpreise Münchens (Stand Dezember 2012) in Abbildung 8 verdeutlicht diese Problematik¹³. Wohnungen zur Miete in München werden auf dem Immobilienportal *IS24* zum Großteil der teuersten Preisklasse von größer als 12 €/m² zugewiesen, und das, obwohl laut Angaben auf *IS24* der Durchschnittspreis beispielsweise im Stadtteil München-Bogenhausen 13, 6 €/m² (Stand Dezember 2014) betrug, während rund um den Viktualienmarkt 17,10 €/m² Durchschnittsmiete verlangt wurden. Es erfolgt eine fast uniforme Zuweisung zur Preisklasse von größer als 12 €/m². Die Klassenzuweisung eines Durchschnittspreises auf Stadtteilebene zur höchsten Preisklasse ist somit verzerrend. Dieselbe Legende wird im Kontrast dazu auch auf Gebiete mit einer sich gänzlich unterscheidenden Marktzusammensetzung angewendet. Beispielsweise wird ganz Osterode am Harz der niedrigsten Preisklasse von geringer als 4 €/m² zugewiesen, obwohl dort der Durchschnittspreis im ersten Quartal 2014 bei 4,3 €/m² lag. Die Klassen passen sich nicht an die im Kartenausschnitt tatsächlichen Mietpreise an. Dieselben Preisklassen werden ohne Rücksicht auf regionale Begebenheiten deutschlandweit angewendet. Die Kartendarstellung orientiert sich daher nicht an den tatsächlichen Marktverhältnissen, es erfolgt keine Anpassung der Legenden an die tatsächliche vor Ort herrschende Werteverteilung der Preise. Zur Bestätigung dieser Beobachtung wurden 1.152 Datenpunkte vom 04.12.2014 mit Angebotsmieten der Stadt München über die Rest-API des Immobilienportals *IS24* bezogen (*ImmobilienScout 24*, 2014). Die danach gebildete deskriptive Statistik der Stadt München weist Extremwerte von 8,84 €/m² bis 36 €/m² bei einem Mittelwert von 17,87 €/m² und einer Standardabweichung von 4,51 €/m² auf. Abbildung 2-6 zeigt ein Histogramm der Preisverteilung. Hiermit wird belegt, dass die Kartendarstel-

13 Die Daten zur Berechnung des Histogramms wurden über die Rest API des Immobilienportals *IS24* bezogen: <http://api.immobilienscout24.de/>.

lung (siehe Abbildung 2-7) und Klassenbildung der veröffentlichten Mietpreiskarte keine plausible Klasseneinteilung zeigen.

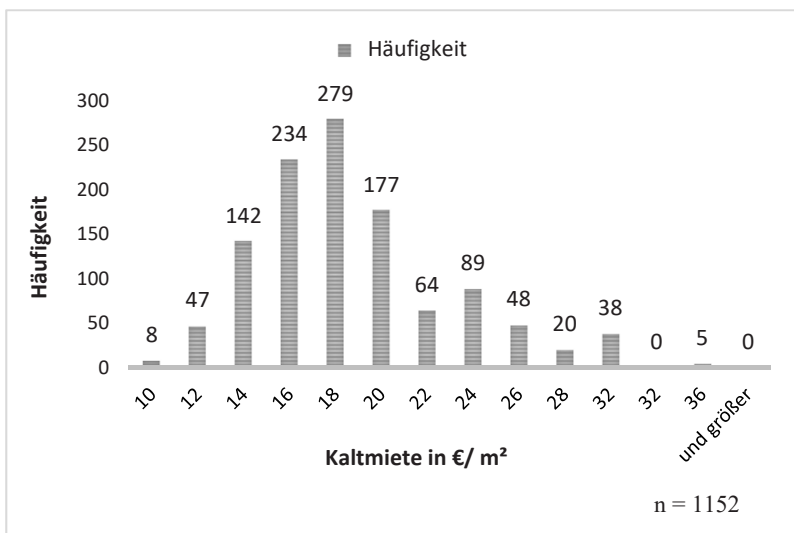


Abb. 2-6: Histogramm der Nettokaltmieten der Stadt München. (Eigene Darstellung).



Abb. 2-7: Ausschnitte aus Mietpreiskarten des Immobilienportals IS24 Oben: München. Unten: Osterode im Harz. (ImmobilienScout 24, Kartendaten © 2015; GeobasisDE/ BKG(©2009), Google)

2.4.4 Punktsignaturen und das „Icon-Cluttering“-Problem

Bei einer ungefilterten Darstellung von Immobilienangeboten als Punktsignaturen kommt es zu dem bekannten Problem der Signaturenüberlagerung (Icon Cluttering). Die Positionsmarker überlappen und überlagern sich (Huang und Gartner, 2012). Die Lesbarkeit der Angebotskarten wird stark eingeschränkt. Es ist Nutzern nicht mehr in sinnvoller Weise möglich, Informationen über Immobilienangebote zu lesen. Besonders stark zum Tragen kommt die Problematik bei mobilen Angebotskarten. Burigat und Chittaro (2006) führen das

Beispiel einer mobilen Kartendarstellung von Restaurants an. In einem großen Maßstab ist die Darstellung der Immobilienangebote problemlos, die POIs und zur Orientierung eingeblendete Straßenverläufe in der Grundkarte sind gut ersichtlich. Bei einem Wechsel in kleinere Maßstäbe tritt das Icon-Cluttering Problem durch die Überlappung von Symbolen deutlich in Erscheinung – Straßenverläufe zur Orientierung werden durch Signaturen überlagert.



Abb. 2-8: Illustration des Icon-Cluttering-Problems nach *Huang und Gartner* (2011)

2.4.5 Integration von Distanz

Die Relevanz der Integration distanzbasierter Suche steigt bei Immobilienportalen. Jedoch wird Distanz in den meisten Fällen über eine rein euklidische Distanz und nicht über eine Netzwerk-Distanz ausgedrückt. Das heißt, dass beispielsweise die Immobiliensuche auf Angebote innerhalb eines in euklidischer Distanz definierten Puffers eingeschränkt wird. Einige Portale wie *Trulia* oder *Immobilienscout* verknüpfen die Suche mit der Berechnung von Reisezeiten, die über die Nutzung von externen APIs berechnet werden. Im Falle von *IS24* ist es möglich, Angebote über Reisezeiten zu definierten Zielen im Öffentlichen Personen Nahverkehr (ÖPNV) zu filtern. Das Portal *Trulia* ermöglicht es Nutzern über definierte Punkte und als Filter definierte Reisezeiten in einem Straßennetzwerk und im ÖPNV zu berechnen und darüber Angebote zu filtern. Visualisiert werden die Reisezeiten bei *Trulia* nicht, es erfolgt nur eine Anzeige der gefilterten Angebote. Eine Möglichkeit, Angebote über eine multimodale Verknüpfung von Reisemitteln zu filtern, existiert in Portalen bis dato nicht.

2.4.6 *Integration von räumlichen Indikatoren*

Immobilienportale setzen vielfach auf die Integration von räumlichen Indikatoren bei ihren Angebotskarten. Jedoch finden sich sehr große räumliche Unterschiede. Bei Immobilienportalen im deutschsprachigen Raum werden als für die Lage wichtig erachtete Punkte als POIs visualisiert. Es kommt zur Darstellung von POIs wie z. B. Gesundheitseinrichtungen, Schulen, ÖPNV-Haltestellen als zusätzlichen Kartenschichten. Bei angelsächsischen Portalen wird mit Lageindikatoren sehr offen umgegangen, es gibt thematische Schichten mit der Bewertung von Schulqualität, Kriminalität, der Gefahr vor Naturgefahren und vieles mehr.

In keinem der Portale wird jedoch transparent offengelegt, wie diese Bewertungen zustande kamen, eine Art räumlicher Lageindex wird in keinem der Portale gebildet.

2.4.7 *Fazit*

Die durchgeführte technologische Recherche zeigt, dass die räumliche Sicht von Immobilienportalen eingeschränkt ist. Es werden keine räumlichen Analysemethoden zur Preisschätzung angewandt. Nicht räumliche Schätzergebnisse werden auf fachfremd verwendete Bezugseinheiten bezogen. Es existieren nur wenige Forschungsarbeiten über Immobilienportale, jedoch gibt es eine Reihe an Forschungsarbeiten, die aus wissenschaftlicher Sicht Teilaspekte der identifizierten Fragestellungen behandeln.

Keine der recherchierten Forschungsarbeiten beschäftigt sich mit der Schätzung von Mietpreisen, sondern vielmehr mit der räumlichen Schätzung von Kaufpreisen. Keine der Arbeiten bringt die untersuchten Methoden in einen operationellen Zusammenhang. Die Parameterkonfiguration für Schätzverfahren und die Sekundärvariablenauswahl wurden in der bisherigen räumlichen Immobilienmarktforschung noch unzureichend analysiert. Der Verfasser der vorliegenden Dissertation sieht in Verfahren der räumlichen Statistik (vor allem Schätzverfahren) und Verfahren aus dem Gridmapping mit einer Visualisierung mittels Map APIs das größte Anwendungspotential zur Belegung der formulierten Thesen.

Räumliche Analyse und Visualisierung von
Mietpreisdaten

Untersuchungen im Anwendungskontext von
Immobilienportalen

Schernthanner, H.

2017, XX, 153 S. 52 Abb., 10 Abb. in Farbe., Softcover

ISBN: 978-3-658-17773-7