

# VERWENDUNGSMÖGLICHKEITEN OFFENER DATEN IN DER MARITIMEN SICHERHEIT

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

## Zusammenfassung

In diesem Beitrag wird gezeigt, wie offene Daten einen Beitrag zur Erhöhung der maritimen Sicherheit leisten können. In der maritimen Domäne sind Informationen über die aktuelle maritime Lage wichtig, um im Einsatzfall notwendige Entscheidungen treffen zu können. Offene Daten sind Daten, die in verschiedenen Quellen zur freien Verfügung stehen. Dieser Beitrag stellt einige dieser Datenquellen vor und beschreibt Szenarien, in denen diese gewinnbringend eingesetzt werden können.

## 1. ECHTZEITDIENSTE FÜR DIE MARITIME SICHERHEIT

Das Projekt EMSec: Echtzeitdienste für die maritime Sicherheit hat zum Ziel, Informationen den Endanwendern situationsgerecht darzustellen. Bei einer durchgeführten Benutzerbefragung [1] wurde offenkundig, dass oftmals behördlich verwaltete Daten nicht die benötigten Informationen enthalten und nur schwer zugänglich sind. Die durch die Befragung identifizierten Anwendungsfälle wurden dahingehend analysiert, ob frei verfügbare Daten als ergänzende Datenquellen einen Beitrag bei der Erhöhung der maritimen Sicherheit leisten können.

Offene Daten werden von verschiedensten Einrichtungen vermehrt veröffentlicht. Portale wie das Open Data Network [2] und GovData.org [3] bieten allgemeine Daten verschiedenster Kategorien, von Bevölkerungszahlen, geographischen Daten, Infrastruktur, hin zu Wetterdaten. Konkret für die maritime Domäne existieren Verbünde wie SafeSeaNet [4] zum Monitoring des Schiffsverkehrs und CleanSeaNet [5] zur satellitenbasierten Detektion von Verschmutzungen und Schiffen. Mit Copernicus [6] stehen frei verfügbare Satelliten-Daten zur Verfügung, die Mehrwerte bei der Verwendung in der maritimen Domäne erzeugen. Aus offenen Daten wurde ein semantisches Netz gebildet, welches Linked Data [7] in Form vernetzter Daten anbietet. Dieses Linked Data Netzwerk ermöglicht es, diese Daten automatisiert und strukturiert abzurufen und so in eigene Datenverarbeitungsprozesse zu integrieren.

Soziale Netzwerke lassen sich nutzen, um Informationen zu (Groß-)Veranstaltungen zu gewinnen, um diese in die Einsatzplanung einzubeziehen. Nutzeraktivitäten bei sozialen Netzwerken lassen sich exemplarisch verwenden, um Bewegungsprofile im Krisenfall in die Entscheidungsfindung einzubeziehen.

In diesem Beitrag wird zunächst vorgestellt, wie aus Daten und Informationen Wissen generiert wird. Dann wird gezeigt, wie maritime Daten mit offenen Daten verknüpft werden können. Dazu werden diverse frei zugängliche Datenquellen vorgestellt. Einen Schwerpunkt bildet das Semantic Web und die Linked Open Data Cloud, welche maschinenles- und verarbeitbare Daten zur Verfügung stellt, die mit Metadaten annotiert sind.

Dann werden in diesem Beitrag Anwendungsfälle vorgestellt, die mit offenen Daten sinnvoll unterstützt werden können. Diese reichen von der Analyse von Schiffsrouten und der Detektion von Anomalien über die Ausbreitung von Ebola über Ländergrenzen hinweg bis hin zur Identifizierung kritischer Infrastrukturen im Krisenfall und der Konsistenzprüfung satellitengestützt erhobener Daten.

Der Beitrag schließt mit einer Zusammenfassung und einem Fazit.

## 2. VERWANDTE ARBEITEN

In [9] werden offene Daten verwendet, um Anomalien zu detektieren. Dabei werden Daten aus frei zugänglichen Quellen verglichen, wie AIS-Informationen, Daten von Hafengebireitern, Behörden, Wetterdaten, Reedereien und Versicherungen. Der Schwerpunkt dieser Arbeit liegt auf der Identifikation maritimer Anomalien, konkret auf auffälligem Schiffsverhalten und der Nutzung von AIS-Daten.

In [10] werden verteilte Daten genutzt, um ein Lagebild der maritimen Situation herstellen zu können. Dabei wird eine maritime Situationsontologie erstellt, in der diverse verteilte Daten abgebildet werden. Dabei wird hauptsächlich auf die Nutzung von maritimen Sensordaten fokussiert, konkret AIS, Sonar, und Radar.

## 3. WISSENSMANAGEMENT

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 den Einsatz semantischer Technologien. In EMSec wird ein „Realtime Maritime Information System“ (RMSAS) entwickelt. Ein zentraler Bestandteil von RMSAS ist ein Enterprise Service Bus (ESB). Der RMSAS ESB schafft die notwendige Infrastruktur zur Erhebung und Verteilung der Daten aus den einzelnen Wissensdomänen. Damit im zweiten Schritt eine maschinelle Verarbeitung und Auswertung dieser Daten erfolgen kann, müssen Daten auf einheitliche, standardisierte und strukturierte Weise erfasst und um Informationen über ihre Bedeutung

(Semantik) erweitert werden. Dies kann in Form von semantischen Datenmodellen erfolgen, wie z.B. Ontologien und den zugehörigen Wissensrepräsentationssprachen, welche das Wissen einer Anwendungsdomäne entsprechend modellieren.

In BILD 1 wird deutlich, wie aus Daten Informationen werden, wenn man die Daten um ihre Bedeutung anreichert. Ebenso wird deutlich, wie aus den Informationen Wissen wird, wenn man die Informationen in einen Kontext setzt und mit Semantik versieht.

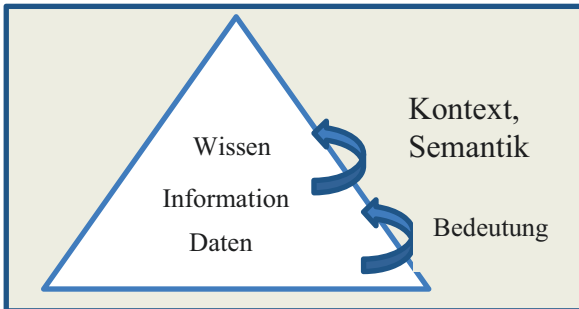


BILD 1. Begriffshierarchie

Werden Informationen systematisch miteinander und mit Daten aus beliebigen Quellen verknüpft, wird Wissen erzeugt. Mit diesem Wissen lassen sich dann Gegebenheiten analysieren, Prognosen erstellen und Handlungsanweisungen erzeugen.

Im folgenden Abschnitt werden Beispiele offener Daten beschrieben, die genutzt werden können, um mit Daten des EMSec-Verbundes verknüpft zu werden.

## 4. OFFENE DATEN

Im Folgenden werden mit Govdata.org und Copernicus zwei Beispiele für offene Datenquellen vorgestellt und diskutiert. Daran anschließend werden notwendige Voraussetzungen für die maritime Domäne abgeleitet.

### 4.1. Govdata.org

Das Datenportal govdata.org ist ein Portal, welches im Rahmen eines Forschungsprojektes im Auftrag des Bundesministeriums des Innern entwickelt wurde. Dieses Portal hat zur Aufgabe, die Daten, die in Deutschland behördlich erfasst und der Öffentlichkeit zugänglich gemacht werden dürfen, zentral zur Verfügung zu stellen. Das Ziel dabei ist, dass diese Daten von Bürgern, Forschungseinrichtungen und der Industrie genutzt und weiterverwendet werden können. So können durch die geeignete Kombination von Daten gegebenenfalls neue Erkenntnisse gewonnen werden.

Die Daten, die bei govdata.org zur Verfügung stehen, stammen aus fast allen deutschen Bundesländern und kommen aus den unterschiedlichsten Domänen. So finden sich Daten zur Bevölkerungsentwicklung, aus Bildung und Wissenschaft, geographische Daten, Daten zur Infrastruktur, Daten zu Transport und Verkehr und zu Wirtschaft und Arbeit.

Für die maritime Sicherheit ist govdata.org im Rahmen von EMSec nur begrenzt einsetzbar, da die angebotenen Datensätze meist nur in menschenlesbarer Form, aber

nicht maschinenverarbeitbar angeboten werden. Konkret stehen viele Daten in Form von Berichten bspw. als PDF-Dokumente oder als Bild zur Verfügung, selten jedoch bspw. als XML- oder CSV-Dokument.

### 4.2. Copernicus

Copernicus<sup>1</sup> ist das europäische Programm, welches eine Infrastruktur zur Erdbeobachtung schafft und Services anbietet, die zur Geoinformation genutzt werden können. Dabei nutzt Copernicus sowohl Daten, die im Weltraum erhoben werden, als auch welche, die in situ erfasst werden. Für die Erfassung von Daten aus dem Weltraum heraus wurden und werden verschiedene Sentinel-Satelliten gestartet, die jeweils unterschiedliche Instrumente tragen. Beispiele für solche Instrumente sind welche zur Messung des Meeresspiegels, zur Erfassung von Spurengasen, und zur Beobachtung von Land- und Ozeanoberflächen.

In-situ-Daten sind Daten aus meteorologischen Einrichtungen, von Wetterballonen, Messbojen, Flusspegel und Fernerkundungsinstrumente. Hieraus lassen sich beispielsweise topographische Karten, Höhenmodelle von Deichen, Übersichten von Zufahrtsstraßen, Gewässerbeschaffenheiten, und Bevölkerungszahlen erzeugen.

Das nationale Verbundprojekt DeMarine-2<sup>2</sup> verfolgte das Ziel, die marinen Copernicus-Dienste nachhaltig zu nutzen. Dabei wurde besonders auf das Monitoring des Seegangs, des Wattenmeers und der Eistätigkeit fokussiert.

Hervorzuheben ist, dass die Daten des Copernicus-Programms online zur Verfügung stehen und für verschiedenste Zwecke genutzt werden können.

### 4.3. Implikationen

Die genannten Beispiele für offene Daten zeigen, dass es für die Nutzung in der maritimen Domäne wichtig ist, dass sich Daten maschinenverarbeitbar nutzen lassen. Dies ist im Bereich des Semantic Web [8], insbesondere durch Linked Open Data [7], gegeben, so dass dies im Folgenden eingeführt wird.

## 5. SEMANTIC WEB

Das Semantic Web [8] erweitert das Web, um Daten zwischen Rechnern einfacher austauschbar und für sie einfacher verwertbar zu machen; so kann beispielsweise der Begriff „Bremen“ in einem Webdokument um die Information ergänzt werden, ob hier ein Schiffs-, Familien- oder der Stadtname gemeint ist. Diese zusätzlichen Informationen explizieren die sonst nur unstrukturiert vorkommenden Daten. Zur Realisierung dienen Standards zur Veröffentlichung und Nutzung maschinenlesbarer Daten, insbesondere das Resource Description Framework (RDF).

Während Menschen solche Informationen aus dem gegebenen Kontext schließen können (aus dem Gesamttext, über die Art der Publikation oder der Rubrik in selbiger, Bilder, etc.) und derartige Verknüpfungen unbewusst aufbauen, muss Maschinen dieser Kontext erst

<sup>1</sup> <http://www.d-gmes.de>

<sup>2</sup> <http://www.demarine.de/lr/>

beigebracht werden; hierzu werden die Inhalte mit weiterführenden Informationen verknüpft. Das Semantic Web beschreibt dazu konzeptionell einen „Giant Global Graph“ (engl. ‚gigantischer globaler Graph‘). Dabei werden sämtliche Sachen von Interesse identifiziert und mit einer eindeutigen Adresse versehen als Knoten angelegt, die wiederum durch Kanten (ebenfalls jeweils eindeutig benannt) miteinander verbunden sind. Einzelne Dokumente im Web beschreiben dann eine Reihe von Kanten, und die Gesamtheit all dieser Kanten entspricht dem globalen Graphen.

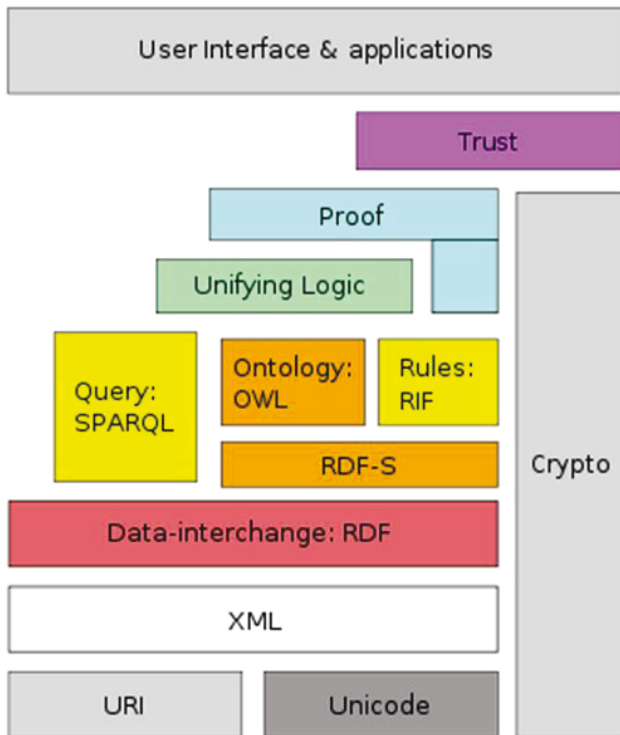


BILD 2. Semantic Web Stack

Das Konzept des Semantic Web beruht auf einem Vorschlag von Tim Berners-Lee, dem Begründer des World Wide Web: das Semantic Web ist eine Erweiterung des herkömmlichen Webs, in der Informationen mit eindeutigen Bedeutungen versehen werden, um die Arbeit zwischen Mensch und Maschine zu erleichtern („The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation“<sup>3</sup>). Das Semantic Web baut auf existierenden Web-Standards und Arbeiten im Bereich Wissensrepräsentation auf. Der in BILD 2 gezeigte sogenannte Layer Cake zeigt die verwendeten Standards: Als Basis dienen Unicode und URIs, die eindeutige Bezeichner für Informationen darstellen. Darüber wird XML als Markierungssprache verwendet und darüber RDF als gemeinsames Datenmodell zur Repräsentation von Aussagen. RDF-S dient zur Deklaration des Vokabulars für RDF. Darüber aufbauend wird OWL zur formalen Definition des in RDFS deklarierten Vokabulars in einer Ontologie verwendet. RIF dient zur Definition und zum Austausch von Regeln in Ontologien; SPARQL ermöglicht die semantische Abfrage des in Ontologien enthaltenen Wissens.

<sup>3</sup> <http://www.w3.org/2001/sw/EO/points>

## 5.1. Linked Open Data

Linked Open Data bezeichnet im Internet frei verfügbare Daten, die mit einer URI identifiziert sind. Somit sind diese Daten direkt abrufbar und können selbst wieder auf andere Daten verweisen. Erstellt sind diese Daten meist mit RDF oder OWL und bilden somit einen Teil des Semantic Web. Tim Berners-Lee erstellte vier Grundregeln für Linked Open Data:

- 1) Verwende zur Bezeichnung von Objekten URIs.
- 2) Verwende HTTP-URIs, so dass sich die Bezeichnungen nachschlagen lassen.
- 3) Stelle zweckdienliche Informationen bereit, wenn jemand eine URI nachschlägt (mittels der Standards RDF und SPARQL).
- 4) Zu diesen Informationen gehören insbesondere Links auf andere URIs, über die weitere Objekte entdeckt werden können.

In BILD 3 ist die Struktur der Linked Open Data Cloud erkennbar. Jeder Kreis repräsentiert eine offene Datenquelle, die Pfeile dazwischen deuten an, dass die jeweiligen Quellen einander referenzieren. Hellblau sind alle Quellen aus der „Media“-Kategorie, gelb sind „Geography“-Quellen, „Publications“ sind hellgrün, „user generated content“ ist hellrot, „Government“ ist blaugrün, „Cross-domain“ ist blaugrau, und Quellen der Kategorie „Life Sciences“ sind pink.

Jeder einzelne Knoten der Linked Open Data Cloud stellt Informationen zur Verfügung, die separat zugreifbar und abrufbar sind. Manche Knoten bieten eigene Programmierschnittstellen an, andere Knoten machen ihre Informationen über einen SPARQL-Endpoint zugänglich. Darüber hinaus enthalten die meisten Informationen in einem Knoten wiederum URIs auf Informationen in einem anderen Knoten.

### 5.1.1. SPARQL-Endpoints

Ein SPARQL-Endpoint stellt RDF-Daten, die in einer Ontologie definiert sind, über ein standardisiertes Protokoll zur Verfügung. Somit sind die Instanzen einer Ontologie zugreifbar und lassen sich analysieren, indem SPARQL-Abfragen an einen SPARQL-Endpoint gesendet werden. In der Regel ist jeder Knoten der Linked Open Data Cloud über einen SPARQL-Endpoint zu erreichen. Beim Entwurf und der Implementierung eigener Systeme werden in der Regel auch eigene SPARQL-Endpoints definiert, die dann über web-basierte Technologien zugänglich sind.

### 5.1.2. DBpedia

Mit <http://dbpedia.org> steht der Inhalt der Wikipedia in maschinenlesbarer Form zur Verfügung. Dazu wurde eine Struktur definiert, die es ermöglicht, Anfragen an Wikipedia zu stellen. So ist es mit DBpedia beispielsweise möglich, alle Personen zu finden, die im 19. Jahrhundert in Bremen geboren sind. Dann kann das „placeOfBirth“-Attribut der DBpedia gemeinsam mit dem Geburtsdatum ausgewertet werden. Als Ergebnis wird dann eine Liste der Wikipedia-Artikel dieser Personen geliefert. Da in der Wikipedia sehr viele Objekte und Personen einen eigenen Eintrag haben, ist es somit möglich, für RMSAS bspw. eine Liste relevanter Objekte für den maritimen Einsatzfall abzufragen. Dies ist dadurch möglich, dass viele



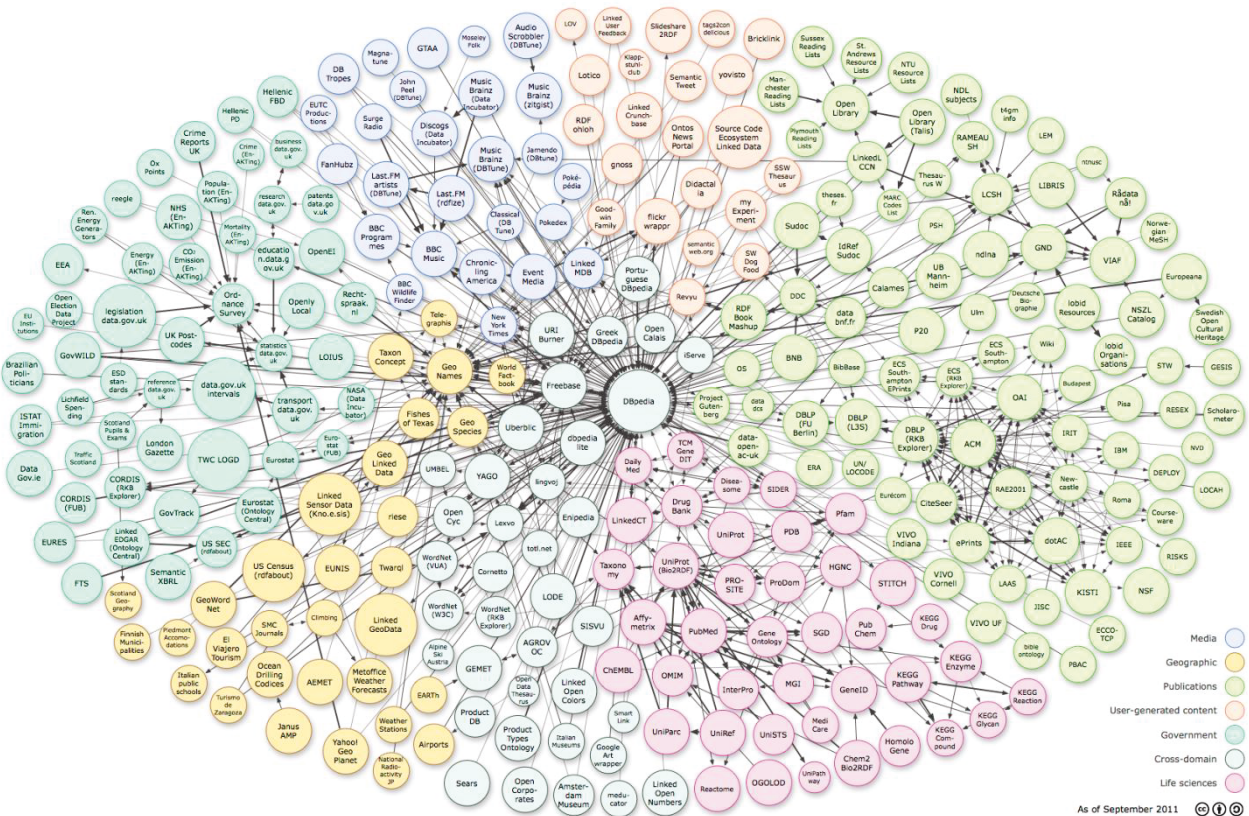


BILD 3. Struktur der Linked Open Data Cloud<sup>4</sup>

geographische Objekte mit Geokoordinaten annotiert sind.

In BILD 4 sind exemplarisch dbpedia-Daten zu einem Offshore-Windpark dargestellt. Die Daten lassen sich über [http://de.dbpedia.org/page/Alpha\\_ventus](http://de.dbpedia.org/page/Alpha_ventus) abrufen. Konkret werden hier Geo-Koordinaten, Luftbilder, Homepage, Betreiber, Gesamteinspeisungen und andere interessante Daten verfügbar gemacht.

owl:sameAs	dbpedia:fr:Parc_eolien_offshore_Alpha_Ventus db:Alpha_Ventus_Offshore_Wind_Farm <a href="http://wikidata.dbpedia.org/resource/Q2645261">http://wikidata.dbpedia.org/resource/Q2645261</a> <a href="http://wikidata.org/entity/Q2645261">http://wikidata.org/entity/Q2645261</a>
geo:lat	54.0 (owl:float) 54.0267 (owl:float)
geo:long	6.5733 (owl:float) 6.62167 (owl:float) 6.62338 (owl:float)
<a href="http://www.kld.org">http://www.kld.org</a>	wiki-de:Alpha_ventus?oldid=131188229
foaf:depiction	<a href="http://commons.wikimedia.org/wiki/Special:FilePath/2012-05-13_Alpha_ventus_-_Norwegian-Lubster_DSCF8001.jpg">http://commons.wikimedia.org/wiki/Special:FilePath/2012-05-13_Alpha_ventus_-_Norwegian-Lubster_DSCF8001.jpg</a>
foaf:homepage	<a href="http://www.alpha-ventus.de/">http://www.alpha-ventus.de/</a>
foaf:isPrimaryTopicOf	wiki-de:Alpha_ventus
foaf:name	alpha ventus NO <span>de:de</span>
is dbpedia:PageRedirect of	dbpedia-de:Bursum_West dbpedia-de:Chirono-Windpark_Bursum_West dbpedia-de:Offshore-Windpark_Bursum_West dbpedia-de:Windpark_Bursum_West
is dbpedia:PageRedirect of	dbpedia-de:April_2010 dbpedia-de:Ö.Ö.N. dbpedia-de:Geschichte_der_Windenergieutzung dbpedia-de:Huber_Nordesteie dbpedia-de:Rawe_(Begriffsklärung)
is foaf:primaryTopic of	wiki-de:Alpha_ventus

BILD 4. Exemplarische Daten der dbpedia zu einem Offshore-Windpark

### 5.1.3. Geonames

Die Geonames-Datenbank <http://geonames.org> enthält über 10 Millionen geographische Namen, die über neun Millionen topographischen Objekten entsprechen. Diese Objekte sind einer von neun Klassen zugeordnet und mit einem von 645 Codes kategorisiert.

Als interessante Informationen sind dort sowohl latitude, longitude, Höhe über NN, Einwohnerzahl, als auch Postleitzahlen und administrative Unterteilungen gespeichert.

Die neun Klassen sind Administrative, Hydrographic Features (Fluss, See,...), Area Features, Populated Place Features, Road and Railroad Features, Spot Features (Orte, Gebäude), Hypsographic Features (mountain,hill,rock,...), Undersea Features, Vegetation Features (forest,heath,...), und "Undefined".

Für RMSAS besonders relevant sind hier beispielsweise die Klasse und Kategorie „H.HBR“, die aktuell 5205 Häfen enthält sowie die Klasse „L.PRT“, die 899 Anlegeplätze enthält. Ersterer enthält „a haven or space of deep water so sheltered by the adjacent land as to afford a safe anchorage for ships“, letzterer enthält „a place provided with terminal and transfer facilities for loading and discharging waterborne cargo or passengers, usually located in a harbor“. Dies kann in RMSAS beispielsweise genutzt werden, um die Liste der bereits angelaufenen Häfen eines Schiffes zu validieren.

In BILD 5 ist ein Ausschnitt der Geonames-Daten der Elbmündung zu sehen. Farblich kodiert sind die Kategorien und Klassen erkennbar und lassen sich entsprechend ein- und ausblenden.

<sup>4</sup> <http://lod-cloud.net/>

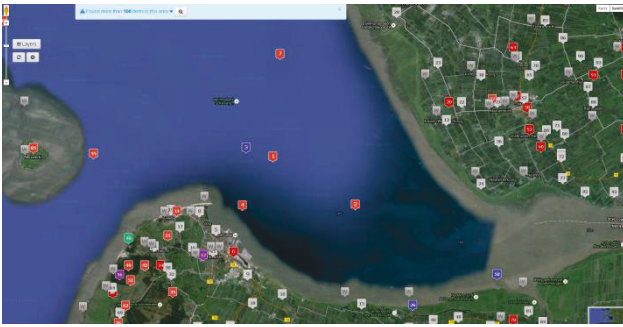


BILD 5. Geonames-Informationen über die Elbmündung

#### 5.1.4. Openstreetmap / LinkedGeoData

In der Linked Open Data Cloud steht mit dem Openstreetmap-Dienst <http://www.openstreetmap.de/> ein Kartendienst zur Verfügung, der von Benutzern frei genutzt werden kann. Damit diese Daten auch maschinenlesbar zur Verfügung stehen, hat <http://linkedgeo.org> diese Daten als RDF-Datenquelle zur Verfügung gestellt. Es verknüpft diese Daten gemäß der oben erwähnten Linked-Data-Prinzipien mit Daten aus anderen Quellen, wie bspw. Geonames und DBpedia. Aktuell stehen eine Milliarde Knoten und 100 Millionen Wege zur Verfügung. Der durch LinkedGeoData resultierende RDF-Graph enthält 20 Milliarden Tripel. Diese durch Openstreetmap bzw. LinkedGeoData zur Verfügung gestellten Daten lassen sich sowohl remote über das Internet nutzen als auch lokal in eine Datenbank importieren. In beiden Fällen stehen Daten zur Verfügung, die beispielsweise in maritimen Lageanwendungen genutzt werden können.

Im folgenden Abschnitt werden Möglichkeiten gezeigt, wie diese offenen Daten in der maritimen Domäne verwendet werden können.

### 6. VERWENDUNGSMÖGLICHKEITEN IN DER MARITIMEN SICHERHEIT

In diesem Abschnitt werden einige Möglichkeiten gezeigt, wie offene Daten in der maritimen Domäne gewinnbringend eingesetzt werden können. Einige Szenarien wurden beispielhaft implementiert, sollen aber nur generelle Konzepte aufzeigen und nicht als vollständig implementierte Beispiele dienen.

#### 6.1. Verwendung offener Geodaten

BILD 6 zeigt, wie Openstreetmap-Daten in eine Lageanwendung integriert werden können. Hier ist das semantische Datenanalysewerkzeug Sextant<sup>5</sup> verwendet worden, um Openstreetmap-Daten anzuzeigen. Mit Sextant ist es möglich, beliebige SPARQL-Endpoints einzubinden und per SPARQL-Abfrage zu analysieren. So lassen sich mit Sextant thematische Karten erstellen und editieren, indem Daten aus der Linked Data Cloud mit anderen Daten, bspw. KML, JSON und GeoTIFF kombiniert werden. Weiter lassen sich Daten temporal analysieren, so dass sich beispielsweise die Route eines oder mehrerer Schiffe analysieren lässt.

<sup>5</sup> <http://hg.strabon.di.uoa.gr/Sextant-New>

Konkret wurden hier die Positionen von Objekten des Typs "Harbour" aus der OpenStreetmap abgefragt, um Häfen in Sextant anzeigen zu können. Dies kann sowohl mit Daten in einer lokalen Datenbank als auch mit Daten in einer entfernten SPARQL-Endpoint-Anwendung.

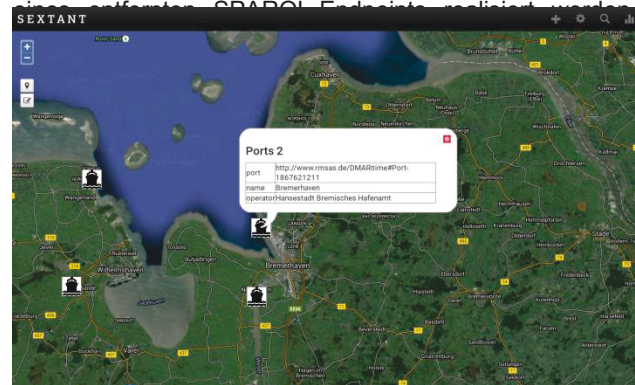


BILD 6. Nutzung von Openstreetmap-Daten

In BILD 7 sind die Schiffspositionen einer Fähre dargestellt, die zwischen Langeoog und Benseniel verkehrt. Wird diese Fährverbindung mit entsprechenden semantischen Informationen annotiert, ist es möglich, Auffälligkeiten wie bspw. Kursabweichungen zu identifizieren. Weiter ist es möglich, solche Routinefahrten in einer Lageanwendung explizit ein- oder auszublenden, um sich nur für eine bestimmte Situation relevante Schiffsdaten anzeigen zu lassen.

#### 6.1.1. Semantische Analyse von Rohdaten

Da Schiffsdaten oftmals als AIS-Datenstrom oder in anderer Rohdatenform zur Verfügung stehen, ist es notwendig, diese Daten auf eine semantische Ebene zu portieren. Dazu bieten sich Ontologien an, die Daten eine Semantik geben. In der maritimen Domäne ist die CISE-Ontologie [12] populär, aber auch andere und eigene Ontologien sind möglich.

Mit Ontologien lassen sich Zusammenhänge in der maritimen Domäne beschreiben, so lässt sich beispielsweise beschreiben, dass ein Schiff eine Besatzung hat, eine Herkunft, einen Reeder, einen bestimmten Hafen anläuft oder eine bestimmte Route gefahren ist. Dies lässt sich verknüpfen mit Informationen über Häfen und ihre Charakteristika wie Kapazitäten, Koordinaten oder Nutzung.

Durch die Verwendung von Ontologien ist es somit möglich, in vorhandenen Rohdaten durch eine geeignete semantische Modellierung Mehrwertinformationen zu finden.

Um Rohdaten in einer ontologiebasierten Form analysierbar zu machen, eignen sich Frameworks und Werkzeuge, die ein Mapping zwischen den Rohdaten und der Ontologie herstellen. Dazu ist beispielsweise das Ontop<sup>6</sup>-Framework in der Lage. Dieses ermöglicht es, Rohdaten, die in einer Datenbank gehalten werden, als SPARQL-Endpoint zur Verfügung zu stellen, wo sie sich dann semantisch analysieren lassen können.

<sup>6</sup> <http://ontop.inf.unibz.it>



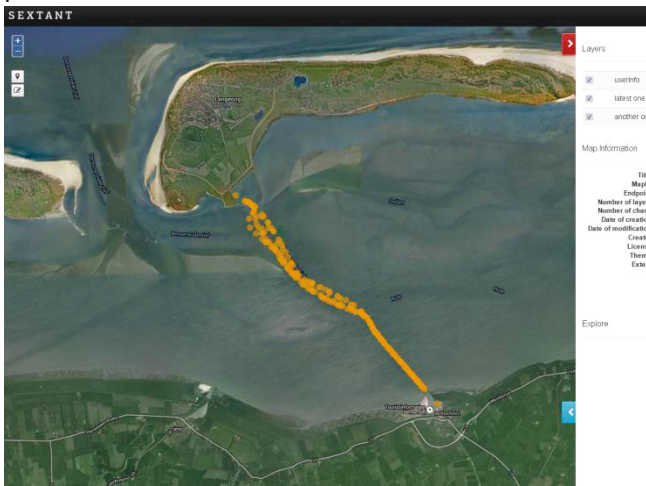


BILD 7. Zeitlicher Verlauf einer Fähre

### 6.2. Verifikation von Schiffsbewegungen

Schiffe, die deutsche Gewässer anlaufen, werden in der Regel erkennungsdienstlich behandelt und es wird entschieden, ob diese Schiffe beispielsweise mit polizeilichen Maßnahmen untersucht oder zolltechnisch behandelt werden. Um präventiv agieren zu können, wird eine manuelle Kategorisierung der Schiffe durchgeführt, die auf Grund definierter Eigenschaften relevant für die maritime Sicherheit sind. Für die sich im jeweiligen Hoheitsgewässer befindlichen oder zu erwartenden Schiffe werden die für das sog. Schiffsprofil notwendigen Schiffsinformationen ermittelt, so dass die Schiffsanmeldungen nach den zuvor definierten Kriterien analysiert und gefiltert werden können, um eine potentielle Relevanz für die maritime Sicherheit rechtzeitig bemerken und entsprechend reagieren zu können [11].

Das Schiffsprofil überprüft unter anderem die Liste der zuletzt angelaufenen Häfen. Hierbei helfen offene Verzeichnisse von Häfen, etwa die Geonames-Kategorien „H.HBR“ und „L.PRT“, welche weltweit Häfen kennzeichnen und klassifizieren, so dass sich hier Informationen und Kapazitäten dieser Häfen abfragen lassen. So kann eine Konsistenzprüfung durchgeführt werden, in der die zu einem Schiff angegebenen Häfen überprüft werden.

In einer Ontologie können Häfen bestimmten Kategorien zugeordnet werden, die beispielsweise Risikofaktoren abbilden. Dies kann genutzt werden, um mit Hilfe eines Reasoners Schiffsanmeldungen zu klassifizieren. So lassen sich Kategorien nach Staaten, Ereignissen, politischen Dimensionen und anderen Eigenschaften kategorisieren.

### 6.3. Ausbreitung von Epidemien

Die Überprüfung bereits angelaufener Häfen betrachtet immer auch die Ausbreitung von Krankheiten in den jeweiligen Häfen. Im Falle von Ebola beispielsweise gelten die Häfen eines Landes als nicht vertrauenswürdig, wenn in dem Land Vorkommen von Ebola gemeldet wurden. Da Epidemien aber nicht an Staatsgrenzen Halt machen, sind auch die Häfen in der Nähe eines Ausbruchs relevant, die

in angrenzenden Staaten liegen. In diesem Fall lassen sich offene Daten wiederum nutzen, um die Menge aller Häfen in der Nähe zu bestimmen. Dies ist beispielsweise mit Geonames dadurch möglich, dass dort Häfen geoannotiert sind. Eine SPARQL-Abfrage kann nun geographische Funktionen wie beispielsweise eine „NEARBY“-Abfrage nutzen, um Häfen, die innerhalb eines definierten Radius um einen Epidemie-Ausbruch liegen, entsprechend auch zu kategorisieren.

Dieses Beispiel zeigt, dass sich Risikofaktoren nicht nur staatenbezogen definieren lassen, sondern dass ein regionales Clustering ermöglicht wird. So lassen sich Ausbrüche von Krankheiten effektiver verwalten und vermeiden.

### 6.4. Brand im Hamburger Hafen

Im folgenden Abschnitt wird ein Szenario beschrieben, bei dem im Falle eines maritimen Notfallereignisses wie einem Brand, beispielsweise in einem Hafen, offene Daten genutzt werden können, um die allgemeine Situation einschätzen zu können. Gegeben sei ein Ereignis in der Nacht, bei schlechtem Wetter und ohne weitere Kenntnisse der Umgebung oder des havarierten Schiffes. Da vom unter fernöstlicher Flagge fahrenden Schiff nur die MMSI-Nummer bekannt ist, sind Kapitän, Reeder, Ladung und weitere Informationen zunächst unbekannt. Über eine offene Datenquelle, in der statische Informationen zum Schiff gespeichert sind, lassen sich diese Angaben aber ermitteln. Der Typ der Ladung ist schnell als Gefahrgut identifiziert. BILD 8 zeigt das havarierte Schiff in einer Lageanwendung. Die Auswirkungen des Brandes auf dieses Schiff sind für die weitere Umgebung zunächst unklar, da keine Detailinformationen über die Umgebung vorliegen. Somit ist nicht klar, ob beispielsweise Autobahnen durch Rauch behindert oder Personen durch giftige Dämpfe gefährdet sind. Weiter ist nicht klar, ob Schiffe in der Nähe sind, die ebenfalls betroffen sind.

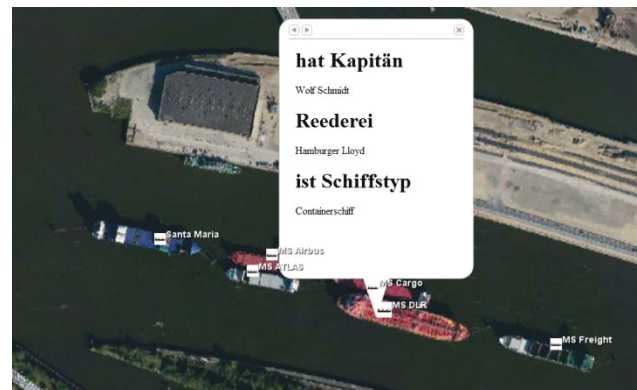


BILD 8. Havariertes Schiff

Die Nutzung der Geonames-Ontologie ermöglicht es nun, kritische Infrastruktur in der Nähe zu identifizieren. Dies wird dadurch ermöglicht, dass die "Nearby"-Funktionalität von SPARQL genutzt werden kann, um Dinge in der Nähe eines Punktes zu identifizieren. In der Geonames-Datenbank sind Orte und Einrichtungen entsprechend georeferenziert und so liefert eine Nearby-Abfrage zum Beispiel einen Kilometer um ein havariertes Schiff herum Infrastruktur, die in der Nähe ist. BILD 9 zeigt diese Infrastruktur in der Nähe des Schiffes.

Um weitere Entscheidungen treffen zu können, ist es wichtig herauszufinden, welcher Art diese kritische Infrastruktur ist. Dazu wird ausgenutzt, dass die Daten in der Geonames-Datenbank Referenzen auf die DBPedia-Datenbank enthalten, welche wiederum die Wikipedia-Daten nutzt. So lässt sich für jeden Eintrag aus BILD 7 anzeigen, um welche Einrichtung es sich handelt.

In BILD 10 ist exemplarisch ein Wikipedia-Eintrag zu einer kritischen Infrastruktur eingebildet. Im konkreten Fall ist identifiziert worden, dass die Havarie in unmittelbarer Nähe zur Hamburger Speicherstadt liegt. Maritime Entscheider können nun weitere Schritte planen und einleiten. Die Hamburger Speicherstadt ist sicherlich ohne weiteres bekannt, noch relevanter werden solche Analysen, wenn Dinge gefunden werden, die nicht unmittelbar bekannt sind, seien es Pipelines, Einrichtungen von Energieversorgern oder andere besonders schützenswerte Infrastruktur.

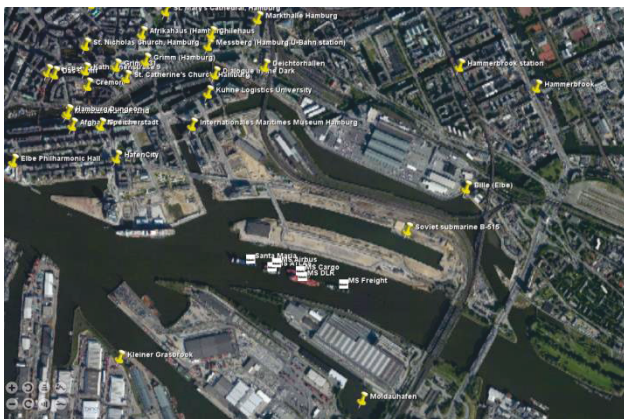


BILD 9. Kritische Infrastruktur in der Nähe eines havarierten Schiffs

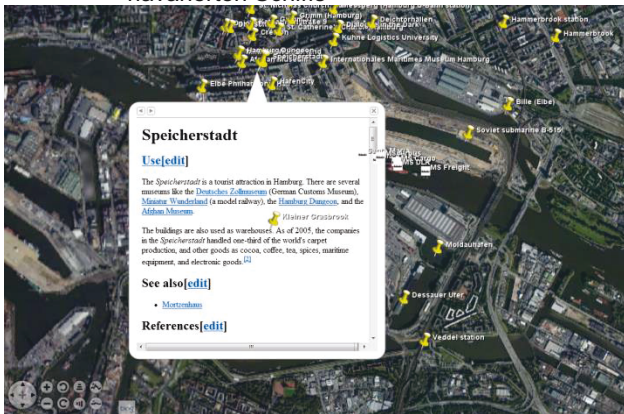


BILD 10. Wikipedia-Eintrag zu einer kritischen Infrastruktur

### 6.5. Satelliten-Beams aus TLE-Daten

Im Falle einer komplexen Schadenslage sind Satellitenbilder oftmals sehr hilfreich. Je nach Nutzlast kann ein Satellit beispielsweise Bilder einer Oberfläche produzieren, aus denen sich Informationen zur aktuellen Lage ableiten lassen:

- Im Falle eines Deichbruchs können sich Einsatzkräfte so über die Beschaffenheit der Infrastruktur informieren und Anfahr- und Evakuierungspläne aufstellen.
- Im Falle einer Havarie lässt sich die Schadenslage bewerten, so dass Einsatzleiter die notwendigen Schritte einleiten können.
- Die Daten der Verbunde SafeSeaNet [4] lassen sich zum Monitoring des Schiffsverkehrs und bei CleanSeaNet [5] zur satellitenbasierten Detektion von Verschmutzungen und Schiffen nutzen.

Allen Szenarien gemeinsam ist, dass Entscheider rechtzeitig wissen müssen, ob Satelliteninformationen rechtzeitig zur Verfügung stehen könnten. Dazu ist es nötig, die Bahnen und Positionen von Satelliten vorab zu bestimmen, um diese im konkreten Schadensfall nutzen zu können.

Für die meisten Satelliten werden vom amerikanischen Air Force Space Command Satellitenbahnelemente zur Verfügung gestellt, die als offene Daten nutzbar sind. Diese liegen in der Form von Two Line Elements<sup>7</sup> (TLE) vor, mit denen sich Lage, Position des Satelliten und Form der Umlaufbahn bestimmen lassen.

Aus diesen Informationen lässt sich dann ableiten, ob ein Satellit zu einem bestimmten Zeitpunkt an einem definierten Ort gewesen ist oder ob er in einem benötigten Zeitfenster an einer bestimmten Position sein kann. Weiter wird dann deutlich, ob er die benötigten Instrumente zur Verfügung hat, um die verlangten Daten zu erheben und eine geeignete optische Aufnahme machen zu können.

In BILD 11 ist exemplarisch der Beam eines Satelliten skizziert. Dies bedeutet, dass ein Instrument eines Satelliten eine bestimmte Funktionalität hat, die auf einen Teil der Erdoberfläche angewendet werden kann, in diesem Beispiel die hochpräzise Landvermessung mit TanDEM-X<sup>8</sup>.

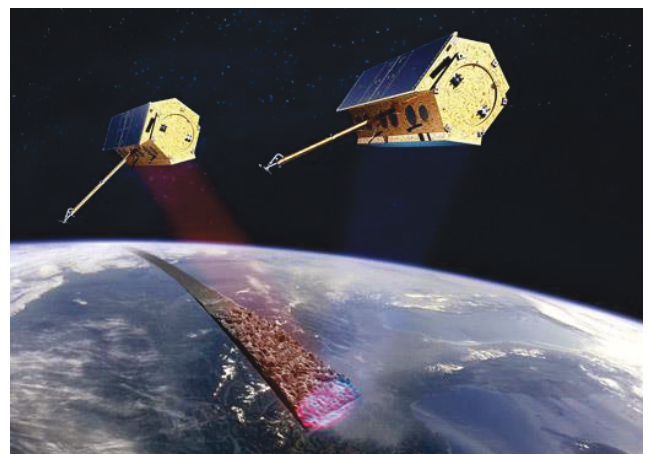


BILD 11. Beam eines Satelliten<sup>9</sup>

<sup>7</sup> <http://www.celestrak.com/NORAD/elements/>

<sup>8</sup> <http://www.space-airbusds.com/de/programme/tandem-x-Geo.html>

<sup>9</sup> <http://www.dailygalaxy.com>



## 6.6. Wetterdaten

In der maritimen Sicherheit ist es von besonderer Bedeutung, die aktuelle und künftige Wetterlage einschätzen zu können. Es existieren diverse Wetterinformationsdienste, die Daten in unterschiedlicher Genauigkeit und mit verschiedenartiger Eignung für die maschinelle Verarbeitung anbieten. So existieren Daten des Deutschen Wetterdienstes<sup>10</sup>, von OpenWeatherMap<sup>11</sup> oder von den Nasa Earth Observations<sup>12</sup>. All diese Dienste liefern verschiedenste Daten unterschiedlichster Quellen. NASA bspw. liefert Ozeographie-Daten, die Daten über Wassertemperatur, Eiskonzentrationen, biologische Faktoren sowie Anomalien enthalten.

Die Dienste von Openweathermap liefern beispielsweise Satellitenbilder, die frei abrufbar sind<sup>13</sup>. Diese Daten sind in Nahe-Echtzeit verfügbar. Weiter liefert OpenweatherMap Vorhersage-Daten zu Temperaturen, Windbewegungen und anderen meteorologischen Daten.

Insgesamt bieten diese Wetterinformationsdienste Daten an, die sich in der maritimen Domäne nutzen lassen, um beispielsweise automatisiert in einer Lageanwendung angezeigt werden zu können. Dies ermöglicht es den Anwendern, im Havariefall die geeigneten Maßnahmen einzuleiten. Beispielsweise haben Schiffe mit einem bestimmten Tiefgang wetterabhängig nicht die Möglichkeit, gewisse Bereiche zu passieren oder zu verlassen. Sich ändernde Wassertiefen müssen hierbei genauso berücksichtigt werden wie Tide-Informationen. In der Einsatzphase einer maritimen Schadenslage lassen sich so evtl. Schiffe aus einer Lage herausfiltern, die nicht geeignet sind, bei einer Havarie zu helfen.

## 7. ZUSAMMENFASSUNG

In diesem Beitrag wurde gezeigt, wie sich offene Daten in der Domäne der maritimen Sicherheit nutzen lassen können. Dazu wurde zunächst mit dem EMSec-Verbundprojekt der maritime Kontext definiert. Danach wurde gezeigt, in wie weit aus Daten und Informationen Wissen generiert werden kann, damit es von Entscheidern in der maritimen Welt genutzt werden kann. Daraufhin wurden verschiedene offene Datenquellen vorgestellt. Ein besonderer Schwerpunkt wurde dabei auf die Nutzung der Datenquellen der Linked Open Data Cloud gesetzt, da diese Daten enthält, die sich besonders gut automatisiert nutzen und beispielsweise in Lageanwendungen einbinden lassen. Ein Schwerpunkt des Beitrags lag auf der semantischen Annotation von Daten; auch dies ein wesentliches Merkmal der Linked Open Data Cloud, in der Daten mit Metadaten versehen sind, um den Daten einen Kontext und eine Bedeutung zu geben.

Von den vorgestellten offenen Datenquellen ausgehend wurden dann verschiedene Anwendungsfälle für diese Daten in der maritimen Domäne vorgestellt. Konkret wurde gezeigt, wie sich geo-basierte Rohdaten semantisch annotieren und analysieren lassen, wenn sie in eine ontologiebasierte Form gebracht werden.

<sup>10</sup> <http://www.dwd.de>

<sup>11</sup> <http://openweathermap.org/>

<sup>12</sup> [http://neowms.sci.gsfc.nasa.gov/dataset\\_index.php](http://neowms.sci.gsfc.nasa.gov/dataset_index.php)

<sup>13</sup> <http://openweathermap.org/news/post/new-interactive-maps-satellite-images/>

Es wurde weiter gezeigt, wie offene Daten das Schiffsprofiling unterstützen können. Dabei werden vorhandene Daten, beispielsweise Metadaten zu Häfen, genutzt, um die von Schiffen gemachten Angaben zu validieren und um die Schiffe nach Gefahr und anderen Relevanzen zu klassifizieren.

Weiter wurde mit dem Ausbruch von Epidemien ein Anwendungsfall vorgestellt, bei dem geographische Informationen die Risikoklassifikation von Herkunftshäfen optimieren können.

Mit einem fiktiven Brand wurde dann ein Szenario vorgestellt, in welchem unterschiedliche Informationsquellen die Einsatzphase unterstützen können. So können Daten über sich in der Nähe befindende interessante Orte genutzt werden, um diese im Bedarfsfall rechtzeitig evakuieren zu können.

Frei verfügbare Satellitendaten lassen sich nutzen, um bestimmen zu können, ob und wann ein Satellit genutzt werden kann, um zusätzliche Informationen einer maritimen Schadenslage zu gewinnen. Hierbei lässt sich bestimmen, wann ein Satellit ein bestimmtes Gebiet überfliegt oder überfliegen hat. Dazu kann weiter herausgefunden werden, ob ein solcher Satellit geeignete Instrumente als Nutzlast trägt, die dann entsprechend einen Beitrag zur Bewältigung einer Schadenslage leisten können.

Schließlich wurden diverse Wetterdienste vorgestellt und es wurde gezeigt, wie mit diesen Diensten Mehrwertinformationen für Entscheider generiert werden können.

## 8. FAZIT

Interviews mit Experten der maritimen Domäne [1] haben gezeigt, dass Entscheider sowohl im Routinebetrieb als auch während einer maritimen Schadenslage einen wachsenden Bedarf an vielfältigen Informationen haben. Diese Informationen stammen nicht nur aus den, oft streng zugangsbeschränkten, internen Datenbanken der Behörden der maritimen Sicherheit, sondern lassen sich oft auch aus offenen Datenquellen extrahieren. Besonders statische Daten wie geographische Informationen und dynamische Daten, wie Wetterdaten, Driftmodelle oder die Verfügbarkeit von Satelliten lassen sich so bereitstellen. Die Informationssammlungen der Linked Open Data Cloud bieten sich hierzu besonders an, da diese sowohl semantisch annotiert als auch maschinenlesbar und in eigene Programme integrierbar sind.



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] Open Data Network - Deutschland - Netzwerk zur Förderung von Open Data, Open Government, Transparenz und Partizipation: [opendata-network.org](http://opendata-network.org)
- [3] Govdata.org: Das Datenportal für Deutschland
- [4] SafeSeaNet, <http://www.emsa.europa.eu/ssn-main.html>
- [5] CleanSeaNet, <http://www.emsa.europa.eu/operations/cleanseanet.html>
- [6] Copernicus:  
[http://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Copernicus/Free\\_access\\_to\\_Copernicus\\_Sentinel\\_satellite\\_data](http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Free_access_to_Copernicus_Sentinel_satellite_data)
- [7] Christian Bizer, Tom Heath and Tim Berners-Lee. Linked Data - The Story So Far. International Journal on Semantic Web and Information Systems, Special Issue on Linked Data, 2009
- [8] Berners-Lee, Tim, James Hendler, and Ora Lassila. "The semantic web." Scientific american 284.5 (2001): 28-37.
- [9] Kazemi, S., Abghari, S., Lavesson, N., Johnson, H., & Ryman, P. (2013). Open data for anomaly detection in maritime surveillance. *Expert Systems with Applications*, 40(14), 5719-5729.
- [10] Cetin, F. T., Yilmaz, B., Kabak, Y., Lee, J.-H., Erbas, C., Akagunduz, E., & Lee, S.-J. (2013). Increasing Maritime Situational Awareness with Interoperating Distributed Information Sources. 18th International Command and Control Research and Technology Symposium (pp. 9-22).
- [11] Suhaila Wahab. Verbesserung des Situationsbewusstseins innerhalb der zivilen maritimen Domäne durch semantische Datenmodellierung. Masters Thesis, Hochschule Bremen, 2015.
- [12] 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.