

Das MeteoSchweiz Data Warehouse System

Milliarden von Datensätzen werden im Data Warehouse System der MeteoSchweiz erfasst, gespeichert und für die weitere Nutzung aufbereitet.

MeteoSchweiz ist der nationale Wetterdienst für die Schweiz. Eine seiner Hauptaufgaben ist es, meteorologische und klimatologische Daten zu erfassen, aufzubereiten, bereitzustellen und zu archivieren. Milliarden von Datensätzen sind bereits im Data Warehouse (DWH) gespeichert. Täglich kommen über 10 Mio. Messwerte aus einer Vielzahl von Quellen in unterschiedlichen Formaten hinzu.

Um dieser anspruchsvollen Aufgabe langfristig effizient gerecht zu werden, betreibt MeteoSchweiz ein so genanntes Data Warehouse System. Es beinhaltet verschiedene Datenbanken zur Datenaufbereitung, eine zentrale Datenbank, die für Analyse- und Informationszwecke optimiert ist, sowie ein zentrales Metadaten Repository. Die leistungsfähige und erweiterbare Infrastruktur umfasst weitere Werkzeuge für Aggregation, Qualitätskontrolle und Analyse der meteorologischen und klimatologischen Daten.



Wartung der MeteoSchweiz Messstation auf dem Säntis. Wohin mit den Daten?

Massgeschneiderte Datenpakete für die Kunden

Das Data Warehouse System bietet massgeschneiderten Datenzugriff für interne und externe Benutzer. Dabei können sowohl aktuelle Daten als auch historische Zeitreihen bezogen werden. Der grösste Teil der meteorologischen und klimatologischen Produkte von MeteoSchweiz werden mit Daten aus dem DWH erstellt.

10 Jahre Entwicklung

1999 hat die Geschäftsleitung der MeteoSchweiz entschieden, die bestehende Datenbanklandschaft neu zu organisieren. Dabei war die wesentliche Grundlage eine Expertise des Department of Computer Science der ETH Zürich und des Institutes für Informatik von der Universität Zürich. Die neue Infrastruktur wurde stufenweise in Betrieb genommen und ersetzte so ein Vorgängersystem nach dem andern. Die Anforderungen aus allen Aktivitätsfeldern von MeteoSchweiz wurden in sieben Releases während zehn Jahren umgesetzt.

Das Nachfolgeprojekt RENAISSANCE modernisiert die auf dem Data Warehouse basierenden Produktionsdienste und baut die Prozesse für die Datenversorgung nachgelagerter Systeme aus.

Architektur Data Warehouse System von MeteoSchweiz

Im DWH System der MeteoSchweiz werden Elemente klassischer relationaler Datenbanken kombiniert mit klassischen Data Warehouse-Konzepten.

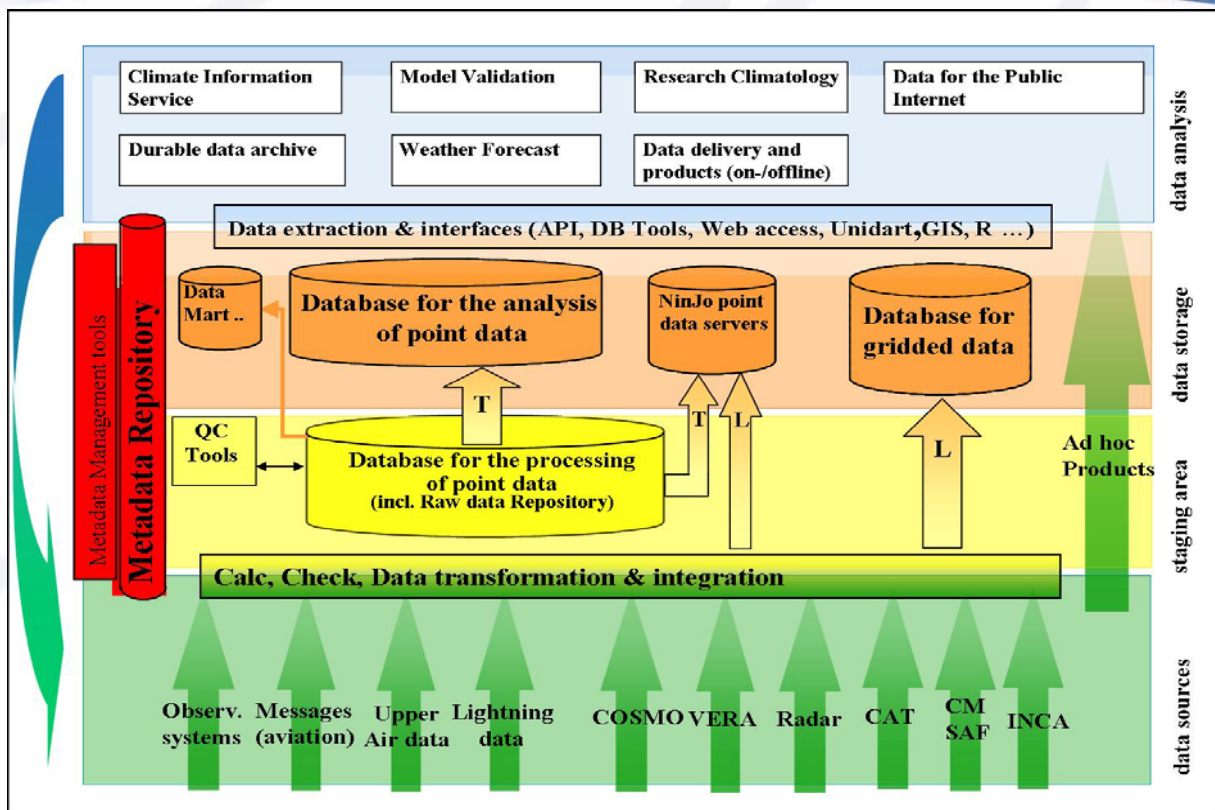
Das DWH System ist in vier Schichten gegliedert (Abbildung auf der Rückseite) und dient als unternehmensweite Integrationsplattform für alle Fachdatenbestände von MeteoSchweiz.

Das zentrale Element des Systems ist das Metadaten Repository. Es enthält alle Informationen zur Beschreibung der eigentlichen Messwerte, wie z.B. Stationskoordinaten, Messprogramme oder Detailinformationen über Messgeräte aber auch alle Informationen zur Steuerung der DWH-internen Prozesse. Die vier Schichten des DWH System sind im Metadaten Repository verankert. Das Repository wird auch von um das DWH angesiedelten Systemen verwendet



Riesige Datenmengen entstehen z.B. bei einer Ballonsondierung.

Die Funktion der vier Schichten kann wie folgt beschrieben werden:



Vier-Schichten-Architektur des Data Warehouse Systems der MeteoSchweiz

Quellschicht: Die ins DWH fließenden Daten stammen aus einer Vielzahl von Quellen. Dazu gehören Beobachtungssysteme (Bodenmessnetze, biometeorologische Messnetze, Radiosondierungen), numerische Wettervorhersagemodelle, das Global Telecommunications System (GTS) der WMO, Wetterradare und Satelliten, Datenarchive mit historischen Datenreihen und Daten aus Beobachtungssystemen von Partnermessnetzen.

Datenaufbereitungsschicht: Die Aufbereitungsschicht (staging area) beinhaltet die Arbeitsdatenbanken und sämtliche Methoden und Werkzeuge, die für das Erfassen, Aufbereiten, Laden sowie die Qualitätskontrolle und Bereinigung der

meteorologischen Daten benötigt werden.

Datenspeicherschicht: In der Datenspeicherschicht befindet sich das „core warehouse“ von MeteoSchweiz. Dabei handelt es sich um eine analytische Datenbank, die den Meteorologen und Klimatologen ein integriertes Set an Daten zur Verfügung stellt, um sie in der Wettervorhersage und Klimaanalyse zu unterstützen. Mit so genannten Data Marts, die ausgewählte Teilbereiche des Data Warehouse enthalten, stehen Datenbestände bereit, die auf die spezifischen Anforderungen bestimmter Anwendergruppen zugeschnitten sind.

Datenanalyseschicht: Die Datenanalytischschicht umfasst alle Anwendungen für Abfragen und Analysen auf das Data Warehouse bzw. die Data Marts. So können z.B. mit der Applikation CLIMAP Messwerte in Form von Tabellen, Karten oder Verlaufsgrafiken schnell und übersichtlich angezeigt werden. Zugang zu den Daten haben sowohl MeteoSchweiz-interne Anwender als auch interessierte Institutionen und Privatpersonen mittels flexibler und individuell konfigurierbarer Software wie CLIMAP-net oder IDAWEB. Ausserdem werden die Daten des DWH in den Produkten der MeteoSchweiz weiterverarbeitet, wie sie zum Beispiel im Internet angeboten werden.

06/2011 © MeteoSchweiz