

DMARITIME: DATENMANAGEMENT UND -ANALYSE FÜR ECHTZEITDIENSTE IN DER MARITIMEN SICHERHEIT

S. Brüggemann, S. Sievi,
Airbus DS GmbH, Airbus-Allee 1, 28199 Bremen, Deutschland

Zusammenfassung

Das durch das Bundesministerium für Bildung und Forschung (BMBF) geförderte Verbundprojekt "Echtzeitdienste für die Maritime Sicherheit - Security" (EMSec) strebt die Verbesserung der maritimen Sicherheit durch erhöhte Verfügbarkeit und Erreichbarkeit relevanter Informationen an Land und auf See an. In diesem Beitrag wird mit DMARitime ein Ansatz zum Datenmanagement und zur Datenanalyse für EMSec vorgestellt. Dabei werden zunächst zu verarbeitende Daten beschrieben und dann klassifiziert, bevor daraus Anforderungen an Kernkomponenten für DMARitime erarbeitet werden. Diese Kernkomponenten werden vorgestellt.

1. ECHTZEITDIENSTE FÜR DIE MARITIME SICHERHEIT

Das durch das Bundesministerium für Bildung und Forschung (BMBF) geförderte Verbundprojekt "Echtzeitdienste für die Maritime Sicherheit - Security" (EMSec) soll zuverlässigen Zugang zu Informationen aus unterschiedlichen Quellen ermöglichen. Die vorhandenen Quellen werden durch Erdbeobachtungsdaten von Satelliten und Flugzeugen ergänzt, um die Lageerfassung auf See und an der Küste zu verbessern.

Dem Anwender sollen relevante Informationen situationsgerecht bereitgestellt werden. Um dieses zu ermöglichen, müssen die Daten aus den unterschiedlichen Quellen integriert und konsolidiert werden, um sie ihm - der Situation angemessen - in einer Lageanwendung präsentieren zu können. Die schnellere und detaillierte Bereitstellung der Daten soll den beteiligten Behörden und Organisationen mit Sicherheitsaufgaben (BOS) in der maritimen Domäne ermöglichen, kritische Situationen früher zu identifizieren, besser zu vermeiden und schneller zu kontrollieren.

Als zentrale Datenmanagement-Komponente in EMSec dient das Realtime Maritime Situation Awareness System (RMSAS) der Integration und Konsolidierung der Daten. RMSAS beruht auf dem „System of Systems“-Ansatz und implementiert ein föderales Informationssystem aus einzelnen Diensten (SOA). Sobald verfügbar, werden Datenquellen in Echtzeit integriert, auf Basis semantischer Datenmodelle und Technologien konsolidiert und als Informationsprodukte dem Anwender zur Verfügung gestellt.

Der weitere Aufbau des Artikels ist wie folgt: In den Abschnitten 1.1 und 1.2 werden zunächst einzelne Datenquellen des Verbundes und unterstützte Szenarien beschrieben. Abschnitt 2 beschreibt die gesamte Grobarchitektur des RMSAS. Hieraus wird in Abschnitt 3 eine Architektur für DMARitime motiviert, wobei Daten und Datenanalysen kategorisiert und konkrete Anforderungen an DMARitime definiert werden. Abschnitt 4 stellt verwandte Arbeiten vor. In Abschnitt 5 wird dann die Architektur von DMARitime vorgestellt und in Abschnitt 6 zur Datenanalyse genutzt. Die Abschnitte 7 und 8 geben eine Zusammenfassung, einen Ausblick auf zukünftige Arbeiten und schließen mit einem Fazit.

1.1. Verbunddatenquellen

Als weltweit verbreitete Datenquelle in der maritimen Verkehrsführung dienen Meldungen des Automatic Identification System¹ (AIS) als Basis für Lageanwendungen und als Referenzdaten in EMSec. AIS ist ein kooperatives System, welches die aktive Teilnahme der Schiffsbetreiber voraussetzt. Es hat nur eine begrenzte Verlässlichkeit, da es von den Schiffsbetreibern manipuliert oder abgeschaltet werden kann. Ebenso können Übertragung oder Empfang der Meldungen gestört sein.

Terrestrisches AIS, welches auf Empfangsmasten an den Küsten beruht, hat eine eingeschränkte Reichweite, während Satelliten basiertes AIS noch nicht über die gewünschte operative Verlässlichkeit verfügt.

In EMSec wird erforscht, wie zusätzliche Datenquellen auf Basis von Erdbeobachtung aus Satelliten und Flugsystemen die räumliche und zeitliche Auflösung der Informationen in der maritimen Sicherheit signifikant erhöhen können.

Die EMSec Verbundpartner stellen folgende Daten zur Integration in RMSAS zur Verfügung:

1.1.1. AIS

AIS Meldungen dienen nicht nur als Referenzdaten, zusätzlich werden sie Qualitätsüberprüfungen unterzogen und - wo angebracht - regelmäßige Bewegungsmuster erfasst, wie beispielsweise bei Fähren. Es steht terrestrisches - und Flugsystem AIS zur Verfügung. Weiter gibt es Satelliten-AIS² und das AIS-Signal des AIS-Receiver am Columbus-Modul der ISS³.

¹ https://de.wikipedia.org/wiki/Automatic_Identification_System

² http://www.esa.int/Our_Activities/Space_Engineering_Technology/ESA_satellite_receiver_brings_worldwide_sea_traffic_tracking_within_reach

³ http://www.esa.int/Our_Activities/Space_Engineering_Technology/Space_Station_watch_on_world_s_sea_traffic

1.1.2. Satelliten SAR

Satelliten Synthetic Aperture Radar (SAR) stellt mit dem TerraSAR-X hoch aufgelöste Radarbilder zur Verfügung. Mit automatisierten Algorithmen werden Objekte und - wenn möglich - auch Schiffe detektiert, zudem werden die detektierten Objekte und Schiffe mit AIS Meldungen in Beziehung gesetzt.

Darüber hinaus werden Wind und Wellen Informationen aus den Radarbildern extrahiert und mit konventionellen Wetterinformationen in Beziehung gebracht.

1.1.3. Flugsysteme

Ein EMSec Flugsystem ist mit einem AIS Empfänger und mit einem Radarsystem ausgestattet. Die AIS Meldungen werden wie oben beschrieben genutzt. Das Radarsystem wird Objekte und deren Bewegungen in Form von Plots, bzw. Tracks bereitstellen.

Ein weiteres EMSec Flugsystem stellt optische Aufnahmen bereit, und auch in den optischen Bildern werden Objekte detektiert.

Die Objektdetektionen der optischen Bilder sowie die Wetterinformationen werden mit Georeferenzen versehen ausgeliefert.

1.2. EMSec-Szenarien

Zum Abschluss des EMSec Projektes steht eine Validierungskampagne, in der zwei Szenarien der maritimen Sicherheit erprobt werden. Die Validierungskampagne dient der Demonstration aller Beiträge der Verbundpartner im Rahmen einer Übung mit der Bundespolizei See.

Im Rahmen der Simulation der beiden Szenarien werden als zusätzliche Einsatzmittel eine Satellitenmission und die Flugmissionen angefordert und ausgeführt. Die Daten werden in (Nahe-)Echtzeit an das RMSAS System übertragen, dort integriert, weiterverarbeitet und an die Nutzerschnittstelle „Maritime Sicherheit durch Mensch Maschine Interaktion“ (MaSiMMI) [1] übermittelt. Diese wird von einem Verbundpartner bereitgestellt und ist zuständig für die Darstellung der Daten und die Nutzerinteraktion mit dem System.

In der maritimen Sicherheit gelten bestimmte Regionen und Orte wie Naturschutzgebiete oder Off-Shore Plattformen als besonders schützenswert (restricted area). In diesen Arealen kann ein eingeschränkter Schiffsverkehr gelten. Um diesen zu gewährleisten, könnte beispielsweise um eine Ölbohrplattform ein Bereich definiert werden, in den nur Schiffe eines bestimmten Typs oder maximalen Größe oder mit einer maximalen Geschwindigkeit einfahren dürfen. Ab einem bestimmten Seegang soll der Schiffsverkehr vollkommen eingestellt werden.

Anhand dieses Anschauungsbeispiels wird die Motivation und Architektur von DMARitime im weiteren Verlauf des Artikels beschrieben.

2. RMSAS-ARCHITEKTUR

Das Real-time Maritime Situation Awareness System soll bestehende Systeme, die zum Teil seit vielen Jahren im Kontext maritimer Sicherheit eingesetzt werden, unter einem gemeinsamen Dach vereinen. Mit dem hier beschriebenen Architekturkonzept wird eine gemeinsame Basis geschaffen, auf deren Grundlage Informationen effizient und sicher zwischen allen beteiligten Partnern in einem Systemverbund ausgetauscht werden können.

Die Basis bildet eine Service Oriented Architecture (SOA),

deren Prinzip in BILD 1 dargestellt ist. Alle internen und externen Systemkomponenten werden nicht, wie im linken Teil dargestellt, einzelnen miteinander verbunden, sondern, wie rechts dargestellt, in Form von Diensten (Services) an einen Enterprise Service Bus (ESB) angebunden. Über den ESB kommunizieren die einzelnen Dienste mit Hilfe von Nachrichten miteinander.

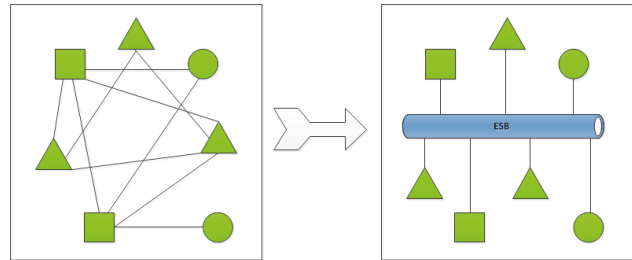


BILD 1 RMSAS Service Oriented Architecture

Diese Architektur bietet die notwendige Flexibilität, eine sehr heterogene Systemlandschaft in ein System zu integrieren, da sich einzelne SOA-Knoten an beliebigen Standorten verteilen lassen. Diese Architektur hat sich allgemein durchgesetzt und wird beispielsweise auch im Kontext von CISE (vgl. [4] [5]) implementiert.

Bei der Integration externer Systeme handelt es sich um Datenquellen und um Systeme mit eigener Geschäftslogik. Die Anbindung an RMSAS erfolgt unter Beibehaltung der Geschäftslogik bei dem jeweiligen System. Daraus folgt, dass die jeweiligen Daten bei dem angebundenen System verbleiben und nicht in die RMSAS interne Datenhaltung dupliziert werden. Die Harmonisierung und Verarbeitung der Daten erfolgt in Mehrwertdiensten auf Basis semantischer Technologien.

3. ARCHITEKTURMOTIVATION FÜR DMARITIME

Wie eingangs beschrieben, hat RMSAS als Datenmanagementkomponente die Aufgabe, Daten aus unterschiedlichen Quellen in unterschiedlichen Formaten zu integrieren und auf Basis eines semantischen Datenmodells zu harmonisieren, zu prozessieren und zu analysieren. Die in EMSec in der maritimen Domäne auftretenden Daten werden nun kategorisiert.

3.1. Kategorisierung von Daten

In RMSAS werden verschiedene Arten von Daten prozessiert. Diese lassen sich folgendermaßen kategorisieren:

- **Datenströme:** Datenströme sind Daten, die in Echtzeit kontinuierlich erzeugt und von RMSAS verarbeitet werden. Das können beispielsweise AIS-Daten sein, die kontinuierlich die Situation der deutschen Bucht melden.
- **Statische Daten:** Statisch sind Daten, die in Datenbanken, auf FTP-Servern oder in externen Systemen gehalten werden. Das können Daten über Schiffe sein, die nicht ausschließlich in AIS Meldungen enthalten sind, wie der Schiffstyp, Ladung, Herkunftshafen oder historische Daten wie die angelaufenen Häfen. Die Daten einer Satellitenmission werden als Paket verfügbar gemacht, nachdem sie auf die Erde übertragen und ausgewertet wurden.
- **Offene Daten:** Offene Daten sind Daten aus der Linked Open Data Cloud oder aus anderen externen

Datenbeständen, welche für bestimmte Analysen Mehrwerte oder Metadaten liefern können. Das sind beispielsweise Ontologien, welche real existierende Häfen verwalten (z.B. Geonames⁴), Geografische Informationen, besonders relevante Orte (Points of Interest, dbPedia⁵), Wetterdaten (Openweathermap⁶), oder ähnliche Informationen (Openstreetmap⁷).

3.2. Kategorisierung von Datenanalysen

Die in diesem Beitrag beschriebene Datenmanagement-Architektur hat zum Ziel, verschiedene Arten von Datenanalysen zu ermöglichen. Diese lassen sich wie folgt unterscheiden:

- Fusion: Die Daten, die von den RMSAS-Partnern zur Verfügung gestellt werden, sind zum Teil bereits mit Daten von anderen Sensoren fusioniert worden. Dennoch kann es sinnvoll sein, die Daten oder Metadaten verschiedener Partner in RMSAS miteinander zu fusionieren, um Dinge zu erkennen, die eine einzelne Quelle nicht erkennen könnte. Weiter macht es Sinn, Daten der Partner mit offenen Daten zu fusionieren. Hierbei können bspw. Geodaten wie Häfen mit Geonames abgeglichen werden, um Schiffsangaben zu verifizieren.
- Bildauswertung: Da in EMSec auch Bilder aus Radar- und optischen Systemen zur Verfügung stehen, müssen sich diese auch im RMSAS Datenmanagement wiederfinden. Die Bildauswertung findet im Rahmen der Flug- und Satellitenmissionen statt. Das kann das einfache Weitergeben eines Bild-Datenstroms hin zu einem Empfänger sein oder auch die komplexe Analyse der Bilddaten hinsichtlich der Identifikation von Schiffen oder anderer relevanter Objekte. Die Ergebnisse werden in RMSAS integriert.
- Anomalie Detektion: Anomalien in Daten können Hinweise auf illegales oder verdächtiges Verhalten liefern. Sowohl zur Prävention als auch zur Analyse historischer Daten muss RMSAS geeignet sein. So kann das Einfahren eines Schiffes in ein eingeschränktes Areal (restricted area) eine Anomalie sein. Das in Abschnitt 1.2 eingeführte Beispiel verdeutlicht dies.
- Plausibilitätsprüfung: In der maritimen Domäne kommt es einerseits oft zu Fehlalarmen, wie beispielsweise bei Search And Rescue Transponder (SART)-Events, und andererseits auch zum missbräuchlichen Einsatz von Systemen, bspw. bei AIS-Sendern. Daher ist wichtig, die Plausibilitätsprüfung von Daten unterstützen zu können. Diese werden von Verbundpartnern zur Verfügung gestellt und in RMSAS integriert. Mögliche Anwendungsfälle können folgende sein:
 - Als Einsatzleiter möchte ich eine Plausibilitätsprüfung eines AIS-SART Alarms bekommen, um besser einschätzen zu können, ob es sich um einen echten Notfall handelt. und
 - Als Nutzer der Lage möchte ich Informationen über die Plausibilität von AIS Daten erhalten, da diese fehlerhaft sein können.
- Objektverfolgung: Dem Aspekt der Objektverfolgung kommt in Lageanwendungen eine hohe Bedeutung zu. Dabei werden Objekte (Schiffe) in Daten und

Datenströmen identifiziert und verfolgt, um Tracks zu identifizieren und zu analysieren. So kann es von Interesse sein, ein Schiff mit einer besonderen Ladung gesondert im Blick zu behalten.

- Regelauswertung: Regeln beschreiben zulässiges, unzulässiges und auffälliges Verhalten und lassen sich nutzen, um Daten und Datenströme auszuwerten und um im Falle einer Regelerfüllung weitere Aktionen zu initiieren.
- Inferenz: Durch Ableitung neuen Wissens lassen sich in Daten Umstände erkennen, die nicht oder nicht explizit modelliert worden sind. Beim Schiffsprofiling beispielsweise lassen sich durch den Einsatz von Reasoning-Technologien Aspekte erkennen, die nicht direkt erkennbar sind.

3.3. Anforderungen an DMARitime

Aus den oben beschriebenen Faktoren und Rahmenbedingungen lässt sich eine Reihe von Anforderungen ableiten, die das semantische Datenmanagement erfüllen muss. Das übergeordnete Ziel besteht darin, anhand des semantischen Datenmodells erstens die Daten der maritimen Domäne entsprechend zu strukturieren und zweitens darauf aufbauend einen situationsgerechten Zugriff auf die Daten zu ermöglichen.

Als wichtige Grundlage müssen die Datenquellen in RMSAS integriert und verfügbar gemacht werden. Wie bereits beschrieben, bleiben alle integrierten Daten wenn möglich in ihrem Ursprungsformat und -ort erhalten. Weiterhin werden die Daten in dem Datentyp nach optimierten Datenhaltungssystemen gehalten. Objekt oder Schiffsdaten können als AIS Meldungen mit Position und statischen Schiffsinformationen oder in Form von detektierten Objekten aus Bildformaten stammen. Statische Eigenschaften, die sich nur selten ändern, lassen sich gut in relationalen Datenbanken oder in Triple-Stores speichern. Georeferenzierte Daten wie Positionen, Tracks oder andere geographische Gegebenheiten werden in Geo-Servern gespeichert werden.

Es muss also eine Verbindung zwischen der semantischen Datenmodellierung und der heterogenen Datenhaltung geschaffen werden. Komponenten dieser Architektur müssen folgende Aspekte unterstützen:

- Prozessierung der Daten
 - Daten-Dekodierung und Zuordnung zum Datenmodell
 - Metadatenfusion
 - Metadatenspeicherung
 - Temporäre Datenspeicherung
- Informations- und Datenzugriff
 - Datenzugriff über das semantische Modell

4. VERWANDTE ARBEITEN

In der maritimen Domäne werden Ontologien zunehmend verwendet, um Daten zu integrieren und zu analysieren. In [9] werden Ontologien verwendet, um die taktische Entscheidungsfindung zu unterstützen. Dazu wird eine maritime Domänenontologie verwendet, um relevante Entitäten und ihre Beziehungen zu modellieren.

In [4] wird im Projekt CISE eine maritime Top Level Ontologie vorgeschlagen, mit der eine gemeinsame Umgebung zum Informationsaustausch etabliert wird.

⁴ <http://www.geonames.org/>

⁵ <http://dbpedia.org>

⁶ <http://openweathermap.org/>

⁷ www.openstreetmap.org/

In [8] wird ein ontologiebasierter Ansatz zur Datenintegration vorgestellt, bei dem das Konzept des "ontology based data access (OBDA)" verfolgt wird. Dort wird ebenfalls, wie in Abschnitt 3.1 beschrieben, zwischen statischen und dynamischen Daten unterschieden. Dort wird vorausgesetzt, dass Daten bereits in relationaler Form vorliegen, die dann in eine ontologiebasierte Form überführt werden und dort semantisch analysiert werden können. Dort wird jedoch nicht darauf eingegangen, dass Daten auch in heterogener Form, bspw. in Dateien oder durch Webservices zur Verfügung stehen können.

5. ARCHITEKTUR VON DMARITIME

In diesem Abschnitt wird nun mit DMARitime eine Architektur beschrieben, die in der Lage ist, die in den vorangegangenen Abschnitten motivierten Anforderungen zu erfüllen.

BILD 2 bietet einen Überblick über die DMARitime Funktionsblöcke.

- Request Management: In dieser Komponente wird der Zugriff auf die Daten verwaltet. Zu den Anfragen gehören sowohl Daten- als auch Missionsanfragen.
- Semantic Data Processing: Die semantische Datenverarbeitung ist die Kernkomponente, bestehend aus den semantischen Datenmodellen, - Regeln und einem Reasoner. Dieser wertet zur Laufzeit des Systems Anfragen und Regeln aus.
- Source Mapping: Das Source Mapping übernimmt mithilfe von Regeln die Übersetzung zwischen Datenmodell und Datenhaltung.

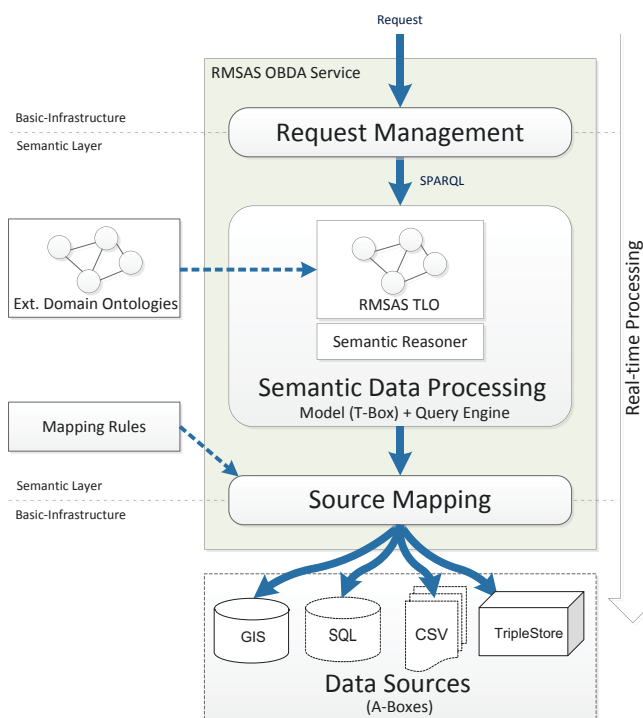


BILD 2 Logischer Aufbau von DMARitime

Auf einige Komponenten wird im Weiteren detailliert eingegangen.

5.1. Source Mapping: Virtualisierung heterogener Datenquellen

Das Framework "ontop" (<http://ontop.inf.unibz.it/>) ist im Rahmen des EU-Projekts Optique (<http://optique-project.eu/>) entstanden. Es ermöglicht die Abfrage von Daten, die in einer relationalen Datenbank (RDB) gespeichert sind. Mit ontop ist DMARitime in der Lage, einen Triplestore zur Verfügung zu stellen, der die Daten einer relationalen Datenbank anbieten kann, ohne diese Daten selbst vorhalten zu müssen. Dabei kann ontop einen sogenannten virtuellen RDF-Graphen erzeugen. Dieser kann auf T-Box-Ebene einer externen Ontologie wie bspw. der CISE-Ontologie genügen. Die Instanzen zu dieser Ontologie werden dynamisch bei jeder SPARQL-Abfrage gegen den Triplestore generiert. In BILD 3 ist die Funktionsweise von ontop beschrieben: Eine relationale Datenquelle wird durch ein Mapping auf eine Ontologie abgebildet. Ein Reasoner nimmt von RMSAS-Diensten Anfragen entgegen und stellt diese dann nicht direkt an die Datenquelle, sondern höherwertig an die Ontologie. Der Reasoner übersetzt die Anfragen dann an die Datenbank.



BILD 3 Ontologiebasierter Datenzugriff mit ontop

Dabei wird eine SPARQL-Abfrage an eine Ontologie in eine auszuführende SPARQL-Abfrage übersetzt. Diese wird dann in eine SQL-Abfrage überführt und gegen die Datenbank ausgeführt. Die Ergebnisse werden dann in A-Boxen der Ontologie überführt und über den TripleStore verfügbar gemacht, der diese Ergebnisse als virtuellen RDF-Graphen bereitstellt und so dem ontop-Reasoner ermöglicht, die Anfrage zu beantworten. Diese Ergebnisse werden dann der Anfrage-Komponente übergeben.

Mit Ontop-spatial steht eine Erweiterung von ontop zur Verfügung, welche geographische Funktionen zur Verfügung stellt. Es implementiert den geoSparql-Standard⁸ und stellt die entsprechenden Funktionen zur Verfügung. geoSPARQL erweitert SPARQL um geographische Filter. So lassen sich Anfragen der Form "Welches Schiff befindet sich in einem bestimmten Bereich" oder "Welche Objekte sind in der Nähe einer Havarie" beantworten.

⁸ <http://www.opengeospatial.org/standards/geosparql>

5.2. Föderation von Datenquellen

Da RMSAS Daten analysieren muss, die aus verschiedensten Datenquellen stammen, wird mit Teiid⁹ ein Werkzeug verwendet, welches diese verschiedenen Daten in eine Datenbank virtualisieren kann. In BILD 4 ist dies schematisch dargestellt: Verschiedene Quellen wie Dateien, Datenbanken oder Web Services werden durch Teiid in eine virtuelle Datenbank zusammengefasst. Diese Datenbank kann dann wiederum ontop zur Verfügung gestellt werden. Somit ist es möglich, SPARQL-Anfragen an die RMSAS-Ontologie (nach CISE) zu stellen, die dann mit Hilfe von ontop mit Daten aus verschiedenen Datenquellen beantwortet wird.

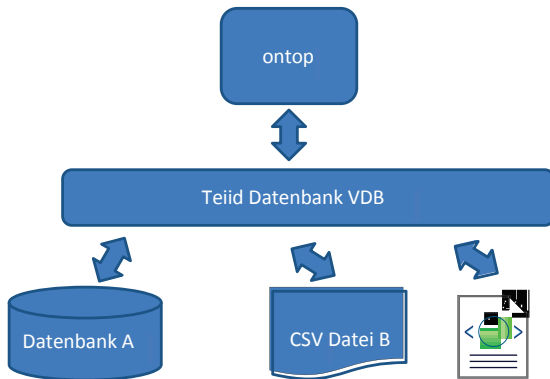


BILD 4 Datenvirtualisierung mit Teiid

Die einzelnen Datenquellen Datenbank A, CSV-Datei B und Webservice C werden mit Teiid in eine separate Datenbank VDB virtualisiert. Diese Teiid-Datenbank ermöglicht dadurch den virtuellen Zugriff auf die einzelnen Datenquellen, da sie für jede Datenquelle weiß, wie sie die benötigten Daten abfragen kann. Für die Datenbank A werden die Daten relational per SQL ausgelesen, die CSV-Datei B wird entsprechend geöffnet und dann ausgelesen, und für den Webservice C wird definiert, über welchen SOAP-Request Daten abgefragt werden können. Der SOAP-Response wird entsprechend definiert. Dies ermöglicht es dann, die virtuelle Datenbank VDB zu erstellen. Ein Triplestore D enthält unabhängig zu den anderen Datenquellen ebenfalls Daten und Metadaten zu Schiffen.

5.3. Semantische Datenmodellierung

Während es für einen Menschen relativ einfach ist, die Bedeutung einer Information in einem bestimmten Kontext zu erfassen und mit anderen Informationen in Beziehung zu setzen, so erfordert die maschinelle Verarbeitung solcher Informationen den Einsatz semantischer Technologien.

In der konventionellen Datenhaltung werden Daten, wie die Informationen zu einem Schiff mit Name="Black Pearl", Kapitän="Jack Sparrow", MMSI Nummer ="123123" in Tabellen in Form von Zeichenketten gespeichert. Die Bedeutung dessen, dass es sich bei dem Kapitän um einen Menschen mit dem Namen Jack Sparrow mit dem Beruf Kapitän handelt, ist in der konventionellen Datenhaltung nicht enthalten. Diese Lücke wird durch die semantische Datenmodellierung geschlossen.

Diese ermöglicht eine formale Beschreibung der Wissensbasis einer bestimmten Domäne. Damit wird eine automatisierte Auswertung (Reasoning) der Daten ermöglicht und daraus folgend die Extraktion von Informationen und Wissen.

5.3.1. Ontologien

Semantische Datenmodelle werden in Form von Ontologien erstellt. Sie stellen eine Art Netzwerk von Begriffen dar, die über logische Beziehungen miteinander verknüpft sind. Es werden die Konzept- (T-Box) und die Individuen-Ebene (A-Box) voneinander unterschieden. Die Konzeptebene wird mit dem semantischen Modell abgebildet, während die Individuen die tatsächlichen Daten und somit die Wissensbasis repräsentieren. Dieses erfolgt über sogenannte Tripel bestehend aus Subjekt, Prädikat und Objekt.

Gerade in einem komplexen und verteilten System-Verbund wie in EMSec ist der einheitliche Zugriff auf Informationen aus den unterschiedlichsten Quellen unverzichtbar für das Funktionieren des Gesamtsystems. Typischerweise ist die Entstehung, Weiterentwicklung und Nutzung von Wissen in solchen Systemen räumlich wie organisatorisch verteilt (siehe [6]). Semantische Technologien stellen die Grundlage für eine solche Informationsaufbereitung zur Verfügung und sind konzeptuell für die Wissensrepräsentation in verteilten Systemen bestens vorbereitet. Nach [7] stellen sie damit ein umfassendes Paradigma zur Entwicklung wissensintensiver IT-Systeme in verteilten Systemlandschaften dar.

Das RMSAS Datenmodell basiert auf semantischen Konzept-Ontologien, als formale und explizite Repräsentation von fachspezifischen (Hintergrund-) Wissen. Das RMSAS Datenmodell stellt ein Basis-Vokabular für den Informationsaustausch der am Systemverbund beteiligten Systeme zur Verfügung und ist damit die Grundlage für einen zukünftigen Wissensaustausch zwischen verschiedenen Akteuren in der maritimen Sicherheit. Dabei erfolgt die sog. Konzeptualisierung, also die Modellierung der grundsätzlichen "Weltsicht" innerhalb des RMSAS, getrennt von der Betrachtung konkreter Objektinstanzen und dem Wissen über deren aktuellen Zustand, da solche Informationen normalerweise systemspezifisch sind, und damit völlig unterschiedlich strukturiert sein können. Der Zugriff auf konkrete Objektinstanzen soll erst zur Laufzeit erfolgen und lediglich über das RMSAS Datenmodell formal gesteuert werden. Dies wird durch das Konzept des "Ontology Based Data Access (OBDA)" realisiert, welches von ontop (vgl. Abschnitt 5.1) implementiert wird.

Für die Konzeptualisierung werden im DMARitime formale Beschreibungssprachen wie z.B. RDF und OWL verwendet, da die Ausdrucksmächtigkeit solcher Sprachen ein wesentliches Element beim Entwurf eines auf semantischen Technologien basierenden Informationssystems ist [11].

Regeln

Regeln sind eine Erweiterung der semantischen Datenmodellierung. Mit ihnen können konjunktive Anfragen auf die Wissensbasis formuliert werden. Die Regeln formulieren logische Ausdrücke, anhand derer dem Ausdruck entsprechende Individuen ermittelt werden. Eine gebräuchliche Regelsprache ist die Semantic Web

⁹ <http://teiid.jboss.org>

Rule Language (SWRL)¹⁰, die in DMARitime angewendet wird. SWRL kann bspw. dazu genutzt werden, für das Beispiel aus Abschnitt 1.2 Regeln aufzustellen, die dann entsprechend „wahr“ werden, wenn ein nicht autorisiertes Schiff eine restricted area befährt. Dies ist die Funktion eines Reasoners, der im folgenden Abschnitt eingeführt wird.

5.3.2. Reasoning

Einer Reasoner überprüft auf der Konzept-Ebene (T-Box), ob das semantische Modell in sich konsistent und widerspruchsfrei ist. Auf der Individuen-Ebene kann mit einem Reasoner neues Wissen über Inferenz abgeleitet werden. In dem Moment, wo es eine Instanz von Kapitän=“Jack Sparrow“ existiert, kann aus dem Modell abgeleitet werden das, Jack Sparrow ein Mensch ist, der den Beruf Kapitän ausübt. Ebenso werden über das Reasoning die semantischen Regeln ausgewertet.

5.3.3. Gemeinsames Datenmodell

Das semantische Datenmodell in DMARitime bildet das Vokabular der der zivilen maritimen Sicherheitsdomäne ab. Darüber hinaus beinhaltet es auch die Struktur und das Vokabular, um die BOS der maritimen Sicherheit in ihren Aufgaben zu unterstützen. Die Modellierung des Datenmodells folgt den Prinzipien des Situationsbewusstseins (Situation Awareness). Zusätzlich wird das DMARitime Datenmodell in Anlehnung an das CISE Modell entwickelt, welches im Kontext des CoopP Projektes [10] entstanden ist. Das Datenmodell gliedert sich in das Kernmodell, das RMSAS Datenmodell als Top Level Ontologie (TLO). Dieses ist in Gruppen gegliedert. Es werden die relevanten Entitäten der maritimen Domäne in den Klassen Objekt, Agent und Position Navigation Time (PNT) abgebildet. Zudem bilden generische Klassen Datentypen ab, sowie Klassen um Nutzerinteraktion und das Prozessieren der Daten in RMSAS abzubilden. Daneben gibt es noch externe Domänenontologien, welche generische von der maritimen Domäne unabhängige Entitäten modellieren. Zu diesen gehören Qualitäts- und Sicherheitseigenschaften.

Die zentralen Entitäten werden im Folgenden beschrieben:

- Object: Objekt beschreibt die wichtigen Objekte der maritimen Domäne, die unter anderem in einer Lageapplikation dargestellt werden können.
- Agent: Agenten sind Personen oder Organisationen, die mit anderen Entitäten in Beziehung stehen. Beispielsweise hat ein Schiff einen Kapitän, oder ein Datenprodukt wird von einer Organisation bereit gestellt.
- PNT: Die Gruppe Position Navigation Time beinhaltet die Entitäten mit einem zeitlichen und geografischen Bezug: DateTime, Geometry, Location, Position, Dimension, Movement
- Environmental Condition: Die Umgebungsbedingungen gliedern sich in Wetterbedingungen und speziell in Meteo Oceanographic Conditions (METOC).
- Datatypes: Datentypen sind grundsätzlich generisch, werden bei Bedarf für RMSAS spezialisiert. DMARitime kennt Events, Activities, Data, Metadata, Quality, und Security.

- Dataprocessing: Schließlich wird mit der Modellierung der Datenverarbeitung dieselbe in RMSAS unterstützt:
 - Product: Beschreibt die Art von Informationen die während der Verarbeitung der Daten entstehen. Die Informationsarten entsprechen den generischen Datentypen. Ein Produkt kann ein Produkt enthalten, so ist nachvollziehbar, welches Produkt auf welchem anderen beruht.
 - Service: Dienste beinhalten die Prozesslogik der Datenverarbeitung, über die Beziehung zwischen Dienst und Produkt und Dienst und Organisation wird nachvollzogen welcher Dienst welches Produkt erzeugt und welche Organisation die Prozesslogik bereit stellt oder nutzt.
 - Sensortype: Sensortypen stehen für die Art von Sensoren, die Werte liefern. Dies kann Satelliten- oder Terrestrisches AIS sein, oder auch eine Beobachtung durch einen Mensch.
 - Message: Zwischen den RMSAS Diensten erfolgt der Austausch von Daten und Informationen über Nachrichten.

6. DATENANALYSE MIT DMARITIME

Viele der im gemeinsamen Datenmodell modellierten Objektklassen bilden Beziehungen untereinander, bzw. werden voneinander abgeleitet. Die aus verschiedenen Quellen bezogenen Informationen über Objekte lassen sich grob in drei Level einteilen, wie in BILD 5 dargestellt ist:

- Detektierte Objekte: Die Feststellung der Existenz eines Objektes. Diese Information kann z.B. aus der Auswertung von Radar-Bildern gewonnen werden
- Klassifizierte Objekte: Zu einem detektierten Objekt können nähere Angaben zur Art / Typ des Objektes gemacht werden. So könnte z.B. zweifelsfrei geklärt sein, dass es sich bei dem Objekt um ein Schiff handelt.
- Identifizierte Objekte: Das Schiff ist eindeutig identifiziert

Das RMSAS Datenmodell ist konzeptionell in der Lage, einzelne Objektdaten auf jedem der drei Level zu erfassen, und zusammengehörige Objektdaten miteinander in Beziehung zu setzen. Ziel ist es, durch Informationsaufbereitung immer den jeweils höchsten Level für ein bestimmtes Objekt zu erreichen. Dies kann bedeuten, dass Schiffe aus den vorliegenden Daten eindeutig identifizieren zu können oder bei Umweltdaten, wie zum Beispiel Wellen, eine "Klassifikation" vornehmen zu können.

¹⁰ <http://www.w3.org/Submission/SWRL/>

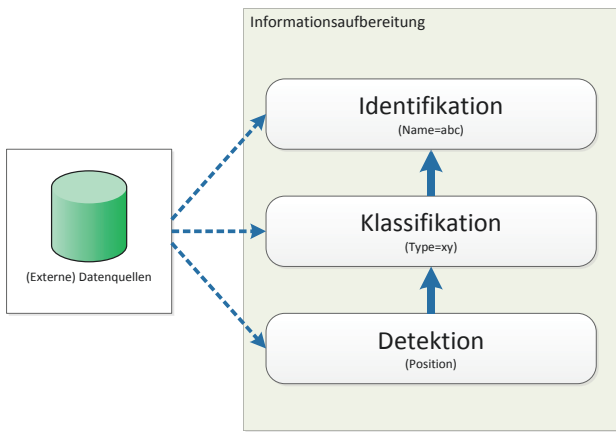


BILD 5 Informationsaufbereitung

6.1. Datenfusion

Eingehende Daten werden dekodiert und dem Modell gemäß strukturiert. Im Fall von EMSec handelt es sich dabei in aller Regel um Objektdaten. Bei der Fusion werden alle gleichartigen Informationen zusammengefasst. So können beispielsweise die Informationen zu einem Schiff über unterschiedliche AIS Quellen in das System integriert werden. In diesem Fall werden die Informationen zu diesem Schiff in ein Informationsprodukt zusammengefasst. Der Beipackzettel gibt Auskunft darüber, welche Datenquellen eingeflossen sind.

Zusätzlich werden die statischen Daten und die Metadaten der eingehenden Daten zusammengetragen und auf Konsistenz geprüft.

In BILD 6 ist gezeigt, wie Daten in DMARitime beispielsweise für eine Lageanwendung fusioniert werden können. Es sind verschiedene Ebenen zu sehen, die mit Daten aus unterschiedlichen Quellen erzeugt wurden. In rot sind dort Schiffe gekennzeichnet, die an Hand ihres AIS-Signals identifiziert wurden. Aus einem GIS-System stammen Informationen zur Land- und Seeseite, und ein Satellit liefert zusätzliche Informationen zur Beschaffenheit des Wassers.

Die Identifikationsklasse und Konsistenz der einzelnen Objektdaten können erst mit der Integration in RMSAS untereinander geprüft werden. Daher sieht RMSAS einen Mehrwertdienst „Beipackzettel“ vor, der mit den Informationsprodukten ausgeliefert wird. Der Beipackzettel enthält miteinander konsolidierte Metadaten und statische Informationen. Er dient dazu, dem Nutzer größtmögliche Transparenz über die ihm zur Verfügung gestellten Informationen zu geben.

Die Erstellung des Beipackzettels beruht auf der Fusion von statischen Daten und Metadaten auf Basis einer Echtzeitdatenanalyse.

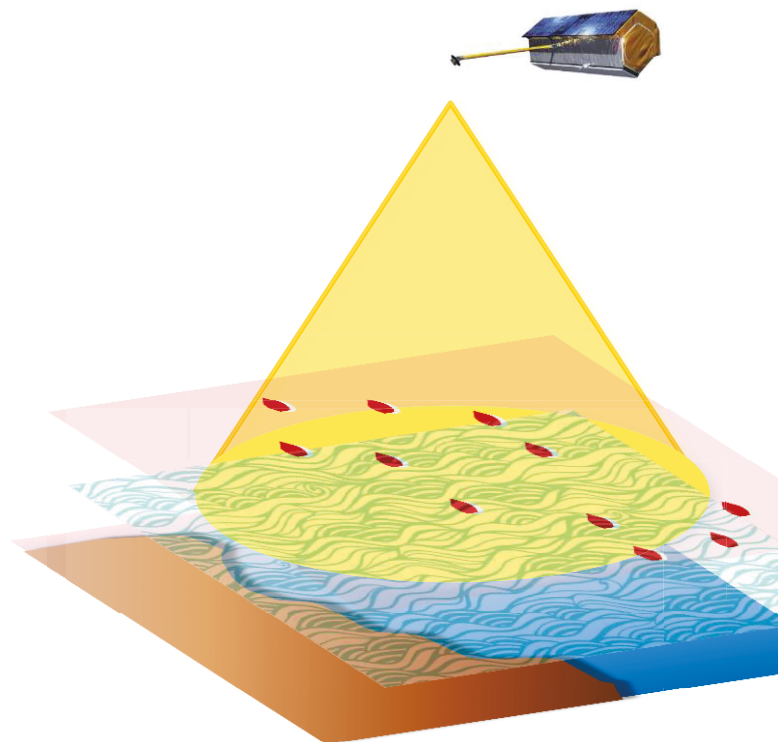


BILD 6 Datenfusion in DMARitime

6.2. Regelauswertung

Regeln werden eingesetzt, um die Datensätze auf Konsistenz zu prüfen. Bestimmte Inkonsistenzen können über eine Nachricht an die Lageanwendung gemeldet werden.

Darüber hinaus werden über Regeln die Qualitätsinformationen der Informationsprodukte ermittelt. Ein über AIS empfangenes Schiff erhält das Identification Level „identified“, während ein Objekt aus einer Luftaufnahme das Identification Level „detected“ erhält.

Wird ein Schiff von mehreren Quellen gesehen, erhöht sich die Glaubwürdigkeit der Information.

6.3. Inferenz

Durch Inferenz lassen sich in Daten Umstände erkennen, die nicht oder nicht explizit modelliert worden sind. Beim aktiven Schiffsprofiling beispielsweise lassen sich durch den Einsatz von Reasoning-Technologien Aspekte erkennen, die nicht direkt modelliert worden sind.

Inferenz bezeichnet die Herleitung neuen Wissens aus bekanntem Wissen, das bereits in strukturierter Form vorliegt. Inferenz bezeichnet somit die Fähigkeit, Beziehungen zwischen einzelnen Informationen herzustellen, die in dieser Form nicht zuvor explizit modelliert wurden. Diese Art der Wissensextraktion ist für DMAritime Konzept besonders interessant, da sie eine Analyse der vorhandenen Daten und Metadaten auf Inkonsistenzen und Anomalien und die Erkennung von wiederkehrenden Mustern durch Regelauswertung erlaubt, die dann wiederum in die geforderte Qualitätsbewertung von Informationsprodukten einfließen.

7. ZUSAMMENFASSUNG UND AUSBLICK

In diesem Beitrag wurde mit DMAritime eine Architektur zum Datenmanagement und zur Datenanalyse zur Echtzeitdatenverarbeitung in der zivilen maritimen Sicherheit vorgestellt und am Beispiel des EMSec-Projektes motiviert. Dazu wurden zunächst relevante Datenquellen beschrieben und es wurden die zu verarbeitenden Daten klassifiziert. Dann wurde vorgestellt, welche Datenmanagement-Komponenten für DMAritime konzeptionell notwendig sind. Weiter wurden verschiedene Datenanalyse-Verfahren beschrieben, für die sich eine Notwendigkeit in der Domäne der zivilen maritimen Sicherheit ergeben.

Im weiteren Projektverlauf werden die für die Validierungsszenarien des EMSec-Projektes relevanten Informationsbedarfe detailliert und im Hinblick auf die Verarbeitung in DMAritime analysiert.

8. FAZIT

In der Domäne der zivilen maritimen Sicherheit stehen Entscheider vor der Situation, dass sie zur Einschätzung einer maritimen Schadenslage und zur Entscheidungsunterstützung eine Vielzahl von Daten benötigen, die oftmals in unstrukturierter Form und in heterogenen Datenformaten in verteilten Datenquellen vorhanden sind. Weiter müssen diese Daten in maritimen Notsituationen in Nahe-Echtzeit zur Verfügung stehen und so detailliert oder abstrahiert sein, dass sie den Nutzer weder überfordern noch enttäuschen.

Der in diesem Beitrag mit DMAritime vorgestellte Ansatz hat zum Ziel, dem Anwender für die maritime Situationseinschätzung genau die Daten zur Verfügung zu stellen, die er in der jeweiligen Situation benötigt. Dazu werden in DMAritime verschiedene Datenmanagement- und Analyse-Funktionalitäten bereitgestellt, die von einer Lageanwendung genutzt werden können, um den Informationsbedarf des Nutzers zu stillen.

Literatur:

- [1] S.Brüggemann und S. Förster, User requirements for real-time services for the maritime security, International Conference on Information on Ships, 2014
- [2] Towards the integration of maritime surveillance: A common Information sharing environment for the EU maritime domain, Brussels, 15.10.2009. Available at:
<http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52009DC0538:en:NO>
- [3] CoopP Final Report ISBN 978-952-491-901-2, ISBN 978-952-491-902-9
- [4] EU CISE 2020 [SEC-2013.5.3-2] European test bed for the maritime Common Information Sharing Environment in the 2020 perspective
- [5] Integrating Maritime Surveillance - ISBN 978-92-79-17050-8
- [6] M. Krötzsch, F. Simancík, I. Horrocks. A Description Logic Primer. 2013
- [7] <http://teiid.jboss.org/about/>
- [8] G. M. Santipantakis, K. I. Kotis, and G. A. Vouros. Ontology-based data integration for event recognition in the maritime domain. In R. Akerkar, M. D. Dikaiakos, A. Achilleos, and T. Omitola, editors, Proceedings of the 5th International Conference on Web Intelligence, Mining and Semantics, WIMS 2015, Larnaca, Cyprus, July 13-15, 2015, page 6. ACM, 2015.
- [9] R. Haberlin, P. da Costa, K. Laskey, An Ontology for Hypothesis Management in the Maritime Domain. 16th ICCRTS, 4444(703), 0-18
- [10] Test Project on Cooperation in execution of various maritime functionalities at subregional or sea-basin level in the field of integrated maritime surveillance (CoopP). Final Report of Work Package 5: Specifications of Common Data Formats and Semantics, 2014.
- [11] Andreas Dengel. Semantische Technologien. 2012. Spektrum Akademischer Verlag. Heidelberg