

Digitale Geodaten waren in der Vergangenheit einer recht kleinen Gemeinde von Spezialisten vorbehalten. Die Nutzer der ersten Stunde kamen aus den Kataster-Verwaltungen und der Energie-Versorgung. Sie erfassten und pflegten Kataster- und Leitungsdaten, dokumentierten somit die vorliegenden Verhältnisse und nutzten sie für Planungszwecke. Mittlerweile hat sich die Situation deutlich geändert. Digitale Geodaten werden heute vielfältig genutzt und sind Grundlage verschiedenartigster Auswertungen.

Koexistenz zwischen Smallworld GIS und Oracle: der Beginn eines Geodata Warehouse

Joachim Figura, CISS TDI GmbH

Zunehmend sind geobasierende Auswertungen in nahezu beliebigen Kombinationen aus den unterschiedlichsten Fachbereichen eines Unternehmens gefragt. Die Fachsysteme stehen dabei nicht zur Diskussion, vielmehr werden die Daten regelmäßig in weniger komplexen Datenmodellen, dafür aber verständlicher aufbereitet, in Kombination mit anderen Datenbeständen benötigt. Fachsysteme sind dafür in der Regel nicht ausgelegt. Es bedarf also anderer Lösungen, um diese vielfältigen Anforderungen bedienen zu können. Eine solche ist das Geodata Warehouse. Es ist sowohl für Fachabteilungen bei Nutzung von Daten aus anderen Bereichen geeignet als auch für Auswertungen auf Management-Ebene, in der häufig eine globale Sicht auf die Daten erforderlich ist.

Bei der Management Service Linz GmbH, der zentralen Dienstleistungseinheit des Ver- und Entsorgers LINZ AG in Österreich, wird aktuell der erste Schritt zum Aufbau des Geodata Warehouse umgesetzt. Hier lässt sich besonders gut anhand der aktuellen Momentaufnahme die Notwendigkeit einer übergreifenden Datennutzung darstellen, die in anderen Kundenprojekten bereits realisiert wurde.

Die Kundensituation

Die Management Service Linz GmbH (LINZ AG) betreibt als operatives System für Planung und Dokumentation das Smallworld GIS der Firma GE mit den Fachschalen Gas, Wasser, Kanal,

Fernwärme, Flurkarte und Naturbestand sowie einer Smallworld-Eigenentwicklung für Strom und Telekom. Die Aufgabenstellungen sind komplex und werden in der Fachabteilung von Spezialisten erledigt.

Das Datenbank-System Version Managed Data Store (VMDS) ist als Bestandteil der Smallworld-Plattform optimiert auf die GIS-Aufgaben der jeweiligen Fachanwendungen, jedoch nicht für den Zugriff durch „Jedermann“. Das führt leider dazu, dass auf flexible Anforderungen häufig auch nur sehr schwerfällig reagiert werden kann.

Die Management Service Linz GmbH betreibt neben dem Netz-Informationssystem Smallworld (NIS) aber auch noch ein weiteres (weborientiertes) GIS-System, den Geomedia Smart Client der Firma Intergraph mit Oracle Locator als Datenhaltungs-System, genutzt für die unternehmensweite NIS-Auskunft, eine spartenübergreifende Grabungskoordination und GeoTrams-Linien für den ÖPNV. Zum Betrieb dieser Systeme und für die Bearbeitung der speziellen Aufgabenstellungen sind jedoch auch die Daten aus dem Smallworld GIS von fundamentalem Interesse. Es wird somit eine abteilungs- und aufgabenübergreifende Lösung zur Nutzung der eingesetzten Geodaten gewünscht, die sich durch folgende Charakteristika auszeichnen soll:

- Offenheit und damit die Möglichkeit des einfachen Zugriffs auf die Daten ohne Spezialkenntnisse

- Einfach zu erstellende Auswertungen
- Flexibilität bei den Auswertungen
- Nutzung von besser geeigneten Werkzeugen für nicht fachabteilungsspezifische Aufgaben
- Nutzung/Integration fremder Daten in der Smallworld-Umgebung – gegebenenfalls auch nur temporär
- Integration von GI-Systemen unterschiedlicher Hersteller mit Nutzung von Standards und Weiternutzung der jeweiligen Stärken der Systeme

Möglichkeit eines Oracle-basierenden Geodata Warehouse

Abstrahiert man die Anforderungen, so finden sich schnell Parallelen zur bewährten Data-Warehouse-Philosophie, jedoch mit der Besonderheit der hier zugrunde liegenden Geodaten. Nach gängiger Meinung lässt sich ein Data Warehouse wie folgt charakterisieren:

- Ein Data Warehouse ermöglicht die Integration von Daten aus verteilten und unterschiedlich strukturierten Beständen, die Konsolidierung, Veredelung und Qualitätssicherung der Daten im Rahmen des Extract-Transform-Load-Prozesses (ETL).
- Ein Warehouse trennt die Auswertungen von der Produktion. Es ermöglicht die übergreifende Nutzung von Datenbeständen, etwa für Analysen, Reports etc. zur Unterstützung von (Geschäfts-)Entscheidungen.

- Eine abteilungsübergreifende, globale Sicht auf die Daten ist daher mithilfe eines Warehouse möglich. Zur Wahrung der Übersichtlichkeit können fachliche Details je nach Nutzer-Anforderungen ein- oder ausgeblendet werden.

Überträgt man dies auf die Wünsche und Anforderungen der Management Service Linz GmbH, liegt der Mehrwert, den ein Data Warehouse, hier auf Basis von Oracle und ergänzt um die Geo-Komponente, bringt, auf der Hand:

- Die Daten werden durch die offene Datenhaltung und den OGC-Standard „Simple Feature Option“ für weitere GIS direkt und ohne weitere Konvertierung nutzbar, also insbesondere auch für Geomedia Smart-Client
- Daten können ohne Smallworld-Expertenwissen mit Standards wie SQL nutzbar gemacht werden
- Flexible Analysen, Reports, Simulationen, Data Mining und Views in Zusammenhang mit anderen Daten sind der Normalfall
- Geodaten heterogener Quellen sind in Oracle (ad hoc) nutzbar

Konzepte der Datenübernahme in das Warehouse

Eine Variante der Datenübernahme ist die „1:1“-Abbildung der Datenmodelle aus Smallworld in Oracle mit vollständiger Synchronisation der Datenbestände und Sicherstellung der Datenintegrität oder alternativ die Speicherung der für die Nachnutzung bereits aufbereiteten Daten in einem Geodata Warehouse. Die „1:1“-Abbildung weist trotz der vollständigen Abbildung der Daten und des Datenmodells in der Datenbank einige Nachteile auf. So sind beispielsweise die technischen Strukturen der Systeme unterschiedlich, sodass die Daten trotzdem modelliert werden müssen. Die Netz-Topologie, also die Speicherung der Zusammenhänge und der Nachbarschaftsbeziehungen, erfolgt im Smallworld zum Beispiel geometrieorientiert, das heißt das System erkennt die Netz-Zusammenhänge anhand der geometrischen Lage auf dem Kartenbild. Oracle hingegen verlangt

eine Knoten-Kanten-Orientierung, die Topologie wird demnach explizit im Datenmodell abgebildet. Es gibt darüber hinaus weitere Beispiele, die eine Datenmodellierung notwendig machen.

Die eigentliche Problemstellung liegt aber in der Komplexität des operativen Smallworld- Datenmodells, das in dieser Form für eine direkte Nachnutzung nicht geeignet ist. In der Regel bauen daher nachgeordnete Auswertungen auf vereinfachten, sehr ähnlichen Datenmodellen auf. Zudem muss die Präsentation, also das Erscheinungsbild der Karte, verständlich aufbereitet werden. Die direkte Aufbereitung zum Zeitpunkt der Auswertung ist jedoch mit Performance-Verlusten verbunden. Somit lohnt sich die Vorprozessierung der Daten analog zu einem klassischen Warehouse in einem ETL-Prozess und die Ablage der Daten in der Oracle-Datenbank.

Der ETL-Prozess mit Geodaten kann aufgrund der Anforderungen an Datenmodell-Änderungen, der Aufbereitung der Präsentation, der regelmäßig großen Datenmengen und der folglich damit verbundenen Forderung nach einer hohen Performance durchaus sehr aufwändig sein, sodass umfangreiche Erfahrungen im Aufbau eines Geodata Warehouse ausschlaggebend für die erfolgreiche Umsetzung sein können. Aufgrund der Komplexität der beteiligten Systeme ist zusätzlich eine detaillierte Kenntnis aller beteiligten GIS-Technologien und der zugehörigen Fachanwendungen für den Projekterfolg von entscheidender Bedeutung.

Die CISS TDI GmbH hat sich auf Projekt-Dienstleistungen im Bereich der Geodaten-Integration spezialisiert. Die Geo-ETL Software CITRA ist ein Lizenzprodukt der Firma. Die Anfänge gehen auf das Jahr 1987 zurück. Die Software wurde seitdem laufend weiterentwickelt und ermöglicht die Modellierung hochkomplexer Geodaten-Bestände.

Die Management Service Linz GmbH hat zum Aufbau des Geodata Warehouse auf CISS TDI und deren CITRA-Produktfamilie gesetzt, da hier alles, sowohl die Dienstleistung als auch die Bereitstellung der ETL-Standardsoftware, aus einer Hand bereitgestellt

werden konnte. „CISS TDI kann eine exzellente Expertise in Smallworld, Intergraph-Produkten und Oracle sowie vielfältige Projekterfahrung mit den beteiligten Systemen vorweisen“, so der Projektleiter Manfred Kurzwernhart von der Management Service Linz GmbH.

Zwar verwendet der Kunde auch andere Geo-ETL-Software, aber die Komplexität der Aufgabenstellung verlangte hohe Erfahrung und eine kompetente Dienstleistung aus einem Haus. So ist die tägliche Differenzdaten-Fortführung mittels CITRA im Warehouse ein wesentliches Merkmal der hohen Performance. Beim Differenzdaten-Update besteht die Schwierigkeit darin, dass immer nur Teile aus einem größeren Zusammenhang aktualisiert werden, es jedoch sichergestellt werden muss, dass der Gesamtzusammenhang erhalten bleibt. Ebenso konnte auch die Übernahme von Smallworld-internen Rasterdaten mit korrekter Geo-Referenzierung realisiert werden.

Obwohl zukünftig noch viele weitere (Geo-) Datenbestände in das Warehouse integriert werden können/sollen, konnte mit dem jetzigen Zustand bereits ein großer Nutzen erreicht werden. Der weitere „Step-by-Step“-Ausbau des Geodata Warehouse ist durch den modularen Aufbau gewährleistet, sodass auch in Zukunft bei überschaubaren Risiken, sowohl in projektplanerischer als auch finanzieller Hinsicht, jeweils ein unmittelbarer Nutzen durch auch nur kleine Ergänzungen erzielt werden kann. Diese Strategie eines nachhaltigen und zukunftssicheren Basis-Konzepts mit der Möglichkeit des schrittweisen Ausbaus ist bei vielen bisherigen Kunden und in Linz ein wichtiger Hintergrund der Projektumsetzung mit CISS TDI.



Joachim Figura
ciss@ciss.de