

Array-Datenbanken für semantische inhaltsbasierte Suche und Analyse in Satellitenbildarchiven

MARTIN SUDMANN¹, DIRK TIEDE¹, NIKOLAUS AUGSTEN²,
ANDREA BARALDI^{1/3}, MARIANA BELGIU¹ & STEFAN LANG¹

Zusammenfassung: Die steigende räumliche und zeitliche Auflösung von Satellitenbildern birgt Herausforderungen beim Durchsuchen und Analysieren in Satellitenbildarchiven dar. Semantische inhaltsbasierte Suchmöglichkeiten und web-basierte Analysen in der Datenbank unterstützen den Umgang mit der steigenden Datenmenge. Dieser Beitrag stellt eine Methode und Implementierung vor, wie Bilddatenbanken durch semantische Abfragen besser in Wert gesetzt werden. Dies basiert auf der Anreicherung der Satellitenbilder mit vollautomatisch vorverarbeiteten semantischen Layern. Dies beinhaltet (i) eine Basisarchitektur mit einer Array-Datenbank und einem web-basierten grafischen Abfragesystem, (ii) ein Vergleich von Datenbanken zur Speicherung von Rasterdaten und (iii) die Implementierung in einem Prototyp (IQ – ImageQuerying). Die Idee und der Prototyp, die mit dem Copernicus Masters 2015 - T-Systems Big Data Challenge ausgezeichnet wurden, sollen dazu beitragen, Satellitenbilder effektiver nutzen zu können.

1 Einleitung

Aufgrund der stetig steigenden räumlichen und zeitlichen Auflösungen von Satellitenbildern und der damit einhergehenden Datenmengen ist die Erdbeobachtung mit zunehmenden Herausforderungen bei der Auswertung dieser Daten konfrontiert. Schon heute können die existierenden Methoden und technischen Möglichkeiten mit dieser Entwicklung kaum Schritt halten, z.B. wird nur ein geringer Prozentsatz aller Satellitenbilder der ESA-Archive heruntergeladen und ausgewertet (BARALDI & BOSCHETTI 2012). In diesem Zusammenhang wird auch der Begriff *dark data* (PURSS et al. 2015) für Daten verwendet, die zwar erhoben und gespeichert, aber niemals genutzt werden. Ein Grund, der hierfür genannt wird ist, dass keine geeigneten Werkzeuge existieren, um mit der Datenmenge intelligent umzugehen und deshalb die für bestimmte Analysezwecke richtigen Satellitenbilder oft gar nicht gezielt gesucht und gefunden werden können, um sie auszuwerten und Redundanzen zu vermeiden (BARALDI & BOSCHETTI 2012).

Ein Beispiel hierfür sind die Datenmengen der neuen Sentinel Satelliten, die im Jahr 2016 geschätzt 5.000 Terabyte an neuen Daten produzieren werden - mit steigender Tendenz (RECK et al. 2016). Die Auswertung dieser Datenmengen, die Selektion von geeigneten Bildern für bestimmte Anwendungszwecke und die Transformation in zuverlässige Informationsprodukte

¹ Paris-Lodron Universität Salzburg, IFFB Geoinformatik, Schillerstraße 30, A-5020 Salzburg, E-Mail: [martin.sudmanns, dirk.tiede, mariana.belgiu, stefan.lang]@sbg.ac.at

² Paris-Lodron Universität Salzburg, Fachbereich Computerwissenschaften, Jacob-Haringer Straße 2, A-5020 Salzburg, E-Mail: nikolaus.augsten@sbg.ac.at

³ Department of Agricultural and Food Sciences, University of Naples Federico II, Italy, E-Mail: andrea6311@gmail.com

stellen ein mittlerweile anerkanntes Problem dar (RECK et al. 2016; SOILLE et al. 2016). Betrachtet man Kosten und Aufwand für Satellitenmissionen und die Tatsache, dass derzeit viel Potential ungenutzt bleibt, ist es aus wissenschaftlicher, gesellschaftlicher und ökonomischer Sicht wichtig, Satellitenbilder effizienter zu nutzen.

Dieser Beitrag beschreibt eine Möglichkeit zur Aufwertung existierender Fernerkundungs-Bilddatenbanken, welche um semantische, inhaltsbasierte Abfrage- und Analysemöglichkeiten erweitert werden. Die erste Herausforderung hierbei ist, Computeralgorithmen so intelligent zu gestalten, damit diese die Inhalte der Bilder erkennen und interpretieren können. Jedoch bleiben computerbasierte Möglichkeiten nach wie vor hinter den Fähigkeiten eines menschlichen Betrachters zurück. Diese Diskrepanz wird als Semantische Lücke (*semantic gap*) bezeichnet. Dadurch sind die Erwartungen der Nutzer/-innen an ein Bildabfragesystem schwer zu erfüllen. Um diese Lücke zu verkleinern, basiert das vorgestellte System auf einer hybriden Methode, die sowohl eine vollautomatische Vorklassifikation der Daten basierend auf physikalischen Modellen (a-priori Wissen) als auch das von Benutzer/-innen als Abfrage eingebrachte Wissen nutzt. Das externe Wissen wird über eine web-basierte grafische Benutzeroberfläche (graphical user interface, GUI) in das System eingebracht, in Datenbankabfragen transformiert und abgespeichert. Dadurch können die Abfragen mit anderen Benutzer/-inne/n geteilt werden. Es ist daher ein inkrementell wachsendes System, das durch die Benutzung zunehmend mehr Möglichkeiten bietet. Die zweite Herausforderung ist verbunden mit den großen Datenmengen, da die klassische Herangehensweisen zur Speicherung von Satellitenbildern für eine datenbankbasierte semantische Suche und Auswertung von Satellitenbildern ungeeignet und ineffizient ist. Alternativ dazu werden die Daten in dem vorgestellten System in einem dreidimensionalen Datenmodell in einer Array-Datenbank gespeichert, um Abfragen durch Raum und Zeit effizient durchführen zu können.

2 Motivation und Stand der Technik

Zur Illustration soll ein gängiges Beispiel dienen: Ein/e Benutzer/in ist auf der Suche nach allen für ihre/seine Anwendung qualifizierten Satellitenbildern für ein ausgewähltes Gebiet (area-of-interest, AOI). Die Bedingung soll sein, dass die Aufnahmen (über diesem Gebiet) wolkenfrei sind. Hierfür muss der/die Benutzer/in in den meisten Fällen zunächst alle Bilder herunterladen (wenn die Datenvorschau nicht ausreicht) und anschließend die unbrauchbaren Aufnahmen manuell aussortieren. Selbst wenn eine Wolkenbedeckung als Qualitäts- und Suchkriterium angegeben werden kann, ist diese für das gesamte Bild global ausgewiesen. Es ist hingegen durchaus möglich, dass ein Bild insgesamt viele Wolken hat, aber das von dem/der Benutzer/in ausgewählte Gebiet wolkenfrei ist – oder umgekehrt. Ein System, das während der Abfrage und basierend auf dem ausgewählten Gebiet die Wolkenbedeckung berechnet und damit die qualifizierten Satellitenbilder automatisch findet, würde dem/der Benutzer/in in diesem Fall viel Arbeit ersparen können oder gewisse Abfragen überhaupt erst möglich machen. Dieses Beispiel lässt sich beliebig komplex erweitern, in dem anstatt Wolken eine Landnutzungs-kategorie oder eine Kombination davon abgefragt wird.

Diese Art der Fragestellung wird allgemein als inhaltsbasierte Suche (content-based image retrieval, CBIR) bezeichnet und gewinnt im Zeitalter von big data, in dem eine manuelle

Auswertung fehlerhaft, langwierig oder gar ganz unmöglich ist, immer mehr an Bedeutung. Derzeitige Bilddatenbanken erlauben jedoch keine inhaltsbasierte Abfragen für benutzerdefinierte Gebiete. Sowohl beispielsweise im USGS Earth Explorer Portal (<http://earthexplorer.usgs.gov/>) als auch im Sentinels Scientific Data Hub (<https://scihub.copernicus.eu/>) sind Abfragen über die geographische Position der Bilder, die Aufnahmezeit und textuelle Metadaten vorgesehen. Es wäre zwar in dem Beispiel möglich, ein spezielles Computerprogramm zu entwickeln, das die heruntergeladenen Bilder nach den gewünschten Kriterien aussortiert. Allerdings ist die Entwicklung eines in diesem Falle sensor- und abfrageabhängigen Programms umständlich, nicht übertragbar und auch nicht für große Datenmengen geeignet. Der derzeitige Ansatz, die Daten herunterzuladen und lokal zu verarbeiten ist nur zu einer begrenzten Datenmenge technisch überhaupt möglich. Für große Untersuchungsgebiete und multitemporale Fragestellungen können bald Terrabytes oder möglicherweise sogar Petabytes an Daten anfallen, die über das Internet gesendet werden müssten. Auch mit steigenden Übertragungsgeschwindigkeiten bleibt es effizienter, nicht die Daten zu den Benutzer/-innen zu transportieren, sondern den Algorithmus zu den Daten. Bestehenden Ansätzen für Online-Analysen auf der Datenbank-Ebene wie dem EarthServer (BAUMANN et al. 2015) oder dem Australian Geoscience Data Cube (AGDC) (PURSS et. al 2015) fehlt die Möglichkeit, beliebige benutzerdefinierte Abfragen und Analysen auf einem höheren semantischen Level durchzuführen.

Ein System, das semantische, inhaltsbasierte Abfragen und Analysen über das Internet möglich macht ist ein neuer Ansatz, um Bilddatenbanken zu verbessern und Satellitenbilder effizienter zu nutzen. Dafür gibt es mindestens zwei Anwendungsgebiete, die im Zusammenhang stehen: (i) inhaltsbasierte Suche als Erweiterung herkömmlicher Suchen und (ii) eine web-basierte Online-Analyse, bei der das Ergebnis entweder ein neuer, georeferenzierter Datensatz (Analyseergebnis) ist mit räumlich-zeitlicher Auswertung (z.B. *Wo hat sich eine Landnutzung in den letzten 5 Jahren verändert?*) oder einer zeitlichen Auswertung (*In welchem Jahr gab es die meisten Änderungen?*).

3 Methodologische Grundlagen

Um Satellitenbilder inhaltlich abfragbar zu machen, müssen diese semantisch durch einen oder mehrere semantische Layer angereichert werden, die den Inhalt auf einem sub-symbolischen oder symbolisch-kategorischen Level beschreiben. Diese können beispielsweise eine Landnutzungs-klassifikation oder auch ein NDVI beinhalten. Um den Ansatz für big data Fragestellungen nutzbar zu machen ist allerdings eine wesentliche Bedingung, die semantischen Layer parameterfrei und vollautomatisch zu erstellen. Für das vorgestellte System wurde die vollautomatische Vorklassifikation nach BARALDI et al. (2010) verwendet. Da die semantischen Layer direkt mit den Sensordaten verknüpft sind, sind inhaltsbasierte Abfragen darauf möglich. Die semantischen Layer bilden dann die Grundbausteine für diese Abfragen. Die Begriffe ‚semantischer Layer‘ und ‚Informationslayer bzw. –ebene‘ werden in diesem Beitrag synonym verwendet.

Die Satellitenbilder und deren assoziierte semantische Layer müssen an einer zentralen Stelle in einem Datenmodell gespeichert werden, das beliebig komplexe Abfragen in akzeptablen Zeiten ermöglicht. Auch wenn derzeit Fernerkundungsdaten als zweidimensionale Bilder direkt im Dateisystem verwaltet werden, bringt die Speicherung in Datenbanken mehrere Vorteile. Die Abfragen erfolgen über eine deklarative Abfragesprache, mit der die Datenbank für jede Abfrage

über die beste Zugriffsstrategie entscheiden und interne Optimierungen durchführen kann (BAUMANN & HOLSTEN 2011). Zusätzlich können ein optimiertes Datenmodell und verschiedene Indexierungsmethoden bestimmte Zugriffsmuster auf die Daten unterstützen. Beispiele sind Zeitreihenanalysen oder wenn das von den Benutzer/-innen gewählte Gebiet mehr als ein Bild oder nur Teile eines Bildes umfasst. Klassische Vorteile von Datenbanken wie erhöhte Sicherheit, Transaktionsmanagement oder bessere Backupstrategien können ebenfalls von Relevanz sein (BRINKHOFF & KRESSE 2011). Die strikte Trennung von logischem Modell (Abfragesprache) und physischem Modell (z.B. Dateiformate, Indexstrukturen, Algorithmen für den Datenzugriff) in Datenbanksystemen erleichtert die Wartbarkeit der Anwendungen, und ermöglicht die Optimierung von Anfragen getrennt von Anwendungsfällen zu betrachten. Verbesserungen im Datenbanksystem führen zu effizienteren Anfragen, ohne dass die Anwendung selbst geändert werden muss. Aus diesen Gründen wird für das vorgeschlagene System eine Array-Datenbank zur Speicherung der Bilder verwendet.

Eine web-basierte grafische Benutzeroberfläche dient als Schnittstelle zwischen den Benutzer/-innen und der Datenbank. Diese ist dabei so konzipiert, dass das Wissen der Benutzer/-innen in maschinenlesbarer Weise in die Wissensbasis und damit in das System codiert und eingebracht werden kann. Hierbei muss berücksichtigt werden, dass semantische, raum-zeitliche Abfragen über große Komplexität verfügen können und daher oft schwierig zu erstellen sind. Die GUI muss daher Möglichkeiten bieten, die Abfragen unterstützt und übersichtlich zu erstellen, diese zu speichern und zu teilen und damit auch anderen Benutzer/-innen zur Verfügung zu stellen.

Für diesen Anwendungszweck wurden vier mögliche Datenbanken zur Speicherung von Rasterdaten untersucht: (1) Oracle Spatial and Graph als Erweiterung der Oracle Datenbank (ORACLE 2016), (2) PostGIS als Erweiterung von PostgreSQL (POSTGIS 2016), (3) Rasdaman (RASDAMAN 2016) und (4) SciDB (SCIDB 2016). Diese lassen sich weiter unterteilen in Datenbanken, die auf dem relationalen Modell aufbauen (1,2) und Datenbanken, die über ein eigenes Array-Datenmodell verfügen (3,4).

Basierend auf den Besonderheiten von Satellitenbildern sowie in der Literatur vorgeschlagenen Datenbankvergleichen (BAUMANN & STRAMERJOHANNIS 2014), wurde eine zugeschnittene Testumgebung geschaffen, die sowohl eine qualitative Analyse als auch einen Performance-Test umfasst. Das Ziel der qualitativen Analyse soll als Entscheidungsgrundlage für oder gegen die Benutzung einer jeweiligen Datenbank für den vorgestellten Anwendungszweck dienen, die Performance-Tests sollen ein Eindruck vermitteln, wie sich die Datenbanken insgesamt verhalten und wo Stärken oder Schwächen liegen.

Auf Grundlage der Datenbankanalyse wurde schließlich eine Kombination der beschriebenen Komponenten in einer neuen, generischen Basisarchitektur und als Nachweis der Machbarkeit in einem Prototyp implementiert.

4 Ergebnisse

Für die Fragestellung sind vier Ergebnisse gesondert zu betrachten. Diese sind (i) ein Vergleich der vier Datenbanken, (ii) die Basisarchitektur des neuen Systems, (iii) das web-basierte Abfragesystem und (iv) der Prototyp des Systems zur semantischen, inhaltsbasierten Suche und Analyse in Bilddatenbanken (IQ – ImageQuerying).

4.1 Evaluation der Datenbanken

Für den vorliegenden Anwendungsfall haben die Datenbanken mit Array-Datenmodell vor allem bei Abfragen durch die Zeit besser abgeschnitten. Das Array-Datenmodell unterscheidet sich von Bildern als Dateien hauptsächlich dadurch, dass die Daten nicht nur auf Basis der räumlichen, sondern auch der zeitlichen Dimension als Würfel angeordnet sind (Abb. 1). Durch die Gleichberechtigung der Dimensionen sind Abfragen durch die Zeit signifikant beschleunigt. Zusätzliche wichtige Elemente sind die deklarative Abfragesprache ähnlich SQL und das intelligente Partitionieren und Indexieren, das es der Datenbank ermöglicht, nur auf die Daten zuzugreifen, die für eine bestimmte Abfrage wirklich benötigt werden.

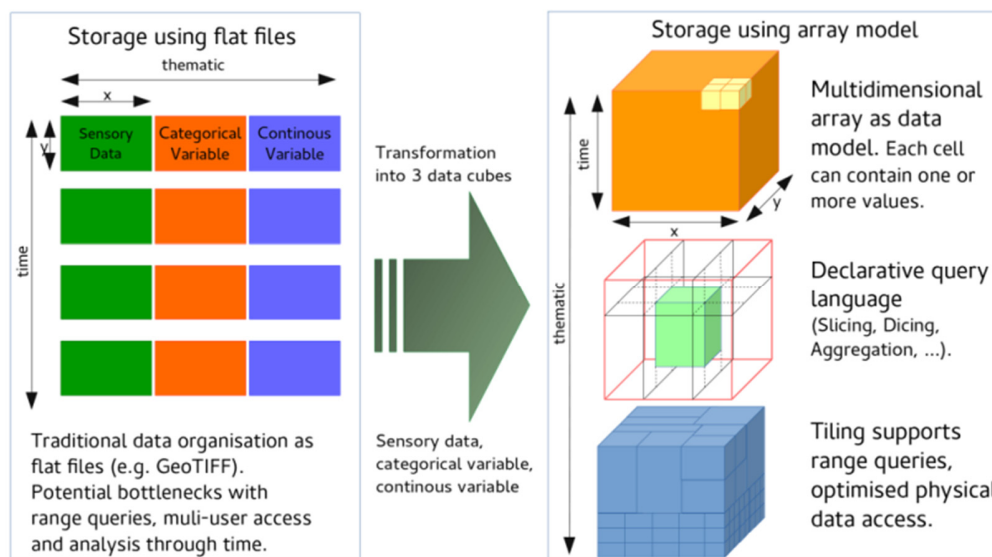


Abb. 1: Eigenschaften des Array-Datenmodells. Eigene Grafik, 2016.

Die experimentellen Ergebnisse haben gezeigt, dass vor allem die Analysen - je nach Art - wesentlich effizienter sein können als mit herkömmlichen Speichermethoden. Hingegen könnten sich die Zeiten für Import und Export in die Datenwürfel negativ auswirken. Es kann daher ratsam sein, die Daten im Vorfeld zu klassifizieren und entweder in import/export-optimierte oder analyse-optimierte Datenmodelle zu speichern. Für den hier genannten Anwendungsfall der inhaltsbasierten Suche ist dies machbar: Für die Analysen können die semantischen Layer in Datenwürfel genutzt, für den Download können die originalen Satellitenbilder genutzt werden. Darüber hinaus konnten Entwicklungs- und Erweiterungsgebiete im Datenbankbereich identifiziert werden, die zwar in klassischen Datenbanken vorhanden sind, aber im betrachteten Anwendungsfall von Array-Datenbanken nicht oder nur mangelhaft unterstützt werden. Dazu gehören erweiterte Elemente für Abfragesprachen, die auch räumlich-zeitliche Referenzierungen innerhalb des Datenwürfels erlauben (z.B. durch benutzerdefinierte Funktionen (user-defined functions, UDF) und benutzerdefinierte Datentypen), Indexierung der semantischen Layer (z.B. B+ tree, Hash Index, R-tree index, Bitmap index) oder Hinweise zur allgemeinen Optimierung (z.B. der Tiling-Strategie oder Pyramid-Layers) und Abfrageplanung.

4.2 Web-basiertes Abfragesystem

Die GUI des Abfragesystems ist eine in JavaScript programmierte Web-Applikation, die sich mit einem Standard-Browser bedienen lässt. Zentrales Element ist das Abfragepanel. In diesem wird die Abfrage sequentiell aufgebaut durch die Kombination von Untergruppen wie die Auswahl des Gebiets (AOI) und der Zeit sowie Kriterien für inhaltsbasierte Abfragen.

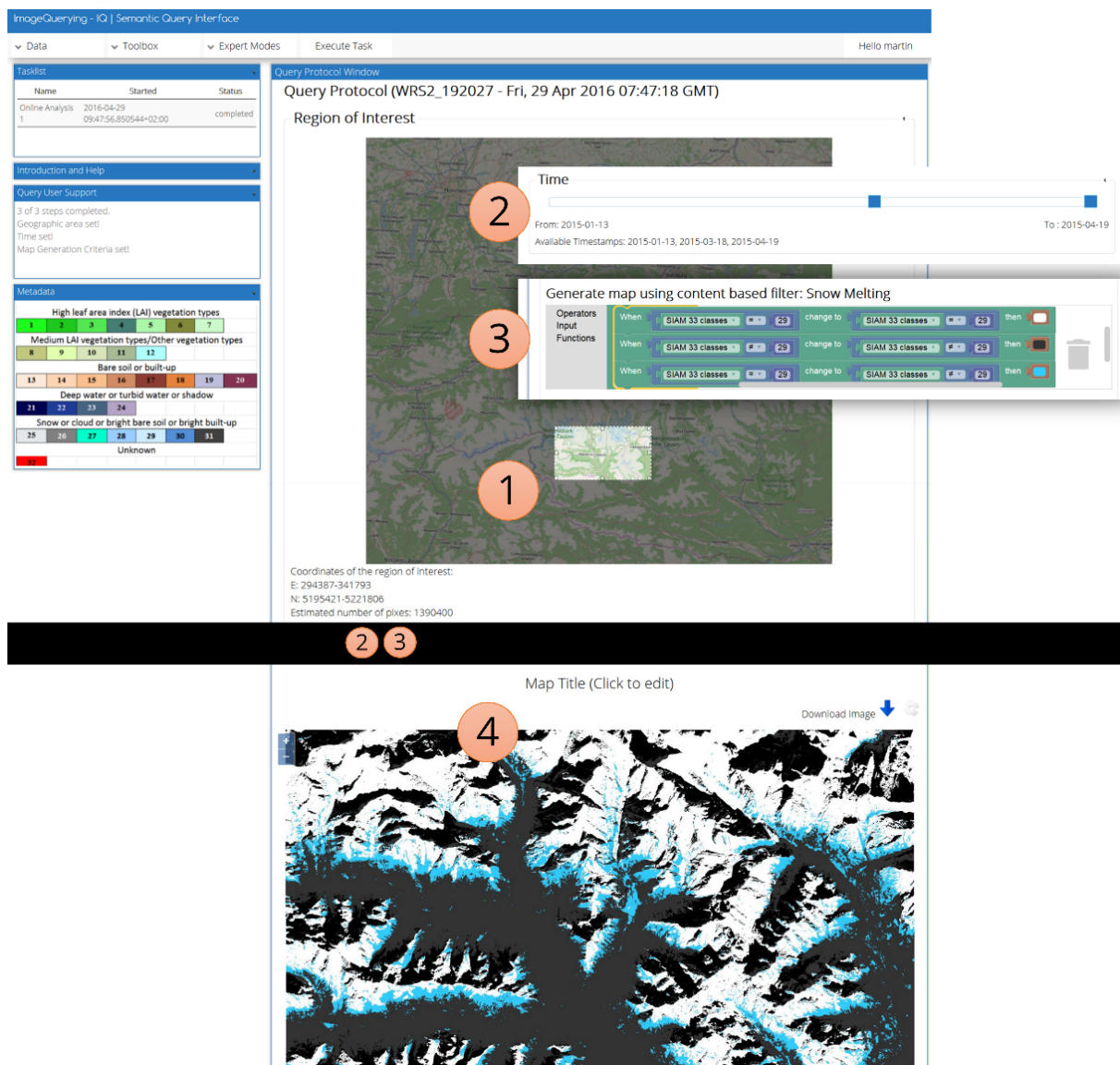


Abb. 2: Das web-basierte Abfragesystem (IQ-Frontend, Stand 29.04.2016) erlaubt es dem Benutzer, die Datenbank grafisch abzufragen. Hierfür selektiert der Benutzer im Abfragepanel zunächst die Area-of-Interest (1), den Zeitbereich (2) und erstellt die Regel zur inhaltsbasierten Abfrage mit grafischen Elementen (3). Nach der Ausführung erscheint das Ergebnis unterhalb der Eingabeparameter. Das vorliegende Beispiel zeigt im Nationalpark Hohe Tauern die Bereiche (blau) mit einer Abnahme der Schneebedeckung zwischen Januar und April 2015 und Bereiche (weiß) mit einer konstanten Schneebedeckung (verändert nach TIEDE et al. 2016).

Formparameter berechnet, die im Datenwürfel für kontinuierliche Variablen gespeichert sind. Eine dritte kontinuierliche Variable ist der *Greenness*-Index für Vegetationsanalysen.

Benutzer/-innen können die Datenbank mit vorgefertigten grafischen Elementen abfragen, die sich wie Puzzle-Teile zu beliebig komplexen Regelsätzen zusammenfassen und kombinieren lassen. Alle Komponenten des Systems sind auf eine Skalierung in den big data Bereich ausgerichtet.

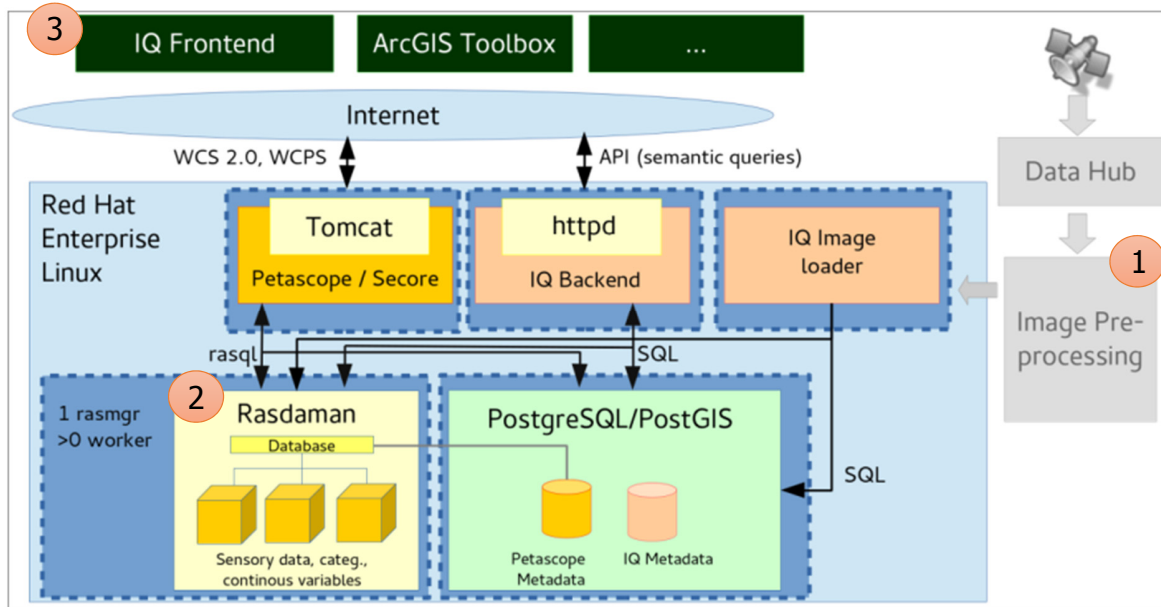


Abb. 4: Implementierung als ImageQuerying (IQ). Eigene Grafik, 2016.

5 Schlussfolgerungen und Ausblick

Die vorgestellte Methode und Implementierung soll Satellitendaten effizienter für Benutzer nutzbar machen. Die Frage, wie man mit den großen Datenmengen, die von den derzeitigen und geplanten Satelliten aufgenommen werden, effizient umgeht, ist relevant für die verbesserte Nutzung und das Ausschöpfen des Potentials von Satellitenbildern. Dieser Beitrag hat eine mögliche Lösung aufgezeigt, die daraus besteht, existierende Bilddatenbanken durch zusätzliche Layer semantisch anzureichern. Diese sind automatisch vorverarbeitet und Grundlage für semantische, inhaltsbasierte Abfragen. Dadurch ist es möglich, zielgerichtete und effizientere Abfragemöglichkeiten anzubieten und darüber hinaus Analysen direkt web-basiert auf der Datenbankebene durchzuführen. Da es derzeit kaum operationelle vergleichbare Systeme oder Alternativen gibt, ist dies derzeit ein innovatives Alleinstellungsmerkmal.

Es hat sich bezüglich Datenmenge und Performance als zielführend herausgestellt, die semantischen Layer mit den Eigenschaften von Array-Datenbanken zu kombinieren. Über eine web-basierte Schnittstelle lässt sich die Datenbank abfragen bzw. Analysen auf den Daten durchführen. Hierfür wurde eine Webapplikation entwickelt, in der die Benutzer/-innen auf eine einfache Art und Weise beliebig komplexe Abfragen erstellen, teilen und benutzen können. Dies ermöglicht die Transformation von Wissen in maschinenlesbare Elemente und verkleinert die semantische Lücke, indem das System inkrementell intelligenter wird.

Als nächster Schritt ist eine Optimierung hinsichtlich der Performance und Verwaltung von größeren Datenmengen angedacht. Für die Entwicklung oder Implementierung von verbesserten Methoden ist eine enge Zusammenarbeit mit der Datenbankforschung notwendig. Darüber hinaus wird die Wissensbasis durch weitere semantische, räumlich-zeitliche Abfragen erweitert. Hierfür wird der Prototyp in einer realen Anwendung getestet, wobei auch der Input von verschiedenen Domain-Experten (für unterschiedlichste Fragestellungen) von Bedeutung sein wird. Weitere Forschung an der Schnittstelle zwischen Fernerkundung und Computer Vision kann zur Verbesserung der automatisch erstellten semantischen Layer beitragen. Beispielsweise haben Neuronale Netze (speziell Bildkontext-sensitive convolutional neural networks, CNN) in den letzten Jahren viel Potential gezeigt (LÄNGKVIST et al. 2016), sodass auch diese ergänzend für die Interpretation der Satellitenbilder und bei der Erstellung der semantischen Layer genutzt werden könnten.

6 Danksagung

Die Implementierung des IQ Systems als Prototyp erfolgte innerhalb des FFG (österr. Forschungsgemeinschaft) geförderten Projektes AutoSentinel-2/3 (Vertragsnummer: 848009) im Rahmen des Austrian Space Application Programme (ASAP) und wurde mit dem *Copernicus Masters Award 2015 – Big Data Challenge* ausgezeichnet (http://www.copernicus-masters.com/index.php?kat=winners.html&anzeige=winner_t-systems2015.html).

7 Literaturverzeichnis

- BARALDI, A. & BOSCHETTI, L., 2012 Operational automatic remote sensing image understanding systems: Beyond geographic object-based and object-oriented image analysis (GEOBIA/GEOOIA). Part 1: Introduction. *Remote Sensing* **4** (9), 2694-2735.
- BARALDI, A., DURIEUX, L., SIMONETTI, D., CONCHEDDA, G., HOLECZ, F. & BLONDA, P., 2010: Automatic spectral rule-based preliminary classification of radiometrically calibrated SPOT-4/-5/IRS, AVHRR/MSG, AATSR, IKONOS/QuickBird/OrbView/GeoEye and DMC/SPOT-1/-2 imagery – Part I: System design and implementation. *IEEE Trans. Geoscience Remote Sensing* **48** (3), 1299-1325.
- BAUMANN, P., MAZETTI, P., UNGAR, J., BARBERA, R., BARBONI, D., BECCATI, A., BIGAGLI, L., BOLDRINI, E., BRUNO, R., CALANDUCCI, A., CAMPALANIA, P., CLEMENTSI, O., DUMITRUA, A., GRANTI, M., HERZIGJ, P., KAKALETRISK, G., LAXTONL, J., KOLTSIDAK, P., LIPSKOCHA, K., MAHDIRAJA, A.R., MANTOVANI, S., MERTICARIUA, V., MESSINAM, A., MISEVA, D., NATALIH, S., NATIVIC, S., OOSTHOEKA, J., PAPPALARDOM, M., PASSMOREN, J., ROSSIA, A.P., RUNDUE, F., SENN, M., SORBERAE, V., SULLIVANO, D., TORRISIF, M., TROVATOM, L., VERATELLIH, M.G. & WAGNERJ, S., 2015: Big data analytics for earth sciences: the EarthServer approach. *International Journal of Digital Earth*, 1-27.
- BAUMANN, P. & HOLSTEN, S., 2011. A comparative analysis of array models for databases. *Database Theory and Application, Bio-Science and Bio-Technology*. Springer Berlin Heidelberg. 80-89.

- BAUMANN, P. & STRAMERJOHANN, H., 2014. Towards a systematic benchmark for array database systems. In: *Specifying Big Data Benchmarks*, Springer, 94-102.
- BRINKHOFF, T. & KRESSE, W., 2011. Databases. *Springer Handbook of Geographic Information*, Springer Berlin Heidelberg, 11-34.
- ORACLE, 2016. Oracle Spatial and Graph Website. <http://www.oracle.com/technetwork/database/options/spatialandgraph/overview/index.html>, letzter Zugriff 29.04.2016.
- LÄNGKVIST, M., KISLEV, A., ALIREZAIE, M. & LOUTFI, A., 2016: Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks. *Remote Sensing* **8** (4).
- POSTGIS, 2016. PostGIS Website. <http://postgis.org/>, letzter Zugriff 29.04.2016.
- PURSS, M.B.J., LEVIS, A., SIMON, O., IP, A., SIXSMITH, J., EVANS, B., EDBERG, R., FRANKISH, G., HURST, L. & CHAN, T., 2015: Unlocking the Australian Landsat Archive – From dark data to High Performance Data infrastructures. *GeoResJ* **6**, 135-140.
- RASDAMAN, 2016. Rasdaman Website. <http://www.rasdaman.org/>, letzter Zugriff 29.04.2016.
- RECK, C., CAMPUZANO, G., DENGLER, K., HEINEN, T. & WINKLER, M., 2016: German Copernicus Data Access and Exploitation Collaborative Infrastructure. *Proceedings of the 2016 conference on Big Data from Space*, Soille, P. & Marchetti, P.G. (Hrsg.), 52-55.
- SCI DB, 2016. SciDB Website. scidb.org/, letzter Zugriff 29.04.2016.
- SOILLE, P., BURGER, A., RODRIGUEZ, D., SYRRIS, V. & VASILEV, V., 2016: Towards a JRC Earth Observation Data and Processing Platform. *Proceedings of the 2016 conference on Big Data from Space*, Soille, P. & Marchetti, P.G. (Hrsg.), 65-68.
- TIEDE, D., SUDMANN, M., BARALDI, A., BELGIU, M. & LANG, S., 2016: ImageQuerying – Prototypische Implementierung eines Systems zur raum-zeitlichen und inhaltsbasierten Satellitenbildsuche und Satellitenbildanalyse. *AGIT - Journal für Angewandte Geoinformatik* [im Druck].