

# Strategien zur Modellierung mobiler, multimedialer Tourismusinformationssysteme im FlyingPix Projekt

Christoph Eigenstetter, Nils Weber & Ilvio Bruder  
IT Science Center Rügen gGmbH  
Universität Rostock

Beim Aufbau mobiler, multimedialer Tourismusinformationssysteme sind eine Vielzahl verschiedener Anforderungen zu beachten. Adäquate Verwaltung, Speicherung, Präsentation der Daten hat direkten Einfluss auf die Modellierung des Systems. Im FlyingPix Projekt sollen Touristen bei der Reise- und explizit bei der Fotoverwaltung unterstützt werden. Hierzu werden Daten und Dokumente zu touristischen Objekten, zu georeferenzierten Informationen, zu semantischen Informationen aus einer Tourismusontologie sinnvoll integriert. So stehen dem Touristen neben herkömmlichen Tourismusinformationen auch zusätzliche, implizite Daten zur Verfügung, die weder in den Reisedaten selbst noch unbedingt durch die Touristen auf der Reise gesammelt werden können. Die notwendige Verknüpfung der verschiedenen Datenquellen erfolgt schon bei der Modellierung des Systems bzw. der konkreten Datenbank. Mittels Assoziationen zwischen Objekten und Dokumenten der unterschiedlichen Quellen wird eine Aggregation von inhaltlich zusammengehörenden Daten zu Komponenten erreicht, die dann bspw. in verschiedenen Ausgabeformaten – als HTML-Seite oder XML-Datei – dargestellt werden können. Ziel ist eine automatische Annotation von georeferenzierten Fotos mit Hintergrundwissen zum Umfeld des fotografierten Objektes.

## 1 Einführung

Ziel des Projekt FlyingPix ist die Entwicklung eines Tourismusinformationssystems, das potentielle Touristen sowohl bei der Urlaubsplanung als auch während des Aufenthaltes im Urlaubsgebiet unterstützt. Bereits vor Urlaubsantritt hat der Tourist die Möglichkeit persönliche Interessen und Vorlieben in seinem Nutzerprofil abzulegen. Das System ist dann in der Lage auf Basis dieser Angaben und der geplanten Urlaubsrouten automatisch passende touristische Objekte vorzuschlagen und diese mit Hilfe von multimedialen Inhalten zu präsentieren. Nach Abschluss der Planung können die Routen- und Objektdaten auf ein mobiles Endgerät (beispielsweise einen PDA) übertragen werden und dienen als Datenbasis für die Routenführung und ortsabhängige Präsentation von Informationen im späteren Urlaubsgebiet. Private Fotos, die während des Urlaubes aufgenommen wurden, können nach Urlaubsende vom System ausgewertet und mit orts- und kontextabhängigen Zusatzinformationen annotiert werden.

Um ein solches Informationssystem zu modellieren, sind verschiedenste Anforderungen zu beachten: Georeferenzierte touristische Objekte, die durch multimediale Dokumente beschrieben werden, stehen im Mittelpunkt. Eine tourismusorientierte Ontologie und geographische Informationen, machen implizites Wissen verfügbar, welches nicht direkt in den Daten enthalten ist. Somit werden Anfragen möglich, die z.B. alle Hotels finden, welche in der Nähe eines Yachthafens liegen und im Stil des Klassizismus erbaut wurden.

Systeme, bei denen ein Datenbankschema durch eine Ontologie bestimmt wird, sind bereits mehrfach untersucht worden. Einen Überblick über "Ontology-Driven" geographische Informationssysteme gibt z.B. [3]. In [2] wurde ein georeferenziertes System auf Basis von Ontologiewissen modelliert, bei dem ein ER-Schema aus einer Ontologiedarstellung mit Synonymwortmenge, eine Menge von Beziehungen zwischen Worten und einer Menge von Features, die die Wortekonzepte beschreiben, generiert wurde. Weitere Quellen, bspw. Dokumente und Dokumentstrukturen,

sind im Modellierungsprozess nicht berücksichtigt worden. Im Projekt GETESS [4] wurde ein Tourismussystem entwickelt, das eine Ontologie sowie Sprachverarbeitung in einer Datenbank integrierte. Die Ontologie und die touristischen Daten wurden getrennt konzipiert, eine Einbindung von Georeferenzen ist nicht erfolgt.

## 2 Anforderungen an die Datenspeicherung

Grundlage des Systems ist ein zentraler Datenspeicher, der alle Informationen zu vorhandenen touristischen Objekten im Urlaubsgebiet und die persönlichen Daten (Interessen, Fotos etc.) des Touristen enthält und diese sinnvoll miteinander verknüpft.

Als touristisches Objekt werden Konzepte betrachtet, die für einen Touristen von Interesse sein können. Dazu zählen beispielsweise Hotels, Sehenswürdigkeiten, besondere landschaftliche Punkte – z.B. Gewässer, Wälder etc. –, Restaurants, Sporteinrichtungen sowie Ereignisse und Personen. Da jeder dieser Objekttypen unterschiedliche Merkmale (Attribute) besitzt, jedoch einige Objekttypen bestimmte Merkmale gemeinsam haben, sollen objektorientierte Techniken wie Vererbung zum Einsatz kommen.

Um geographische Funktionen, wie z.B. Umkreissuche, Entfernungsbestimmung oder Routenführung, nutzen zu können, sollte das System spezielle Datentypen zur Speicherung von geographischen Formen sowie Methoden für deren Verarbeitung zur Verfügung stellen.

Informationen zu touristischen Objekten werden in Form von multimedialen Dokumenten gespeichert. Ein multimediales Dokument kann aus Text-, Bild-, Audio- oder Videodaten bestehen. Jedes Dokument wird durch einen Metadatensatz ergänzt, der sowohl für die automatische Bildannotation als auch für effiziente Suchverfahren benötigt wird und Angaben wie z.B. Auflösung, Farbraum, Copyrightangaben etc. enthält. Grundlage für das Design der Dokumentverwaltung waren Konzepte aus [1].

Für die automatische kontextabhängige Annotation von Fotos wird zusätzliches implizites Wissen benötigt, welches nicht direkt aus den gespeicherten Daten ableitbar ist. Zu diesem Zweck wird eine auf Konzepte aus der Domäne Tourismus beschränkte Ontologie entwickelt, die in das Datenbankschema integriert wurde.

Da Applikationen für den Tourismusbereich bevorzugt von verschiedensprachigen Anwendern genutzt werden, ist es sinnvoll die gespeicherten Textinformationen mehrsprachig verfügbar und durchsuchbar zu machen. Aus diesem Grund wird die Verwendung des XML-Formats für die Speicherung von Textdokumenten vorausgesetzt.

## 3 Ontologie

Eine wesentliche Anforderung an das Design des Datenspeichers war die Bereitstellung von implizitem Wissen, welches nicht direkt aus den gespeicherten Daten ableitbar ist. So sind beispielsweise Informationen über Züge und Bahnhöfe abfragbar, es ist jedoch nicht möglich herauszufinden, welcher Zug durch welche Bahnhöfe fährt. Um Anfragen dieser Art zu ermöglichen, sollte der Datenspeicher einen Mechanismus zur Verfügung stellen, mit dem sich Beziehungen zwischen Objekten darstellen lassen.

Als Lösung für dieses Problem bot sich die Nutzung einer Ontologie an. Erfahrungen aus dem Projekt GETESS [4] zeigen, dass Ontologien bei der Nutzung in Informationssystemen möglichst klein, mit wenig Konzepten und Beziehungen technisch sinnvoll bei der Modellierung umgesetzt werden können. Aus diesem Grund wurde eine Ontologie entwickelt, die auf die wichtigsten Konzepte aus der Domäne Tourismus beschränkt ist (siehe Abbildung 1).

Die Wurzel der Ontologie ist das Konzept "Object". Darunter wurden die vier Hauptkonzepte "TouristicObject", "Person", "Event" und "Media" definiert, von denen alle weiteren spezialisierten Unterkonzepte abgeleitet werden. Das Konzept "TouristicObject" beschreibt Objekte,

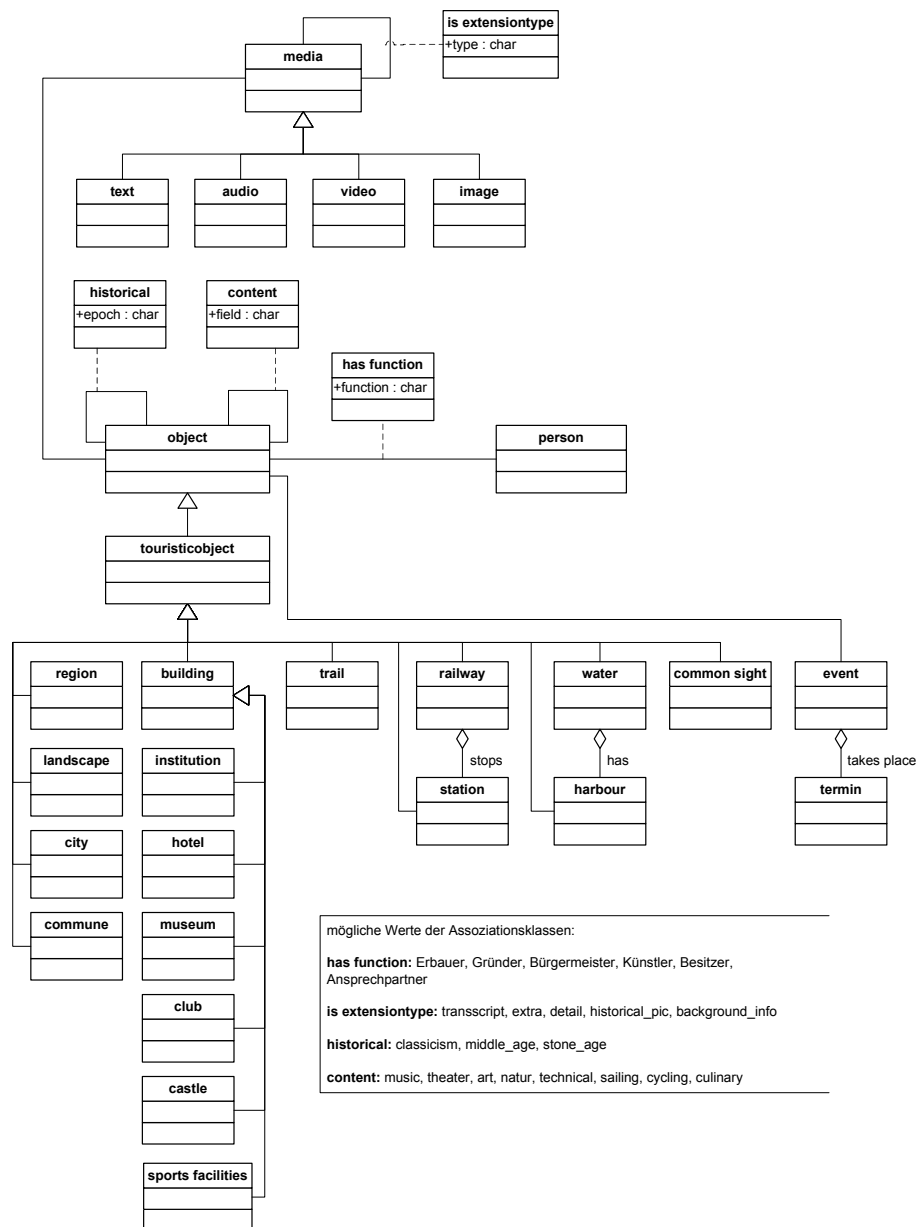


Abbildung 1: Ontologiestruktur

die eine feste geographische Position besitzen und stellt die physische Ebene der Ontologie dar. Die repräsentative Ebene wird im Konzept "Media" gekapselt, welchem Multimediale Dokumente zugeordnet werden, die Informationen zu anderen Konzepten enthalten. Die Konzepte "Person" und "Event" repräsentieren Personen und Ereignisse und gleichzeitig die personelle und zeitliche Ebene.

Beziehungen zwischen den einzelnen Konzepten werden in Form von Assoziationsklassen dargestellt. Eine solche Assoziationsklasse definiert den Typ der assoziierten Konzepte und eventuelle zusätzliche Attribute.

Funktionen, die eine Person hinsichtlich eines touristischen Objekts wahrnimmt (z.B. Gründer, Ansprechpartner) werden mit Hilfe der Klasse "has function" abgebildet.

Für die Darstellung von Beziehungen zwischen Medien, wurde die Klasse "is extensiontype" integriert.

Um die Ontologie möglichst flexibel zu gestalten unterstützt die Assoziationsklasse "content" Beziehungen zwischen beliebigen Konzepten indem als Typ der assoziierten Konzepte das Wur-

zelkonzept "Object" festgelegt wurde. Auf diese Weise können Beziehungen wie "Zug hält am Bahnhof" oder "Wasserfläche hat Hafen" realisiert werden.

## 4 Datenbankschema

Im Rahmen des Projektes wurde ein Datenbankmodell entwickelt, das die in Abschnitt 2 definierten Anforderungen erfüllt. Grundlage für das Modell war die in Abschnitt 3 beschriebene Ontologie. Ein wesentliches Problem des Modellierungsprozesses war die Abbildung der Ontologie auf ein physisches Datenmodell.

Die Konzepte der Ontologie bilden eine Vererbungsstruktur. Um Mechanismen wie "Vererbung" und "Implementierung von objektgebundenen Methoden" zu nutzen, war die Verwendung eines objektrelationalen Datenbankschemas sinnvoll. Aufgrund von bereits bestehenden Erfahrungen wird IBM Universal Database V8.2 (IBM UDB) als objektrelationales Datenbankmanagementsystem (DBMS) eingesetzt.

Da die Konzepte bereits die zu speichernden Objekte repräsentieren, können sie direkt in Entitäten des physischen Datenmodells umgewandelt werden. Dazu werden im DBMS benutzerdefinierte Datentypen erzeugt, deren Attribute die Merkmale des jeweiligen Objekts abbilden. Danach können auf Basis dieser Datentypen typisierte Tabellen generiert werden.

Alle Datentypen und typisierte Tabellen werden vom Supertyp "Object" abgeleitet. Die Typen "Person", "Dates", "TouristObject" und "Document" spiegeln die Hauptkonzepte der Ontologie wieder (siehe Abbildung 2). Der Datentyp "Component" wurde eingefügt, um Kollektionen von Dokumenten zusammenzufassen. Eine Kollektion stellt eine Menge von Dokumenten dar, die ein Objekt in einem bestimmten Kontext beschreiben. Mit Hilfe von Templates können Layouts für die Präsentation der Dokumente definiert werden. Dadurch ist es möglich die Darstellung von Kollektionen an unterschiedliche Ausgabemedien und verschiedene Geräte anzupassen.

Die oben beschriebenen Haupttypen werden im weiteren Modellierungsprozess in immer feiner granulierte Entitäten spezialisiert, die alle weiteren Konzepte der Ontologie repräsentieren. Um die Assoziationsklassen abzubilden, wurden zwei zusätzliche Entitäten – "RelationType" und "Relationship" – eingefügt. Jeder Datensatz der Tabelle "RelationType" enthält die Definition einer Assoziationsklasse. Die Entität "Relationship" speichert die eigentliche Beziehung. Integritätsprüfungen werden mit Hilfe von Triggern implementiert.

Diese Art der Implementierung ermöglicht die Definition von neuen Assoziationen, ohne Änderungen am Datenbankmodell vorzunehmen. Es wird lediglich ein neuer "RelationType" angelegt. Da alle Objekte vom Supertyp "Object" abgeleitet sind, können beliebige Objekte in Beziehung gesetzt werden. In einer Produktionsumgebung könnte sich diese Lösung jedoch nachteilig auf die Antwortzeiten der Datenbank auswirken, da die Tabelle "Relationship" dann eine sehr grosse Menge von Tupeln enthält.

Eine alternative Implementierungsmöglichkeit stellt das Erzeugen einer neuen Entität für jede einzelne Assoziationsklasse dar. Wahrscheinlich würde mit Hilfe dieser Strategie die höchste Performance erreicht, jedoch macht jede Modifikation der Assoziationsklassen eine Anpassung des Datenbankschemas notwendig.

Eine dritte Variante wäre die Definition von Relationship-Entitäten für jeden Haupttyp. Diese Mischlösung gewährleistet vermutlich einen guten Mittelweg zwischen Laufzeitverhalten und Wartungsaufwand.

Antworten auf diese Fragestellungen werden weitere Testimplementierungen, die für die Zukunft geplant sind, geben.

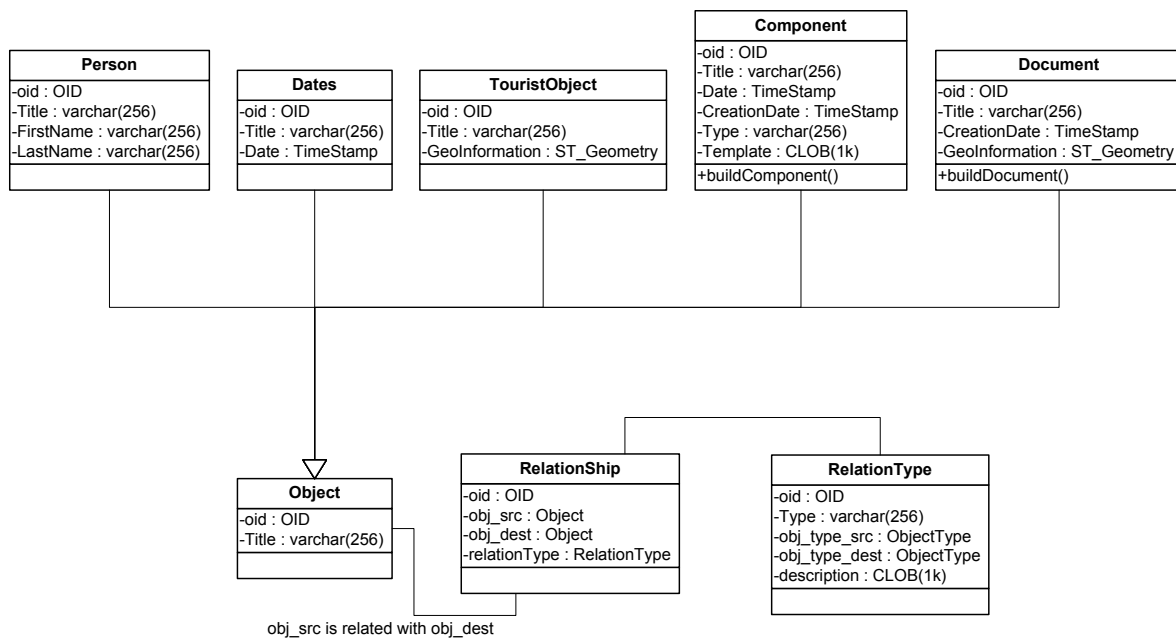


Abbildung 2: Auszug Datenbankschema

## 5 Zusammenfassung und Ausblick

FlyingPix ist ein ontologiegestütztes, geographisches Tourismusinformationssystem. Auf der Grundlage von Anforderungen aus der Sicht touristischer Dokumente, einer Tourismusontologie und georeferenzierten Objekten wurde ein Modell entwickelt, das den Touristen von der Routenplanung bis zur Urlaubsnachbereitung unterstützen kann. Verschiedene eher generische Modelle wurden auf ein physisches Datenbankschema abgebildet.

Zukünftig ist geplant, Strategien für flexible und effiziente Anfragen an die implementierte Ontologie zu entwickeln, Messungen des Laufzeitverhaltens durchzuführen und das Komponentenmodell zur Speicherung von Kollektionen multimedialer Dokumente für die Präsentation auf mobilen Geräten zu optimieren.

## Literatur

- [1] Ignatova, T., Bruder, I.: Utilizing a Multimedia UML Framework for an Image Database Application. Proceedings of the Best Practices Workshop, ER2005, (2005).
- [2] Mariella Gutiérrez, Andrea Rodríguez: Querying Heterogeneous Spatial Databases: Combining an Ontology with Similarity Functions. In: Conceptual Modeling for Advanced Application Domains: ER 2004 Workshops, (2004).
- [3] F. Fonseca, M. Egenhofer, P. Agouris, and C. Camara: Using Ontologies for Integrated Geographic Information Systems. In: Transactions in GIS, 6(3), (2002).
- [4] M. Klettke et. al.: GETESS - Ontologien, objektrelationale Datenbanken und Textanalyse als Bausteine einer semantischen Suchmaschine. In: Datenbank Spektrum – Zeitschrift für Datenbanktechnologie, 1/2001, (2001).