Grant Agreement No: 687591

01/04/16

Big Data Analytics for Time Critical Mobility Forecasting

datAcron

D5.1 Maritime Use Case Description

Deliverable Form	
Project Reference No.	H2020-ICT-2015 687591
Deliverable No.	5.1
Relevant Work Package:	WP 5
Nature:	R
Dissemination Level:	PU
Document version:	1.1
Due Date:	01/04/2016
Date of latest revision:	01/04/2016
Completion Date:	01/04/2016
Lead partner:	CMRE
Authors:	Anne-Laure Jousselme, Cyril Ray, Elena Camossi, Melita Hadzagic, Christophe Claramunt, Karna Bryan, Eric Reardon, Michael Ilteris
Reviewers:	
Document description:	This deliverable provides a description of the maritime use case.
Document location:	Documents/datAcron/WP5/Deliverables/Final

HISTORY OF CHANGES

Version	Date	Changes	Author	Remarks
1.0	14/03/2016		CMRE - NARI	Initial version
1.1	01/04/2016		CMRE - NARI	Final version

EXECUTIVE SUMMARY

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

As stated in the Grant Agreement (GA), datAcron "aims to develop novel methods for real-time detection and prediction of trajectories and important events related to moving entities, together with advanced visual analytics methods, over multiple heterogeneous, voluminous, fluctuating, and noisy data streams from moving entities, correlating them with archived data expressing, among others, entities' characteristics, geographical information, mobility patterns, regulations and intentional data (*e.g.* eplanned routes), in a timely manner. The general goal is to increase the safety, efficiency and economy of operations concerning moving entities in the Maritime domain".

Regarding the application to the maritime domain, the algorithms being developed would thus support the detection of abnormal activity in large vessel fleets across large geographical areas, as well as the assessment of any threat for an early intervention. In other words, datAcron would support the decision maker in reaching his[1] level of situation awareness for an informed decision. Situation awareness, as described by Endsley and Garland [3], is the *perception*, *comprehension* and *projection* of the elements of situation in time and space. On the other hand, situation awareness can also be defined as the product of the situation assessment tasks referred as the level 2 of the Joint Directors of Laboratories (JDL) model of information fusion ([23, 10]).

Effective Maritime Situation Awareness (MSA) requires not only detecting, tracking and classifying vessels but also detecting, classifying and predicting their behaviour, which includes detecting relationships between vessels and vessels' behaviours. This challenging and crucial task is at the core of the compilation of a *maritime picture* [8, 7] which involves extracting relevant contextual information (for instance maritime routes or loitering areas [16]) but also monitoring the real time maritime traffic. The use of a set of sensors mixing cooperative self-identification systems such as the Automatic Identification System (AIS) and non-cooperative systems such as coastal radars or satellite imagery provides the necessary complementarity and redundancy of information to overcome the possible and quite common spoofing of AIS signals while increasing the clarity and accuracy of the maritime picture. In many cases, intelligence information can also be helpful in refining and guiding the search in the huge amount of data to be processed, filtered and analysed, as well as representing the contextual information for some MSA problems [11].

Facing the huge *volume* of *various* information with high *velocity* which often lacks *veracity*, a system to automatically process both historical and timely information would greatly support the operator in monitoring and analysis tasks. The maritime use case described in this document emphasises the use of datAcron in a human-centric automatic processing of data, stressing the role of the user (or decision maker) in his interaction with the datAcron system.

While typical scenarios help developers to understand the requirements of a system and how it works, use cases are rather dedicated to "formalising the capture of these scenarios" and documenting the requirements of a system [13]. From a collaborative research perspective, it is expected that use cases would facilitate collaborative research work among several partners and work packages, as it should integrate the different pieces of works to be developed, allow the necessary flexibility to support research findings, and validate the solutions on a practical operational use. As such, it is seen as a bridge between the operational and the scientific community.

We will rely in this document on the previous experience where use cases were designed to support collaborative research on context-based reasoning in high-level information fusion on a ONRG NICOP project [6, 21].

This document is Deliverable 5.1 of the Work Package 5 (WP5) of the datAcron project. It describes a maritime use case in support to the datAcron research objectives. It will also serve as the basis for preparation of the experimentation plan and evaluation criteria (Deliverable 5.2).

[1] In this document, the use of the masculine is solely to simplify the text.

2 METHODOLOGY FOR THE USE CASE DEVELOPMENT

A *use case* describes the interaction of a user with a system to be designed, to achieve a specific goal or accomplish a specific task. The system requirements can then be derived enabling the user to achieve his objectives in different scenarios [13]. The term *Actor* is often used to describe a person or an external system that interacts with the system to be designed. We will be using the term Actor interchangeably as the *user* or the *operator* of the system.

The different *scenarios* to be defined within a use case illustrate different usages of the system, and eventually define success (if the goal is achieved) or failure (if the goal is not achieved) [13].

2.1 Purpose of the use case

The use case definition provides a tool for datAcron to address different aspects of a large research problem, describing users' needs, operational problems and underlying challenges. Illustrating research findings on a common use case, sharing the same datasets, and utilising outputs from other teams are all benefits of having an integrated picture of the general research problem.

As such, the datAcron maritime use case should:

UC Req. 1 Address challenging problems deemed of interest for the maritime operational community in general;

UC Req. 2 Be aligned with the European Union maritime policy and needs in particular;

UC Req. 3 Be aligned with datAcron 's research objectives and expected outcomes such that the use case challenges the datAcron 's technical solutions to be developed, while accommodating the research interests of the different partners;

UC Req. 4 Describe the problem in a simple way as a kind of "skeleton", flexible enough to allow further evolution and developments as possibly requested by partners' interests;

UC Req. 5 Provide the necessary information to understand the user's goal, from which the corresponding sub-goals, associated levels of granularity required, the information needs and the desired output quality can be deduced;

UC Req. 6 Act as an "integrator" for the different aspects to be pursued so that teams can illustrate their findings within a common story;

UC Req. 7 Provide a background and support for close interactions between the different work packages and teams involved with the team in charge of the maritime use case;

UC Req. 8 Rely on the available datasets (unclassified, shareable) among the teams and others of interest in the research community (*e.g. e*, AIS data, radar datasets, databases of past events, intelligence reports, etc).

These requirements may also be used as qualitative system performance metrics, while, at the implementation level, they may act as result validation measures.

More specifically, the main role of the datAcron 's maritime use case is to act as a guide for the research and development. It focuses on relevant practical challenges and operational questions. It describes possible operational uses of datAcron . In the second step, it will help to validate research findings from an end-user point of view and will provide common examples and illustrations of results.

The goal of a use case is thus to stimulate ideas, define new solutions and improve theoretical findings based on practical considerations. It should thus challenge the research, and never be an obstacle to research.

The consortium being composed of complementary and sometimes overlapping expertise, unifying the individual contributions may be hard beside any good willingness, fruitful meetings and frequent interactions. If each research team takes care of positioning its scientific contribution within the use case in a meaningful way and relatively to the other members' contribution, the final product will *de facto* be unified.

However the use case does not provide specific values for source performances or prior probabilities, for instance. The latter should be simulated, retrieved through real data, elicited or even invented where needed. Each team has thus the freedom (and even the duty) to develop any missing part useful for its needs, and adapt it to the use case. Neither the use case suggests an internal structure for the system nor a design solution, to avoid putting extra constraint on the system's designer [13].

2.2 datAcron research objectives

"Technological developments are to be validated and evaluated in **user-defined challenges** *focusing on increasing the safety, efficiency and economy of operations concerning moving entities in the Air-Traffic Management and Maritime domains."*[2].

The main objectives of datAcron are the development of highly scalable methods for advancing:

O.1 Spatio-temporal data integration and management solutions;

O.2 Real-time detection and forecasting accuracy of moving entities' trajectories;

O.3 Real-time recognition and prediction of important events concerning these entities;

O.4 General visual analytics infrastructure supporting all steps of the analysis through appropriate interactive visualisations;

O.5 Producing streaming data synopses at a high-rate of compression.

Also,

- All analytics components can take full benefit of the computations of others, also taking advantage of interlinking between their results. Thus, the trajectory detection and forecasting methods can benefit from events detected or forecast and vise-versa. Similarly for the visual analytics methods;
- Users can interact and explore data, via integrated data views, being supported for decision-making;
- Advanced processing of data close to the data sources (following the in-situ data processing paradigm).

2.3 Underlying big data challenges

Maritime Situation Awareness (MSA) requires processing in real-time a high volume of information of different nature (numerical, natural language statements, objective or subjective assessments, ...), originating from a variety of sources (sensors and humans - hard and soft), with

[2]From the abstract of the Grant Agreement

a lack of veracity (uncertain, imprecise, vague, ambiguous, incomplete, conflicting, incorrect, etc). The datAcron algorithms to be designed in support to MSA should cope with these big data challenges and this ability being reflected also in the quality of the results provided (see Section 3.6). Also, these challenges underlying the datAcron objectives should be reflected in the description of the use case. The big data challenges are described in the GA as (see Annex 5.2.1 in the GA):

Variety - *Different types of data are available and only if properly combined and integrated these data can provide useful knowledge. Different sensor technologies are being developed and the data coming from multiple sources need to be cleaned up from inconsistencies, standardised in format and summarised.*

The set of sources of information to be considered should cover the wide range of variety of commonly referred sources of information across the different communities. It should contain in particular (non-exhaustive list):

(a) *Physical sensors* such as Automatic Identification Systems, coastal or on-board radars as traditionally used for tracking objects, Synthetic Aperture Imagery, cameras, ESM on board military ships;

(b) *Automated processors* such as trackers, Automatic Target Recognition algorithms or classifiers in general;

(c) *Human sources* including operators or analysts themselves possibly manipulating lower level data (*e.g.* evideos, radar images) to reflect the chain of information processing, from automation to possible subjective assessments, intelligence reports;

(d) *Databases* as records of past events (*e.g.* epiracy, accidents, illegal fishing activities), records of vessels such as the Lloyds database.

Other sources could be considered as well such as social media, or open-source media. The list and description of the available datasets is given in Section 4.

Veracity - *Data measurements have an intrinsic uncertainty which proper fusion and clustering address to solve in a preparation/preprocessing phase (by assessing the quality of data themselves) and combining measurements from complementary sources.*

For instance, AIS data are incomplete, intermittent, with errors, the signal can be spoofed.

Volume - *The growing number of sensors (in coastal and satellite networks) makes the sea one of the most challenging scenario to be effectively monitored; the need for methods able to scale in time and space the data processing of vessel motion data at sea is highly critical for maritime security and safety.*

Velocity - *The analysis of streaming data from multiple sensors is essential to detect as soon as they occur critical events at sea. This poses the emphasis on incremental clustering able to include new data into the data-at-rest already processed and on sequential methods able to detect critical events by continuously processing data.*

Moreover, contextual information is a crucial aspect to be characterised within the use case. It is a relative notion that depends on the user's goal. For instance, the maritime routes are contextual information for vessel's destination prediction, while the maritime navigation rules are contextual information for the route extraction problem. Consequently, the user's goal highly influences what contextual information is. Following the categorization of contextual information proposed in [18] of User context, Physical context, Network context, Activity context, Device context, Service context, we identify the following dimensions of context to be considered with our use case:

1. *Area of interest*: Assembles all the information that could be extracted relatively to the area of interest. For instance, harbour zones characteristics, such as water depths, channels,

restricted areas, protected areas, fishing areas, borders, harbours (fishing, recreational, etc), shipping lanes, ferry lanes, military and (Liquid Natural Gas) LNG anchorage areas, islands, offshore platforms, etc;

2. *Rules*: Gather together the legislation about navigation or electronic emission such as the AIS transmission or other mandatory reports;

3. *Patterns of life*: Give information about past behaviors, usually followed patterns, routes, etc. It concerns both individual vessels or groups of vessels;

4. *Meteorological conditions*: The METOC information is mainly about the sea state, the weather, the wind;

5. *Traffic density*: This is a current contextual information deduced from the number of vessels in a given area;

6. *Time / period*: This could be either the period of the day (night versus day, morning versus afternoon), but also seasonal information;

7. *User*: The user's characteristics (mission, goal, etc) are considered as part of the context and are crucial components. Indeed, information needs are directly derived from his role or mission. The user's characteristics include:

   (a) *Decision maker / Authority*: Role and hierarchical position together with a possible communication network;

   (b) *Mission*: At the current instant in time, from which tasks and sub-tasks are deduced;

   (c) *Decision*: The list of actions at his disposal, the type of decision he can or needs to take and for which the level of situation awareness required.

   (d) *Reaction time*: The time constraints associated to the mission and other contextual elements.

2.4 General methodology for the maritime use case development

The diagram displayed in Figure 1 provides the general idea for developing the use case. The datAcron 's objectives (O.1.1 to O.4.4) of the proposal describe the general goals for the algorithms to be designed. They involve several underlying challenges (UC) or would drive some research foci that may be specific to each partner. In collaboration with the operational partners (US Navy through CMRE, Ecole de la Marine Marchande (ENSM)[3]), we ensure that the use case described is operationally relevant. In particular, the use case describes the general context of use of datAcron algorithms. The operational information needs are captured by the list of relevant Maritime Situational Indicators (MSI) as described in Table 2.

The use case will be supported by the appropriate datasets. From the initial list of datasets provided in the proposal, we will ensure that the data are usable, aligned in time and space, etc. Additional datasets may be identified (see Section 4). The performance criteria specify what the user expects, and are given in Section 3.6. They also help to closely tie the experimental plan to the use case development, resulting in an integrated product. The detailed description of the performance criteria will be the purpose of the experimental plan document, as they will be linked to the criteria identified in the GA.

Figure 2 describes the main components of the maritime use case development, their relations to the datAcron component platform and the experimentation environment for the maritime domain.

[3]Note that the Ecole de la Marine Marchande will perform a second-round validation for the next deliverable.

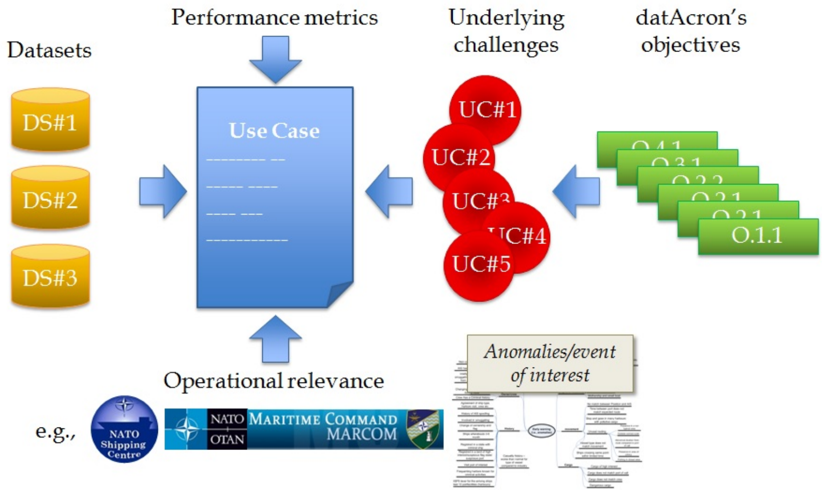


Figure 1: General methodology for the maritime use case development

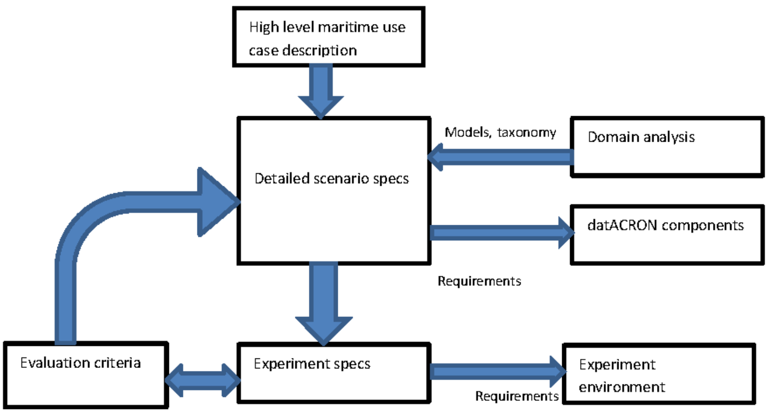


Figure 2: Use case specification components graph

The maritime use case comprises one or more scenarios that describe how actors in the use case perform a set of operations in order to achieve a specific goal. Scenarios describe the current operations that will serve as a base for understanding and validating the datAcron technology, while demonstrating how it can be effectively used in the maritime domain (see Section 3.4).

3 MONITORING FISHING ACTIVITIES

3.1 Maritime surveillance

The general task addressed by the datAcron Maritime Use Case can be referred to as maritime surveillance. A maritime surveillance system (MSS), as depicted in Figure 3[4], is a multi-mission system that is intended to support operators to supervise maritime traffic, to prevent pollution at sea, to control fishing activities, to control borders, to pursue offenders, and to carry out Search And Rescue (SAR) operations. It is usually composed of an aircraft, a set of sensors, a crew and many software artifacts. A MSS is usually a very complex system, because of the large number of possible missions, of the relationships between hardware and software components and of the communication between the system and other entities (base, other MSSs).

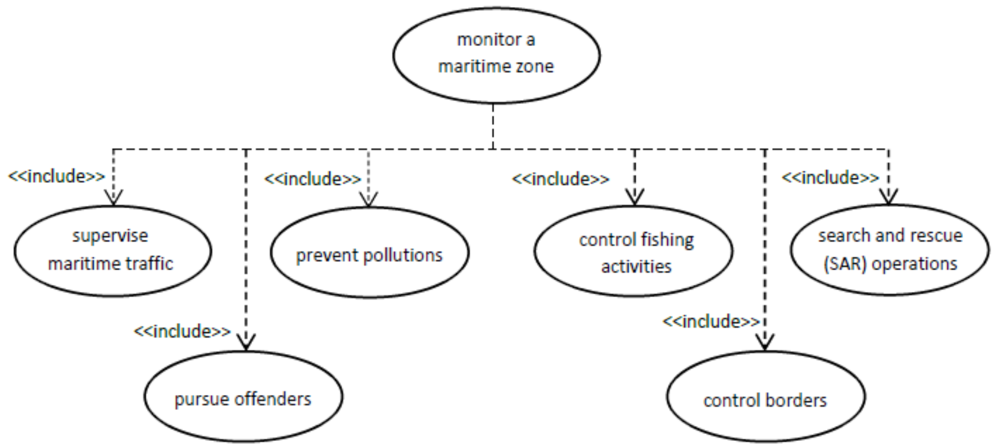


Figure 3: "Use cases of maritime surveillance system" (Excerpt from [14])

While developing the Maritime Use Case for datAcron we will focus on one of the aforementioned missions: specifically, the control of fishing activities. Indeed, fishing activity monitoring fulfills all the requirements identified in Section 2. Moreover, it is a complex mission that encompasses several maritime risks and environmental issues such as environmental destruction and degradation but also maritime accidents, Illegal, Unreported, and Unregulated (IUU) fishing and trafficking problems, which can be described in different scenarios.

datAcron will provide algorithms and analytical tools which would support a user in his analysis, reasoning and decision making tasks in support of the control of fishing activities.

3.2 Motivation for fishing activities monitoring

The European Union Maritime Security Strategy (EUMSS) - Action Plan, published in December 2014[5], defines several strategic interests for the European Union and the Member States in terms of maritime security. datAcron might support the EUMSS ensuring *security* and *control* of fishing activities, which is one of the most important aspect of the strategy. Europe as the world's

[4] Figure 3 is a UML Use Case diagram according to the UML 2.2 specification and notation.

[5] EUMSS Action Plan: http://ec.europa.eu/maritimeaffairs/policy/maritime-security/doc/20141216-action-plan_en.pdf

biggest market for seafood wants to promote better international governance across the world's seas and oceans to keep them clean, safe and secure. Since fishing is an activity that exploits common natural resources, it needs to be regulated to safeguard fair access, sustainability and profitability for all, addressing:

- Total Allowable Catches;
- Fishing licenses;
- Boat capacity management;
- Minimum fish and mesh sizes;
- Design and use of gears;
- Closed areas or seasons;
- Reducing environmental impact.

In order to support this strategy, the European maritime and fisheries fund for the period 2014-2020 will reach 6400 million of Euros[6], making it one of the top five European Structural and Investment Funds. The European Union reinforced this objective with a new Proposal for Regulation, published in March 2016[7], for "the conservation of fishery resources and the protection of marine ecosystems through technical measures". Regulations are established at the regional level for the North Sea, North Western Waters, South Western Waters, the Baltic Sea, the Mediterranean, the Black Sea and Outermost Regions. The proposal defines technical measures grouped into five categories:

- Measures that regulate the operation of the gear;
- Measures that regulate the design characteristics of the gears that are deployed;
- Minimum sizes below which fish must be returned to the sea;
- Measures that set spatial and temporal controls (e.g. closed/limited entry areas and seasonal closures) to protect aggregations of juvenile or spawning fish;
- Measures that mitigate the impacts of fishing gears on sensitive species (e.g. marine mammals, seabirds and turtles) or closed areas to protect sensitive habitats (e.g. coldwater coral reefs).

Unfortunately, and despite European Union's efforts, many offenses are regularly reported. Illegal, Unreported or Unregulated (IUU) fishing is a global threat to the marine environment and honest fishermen alike. The cost of these IUU activities is not harmless as the estimated cost of global IUU fishing is (worldwide) about 10 Billion Euros per Year. The European Union is committed to fighting IUU fishing worldwide. In collaboration with International organisations, EU publishes a list of blacklisted vessels that have been involved in IUU activities and of non-cooperative countries, and penalises offenders.

A first challenge for datAcron will be to support the European Union's control and enforcement strategy. The support for processing, analysis and visualisation of fishing vessels at the European scale - although not worldwide -, the capability of predicting the movement of maritime objects and the identification of patterns of movement and navigational events shall improve existing solutions to monitor the compliance to the European common fisheries policy.

Besides the control of fishing activities, another core issue of the EUMSS is safety. Fishing, in peace situation, is known as one of most dangerous activity and is regularly ranked in the

[6] http://ec.europa.eu/maritimeaffairs/index_en.htm

[7] COM(2016)134 final - 2016/0074(COD), "Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the conservation of fishery resources and the protection of marine ecosystems through technical measures, [...] https://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-134-EN-F1-1.PDF

top five, depending on the years. According to the American Bureau of Labor Statistics (BLS)[8], fishing is often identified as one of the top deadliest jobs. Ensuring security in this context is really challenging. First, because it does not concern only fishing vessels themselves but also the surrounding traffic and more generally all other human activities at sea. Another issue identified is that fishing vessels are repeatedly switching off their AIS device while fishing. Therefore, ensuring fishing security implies to process and predict fishing trajectories in real-time, detect fishing events, movement patterns and fishing areas when AIS is off, computing collision prediction with all surrounding ships (whatever their type) in a time scale typically of 5 to 15 minutes.

Another meaningful challenge for datAcron will be to explore and compare the use of AIS at large scale for fishing vessels monitoring [12]. Indeed in the domain of fisheries, the AIS data are far less used than in other domains, such as cargo monitoring for instance, because another specific system exists for tracking and monitoring commercial fishing vessels. The Vessel Monitoring System (VMS) is a European system set up during years 2000 for the control of fishing activities via satellite communications [19, 17]. Several studies such as [5] have considered this system for processing, analysis and visualisation of fisheries. The system has benefits for the control of fisheries, however VMS data are confidential, i.e, they are not freely broadcasted like the AIS, and positions are communicated at a very lower rate (*e.g.* *e*every hour). While fishing vessels are sometime switching off the AIS while fishing, a potential and significant improvement of existing systems based on analysis of AIS is foreseen for the control of commercial fishing in European waters.

3.3 Maritime operational use of datAcron

From a high-level perspective, the user in accomplishing his routine (still contextual) task of monitoring fishing activities, follows an iterative process of the information gathering and decision making steps which is classically represented by Boyd's OODA loop (Observe Orient Decide Act). The Observe and Orient steps could be further gathered under Situation Awareness while the Decide and Act steps are typically the Decision Making:

- **Observe**: Generation of automatic alerts based on Maritime Situational Indicators (MSI) requirements;
- **Orient**: Investigation of MSIs detected and confirmation of relevant events;
- **Decide**: Notification of relevant authorities;
- **Act**: Appropriate response and/or back to the Observe phase.

Figure 4 illustrates how datAcron could support the different steps of situation awareness and decision making.

During the *Observe* phase, automatic alerts are received corresponding to certain Maritime Situational Indicators previously selected and parametrised according the user's needs. These depend on the specific mission, and may vary from one day or another depending on the specific context (*e.g.* *e*, a high activity of illegal fishing reported in similar areas). Depending on his specific mission (see the different scenarios in Section 3.4), the user may be interested in the automatic detection of vessels heading to a specific port, entering a given area, traveling in a range of given speed values, those which the AIS contact has been lost, etc. The user may be interested in "higher-level" indicators about the vessels' behaviour such the vessels which are loitering in proximity to a given area, the ones which in proximity to the others for possible rendez-vous at sea, etc. As an automatic Maritime Situational Indicators (MSIs) detector of general interest, datAcron would generate relevant alerts upon some user's customisation.

[8] http://www.nbcnews.com/id/49125037

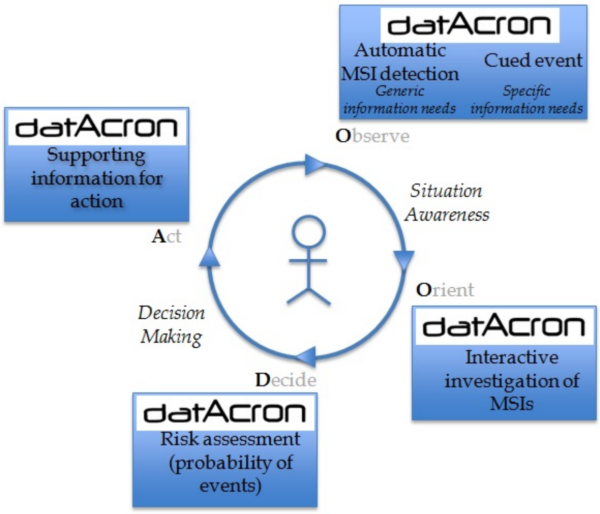


Figure 4: datAcron support to the OODA loop for maritime surveillance.

During the *Orient* phase, these alerts (*i.e.* detected MSIs) are further investigated to filter out the possible irrelevant ones. The user's domain expertise is used to filter out both false alarms due to the system's mistakes and some alerts that may be justified or explained. datAcron would support the interactive investigation of specific MSIs through visualisation and appropriate user interface.

In the *Decide* phase, the confirmed detection of event of interest triggers some notifications. Upon detection of an event of interest, a central component would receive a notification with event details (location, identification information, nature of activity). Depending on the system architecture, alerts could be sent directly to national operational headquarters, thereby removing the requirement for a central operational component. Each national component or member will correlate alerts with their own sensors and decide on a proper course of action. The decision is usually taken based on some risk assessment (probability, cost or consequences). datAcron would support the user in providing the probability of events of interest. In order to properly decide, additional information may be required bringing him back to the observation phase with specific information needs.

Finally, in the *Act* phase the effective action is taken and an intervention is launched. To be able to properly intervene, the authority in charge must have the appropriate information (the precise position of the vessel, the position of the vessels in the vicinity, the type of activity suspected, the predicted time of possible collision, etc). If some information is missing, the loop closes then to the observation phase for a further refinement. datAcron would provide the required information to the authority so they are able to act ("actionable knowledge").

datAcron would thus help the user in reaching the adequate level of situation awareness by answering specific information needs. Understanding possible human interactions with automated tools stresses the need for a suitable parametrisation of the algorithms so that they would actually meet a wide range of user's needs and a greater flexibility.

The conceptual diagram in Figure 5 illustrates the logical flow of datAcron and the main interaction of a datAcron software in the practical use case of maritime surveillance. The Maritime Situational Indicators list (see Table 2) captures the operational information needs while formalising in a generic way the goals of datAcron algorithms. The MSIs drive the datAcron analysis supported by the four work packages and their suitable interaction. The analysis is supported by some visualisation displaying analysis results and allowing possible refinement to the analysis through an updated parametrisation of the MSIs (e.g. change the areas of interest, speed thresholds, distance thresholds, etc).

The user's information needs (summarised in Table 2) are driven by his task to monitor the situation and prevent or detect one of the events described by different scenarios presented

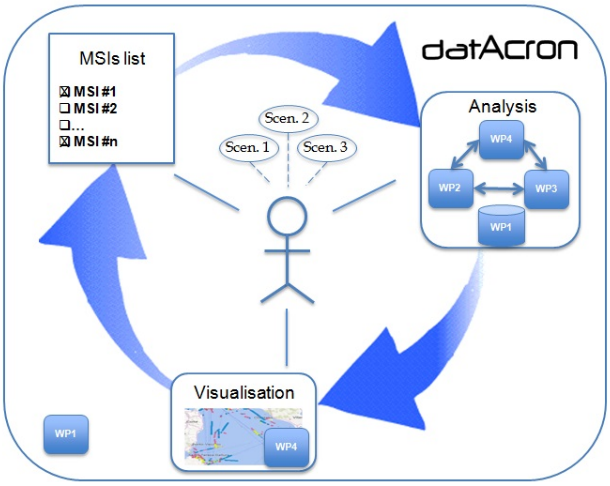


Figure 5: Conceptual diagram of the datAcron logical flow for the Maritime Use Case.

in Section 3.4. Depending on his objectives, the user expresses his needs through the list of MSIs of interest at a given time. He selects the appropriate algorithms and parametrizes them accordingly to perform the analysis. He is able to observe results of the selected algorithms using the visualisation tools and additional visual analytics, allowing him to further query the system upon specific information needs.

3.4 Scenarios

In order to support datAcron 's challenges, several scenarios involving fishing activities have been considered. These scenarios highlight the needs for continuous (real-time) tracking of fishing vessels and surrounding traffic, as well as contextually enhanced offline data analytics, including for instance, cluster and spatial analysis together with motion pattern detection. These scenarios are also designed to favour the use of additional contextual data. Finally they have been elaborated in order to stress datAcron 's algorithms in terms of *velocity*, *veracity*, *variety* and *volume*. They should provide a complete support for trajectory and event detection, prediction and visualisation.

From an operational point of view, the scenarios' aim is fishing security and control. For each scenario, we will describe the specific user information needs, linking them to Maritime Situational Indicators (MSIs). Specifically, the scenarios are:

- **SC11** - Collision avoidance;
- **SC12** - Vessel in distress / Man Overboard;
- **SC21** - Monitoring marine protected areas (from - illegal - fishing);
- **SC22** - Fishing pressure on areas;
- **SC31** - Detection of migrants / refugees and human trafficking;
- **SC32** - Illicit activities.

While the general processing scheme is described in Section 3.3, we detail below the users' goals and possible actions specific to each scenario. The detailed information needs are given in Table 2.

3.4.1 Secured fishing

As mentioned above, fishing is known to be one of most dangerous activity. datAcron can concur to secured fishing through scenarios SC11 and SC12 by detecting and preventing collisions between ships and by optimizing rendez-vous between rescuing ships in proximity of the vessel in danger and emergency services.

SC11 - Collision avoidance

Collisions involving fishing vessels are frequent, not only while fishing. Figure 6 shows a collision between a cargo and a fishing vessel, a typical example of a situation the datAcron 's user wants to prevent. In the situation illustrated in the picture, the collision occurred during the night between the fishing vessel Sokalique and the cargo Ocean Jasper. The cargo vessel continued his route without assisting the sinking fishing vessel, which asked the ships in her vicinity for help.

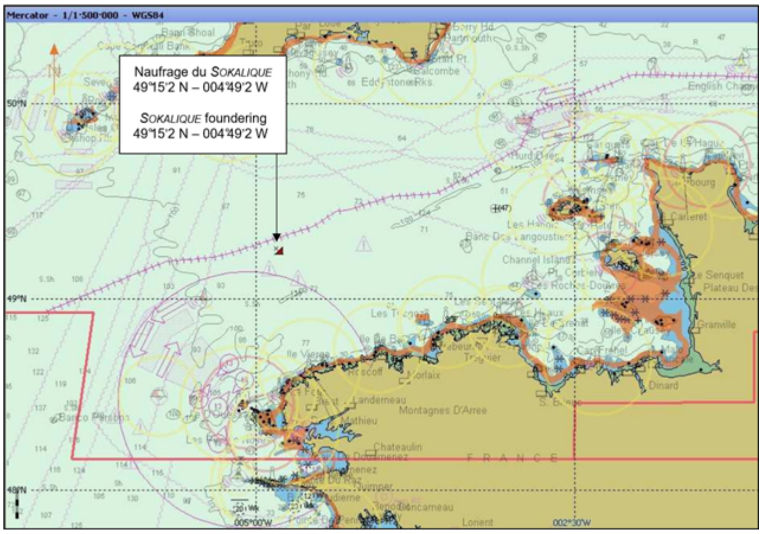


Figure 6: Example of a collision between a cargo and a fishing vessel

While vessels maintain the responsibility to comply with the COLREGs[9], this system can also be used to enhance the situational awareness between vessels, specifically when it is anticipated that a vessel will be required to "give way" to a fishing vessel (sailing vessels and power-driven vessels).

Objective - In order to prevent a collision of fishing vessels with other ships, the user wants to predict which other vessels (such as cargos, tankers, ferries) will cross the areas where the fishing vessels are fishing.

Actions - Upon the prevention of possible collisions, the user can send a warning to the fishing vessel and the vessels identified for possible collision so they can take the appropriate action. He will base his decision on the potential risk highlighted by the monitoring system.

SC12 - Vessel in distress / Man Overboard (MOB)

Beside collisions, a fishing vessel difficulty may be caused by technical problems or accidents occurring on-board. A situation of a fishing vessel having an on-board accident is shown in Figure 7. In this scenario, the fishing vessel left the port, went towards a first fishing zone then

[9] www.imo.org/en/About/Conventions/ListOfConventions/Pages/COLREG.aspx

disappeared (*i.e.* , it did not send any AIS data) for 35 minutes. After 35 minutes, the vessel re-appeared (AIS unit turned on), moved towards another fishing area and again switched off the AIS transmitter for 35 minutes. Then, it re-appeared again, moved towards a third fishing area where it started fishing (AIS off). The accident, highlighted as a red ellipse in the Figure 7, occurred in the morning, 7 hours after the AIS unit was turned off. The AIS on-board was re-enabled immediately after the accident and the ship headed back to the port. The rescue helicopter reached the boat one hour later (red circle in the picture).

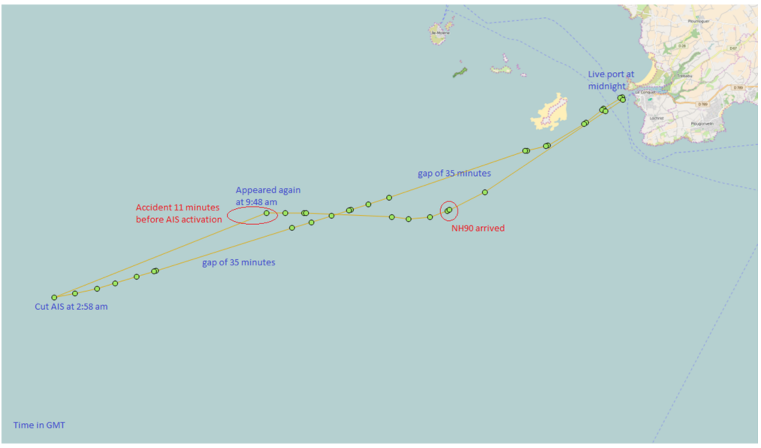


Figure 7: Example of accident on a fishing ship

A system that could alert the vessels in the vicinity of a vessel in distress or in a MOB situation would be a valuable capability to optimise rescuing operations. Moreover, the information from Emergency Position Indicating Radio Beacons (EPIRB)s and MOB equipped with AIS devices (fishermen sometimes wear small individual AIS beacon broadcasting in a range of 5 to 10 miles), if integrated in the system, could provide almost immediate event cueing.

Objective - In order to effectively intervene and help a vessel in distress, the user wants to identify the vessels which are in the vicinity of the vessel in danger, which could be on-site and provide assistance before the emergency services. In the case of a collision and escape of the responsible vessel, the user also wants to predict the trajectory of the fugitive.

Actions - Once the position of the vessel is known or predicted, the user alerts the vessels in proximity to the vessel in distress/MOB, and launches the Search And Rescue (SAR) operation, involving helicopters and, if necessary, rescue vessels. In case of the collision and escape, the Navy or the Coast Guards start an interception of the fugitive.

3.4.2 Maritime sustainable development

Estimating the spatial distribution and the intensity of fishing activities is necessary to natural resources management, impact assessment and to maritime planning. However, the access to such information at high-resolution is still a challenge. Since the European Union adopted the Vessel Monitoring Systems (VMS) to monitor fishing vessels, significant advances have been made[9]. Nevertheless, the results achievable by VMS analysis are limited by the restricted access applied to VMS data, which are confidential, and by their incompleteness, because for small vessels (i.e., less than 12 m long) the use of the VMS is not mandatory. In addition, VMS data analysis is usually conducted at a spatial resolution from 1 km to 10 km, which can be too coarse for secure fishing operations.

The perspective of characterizing the potential impacts of fishing activities, including the illegal ones, on species and on the geographical areas mentioned in the annex of European report COM/2016/0134 final - 2016/074 COD, and providing relevant information for European resources management based on AIS data is therefore challenging. Recently, the AIS data has been considered for the monitoring of fishing activities [12]. Two typical scenarios can be distinguished: Protection of marine areas from fishing (SC21) and fishing pressure on areas (SC22).

SC21 - Protection of areas from fishing

As stated in the recent report from *Global Fishing Watch* (www.oceana.org), the IUU fishing around the world *has escalated rapidly as the chance for profit outweighs concerns about the health and sustainability of our oceans.* Amongst others, the annex of European report COM/2016/0134 final - 2016/074 COD defines regulated areas for fishing [10]. These areas have potentially associated temporal constraints as fishing also depends on specific periods where fishing is legal and periods where fishing is prohibited. In order to protect these known areas from fishing, a monitoring of ships (not only the ones declared as fishing vessels) is required.

In order to control fishing in areas at the European level, the user needs to monitor ships in real-time. Entrances, exits and movements inside the surveyed areas have to be detected. Taking into consideration the identity declared by the ships as well as the type, the algorithms to develop should evaluate the right or not for this vessel to be in these areas of interest using the information from the fleet register, fishing licenses, black list, and historical data. Fishing patterns are detected and visualised and when the AIS has been switched off, predicted fishing polygons are estimated and compared with known fishing grounds to monitor.

Besides the control of fishing itself, there are regions that have to be protected from all activities as they correspond to geographic areas where protected species live (*e.g.* emarine reserves). The user would also be alerted about ships navigating and stopping in such areas. For example, f Figure 8 shows the trajectory of the Komarovo, a trawler registered in Russia. It appears to be fishing five times inside the Dzhugdzhursky State Nature Reserve in September 2013. The Komarovo has fished there at least for 13 days exhibiting different behaviours (slow motions, fast traveling speed, erratic use of the AIS).

Objective - In order to protect known areas from fishing (or navigation), the user wants to locate the set of geographical zones at the European level to be monitored. He wants to know if a vessel enters, exits, sails or spends time in such areas. When a vessel is located in the protected area, the user would like to know if there is fishing activity ongoing.

Actions - Detected offenders are tracked. Their trajectory and destination are verified. Navy, Coast Guards or port authorities control them and verify their freight. Erratic use of the AIS to mask such IUU should result in blacklisting of a fishing vessel.

SC22 - Fishing pressure on areas (density of fishing) Regular and legal fishing activities can also have negative impacts on the maritime environment and sea resources. There is an increasing concern for the preservation of natural resources against overfishing. Developed by Halpern in 2007, the concept of cumulative impacts mobilizes methods of spatial and quantitative analysis. It considers the potential impacts as the result of anthropogenic pressures on the ecosystem components. This has been applied to evaluate spatial and temporal changes in cumulative human impacts on the world's ocean over a period of five years [4]. The identification of these fishing efforts, based on AIS data, has been also recently considered by [15].

Figure 9 shows the estimated fishing activity for scallop dredging vessels per month in Brest bay. Fishing areas and fishing efforts on them had been identified by the analysis AIS data and the discovery of typical dredging patterns (speed-based).

A system that could count and display cumulative fishing activities over time and seasons would be valuable for the identification of overfishing areas and consequently for the preservation of marine resources.

[10] There are about 6600 marine protected areas covering about 2 percent of the world's oceans

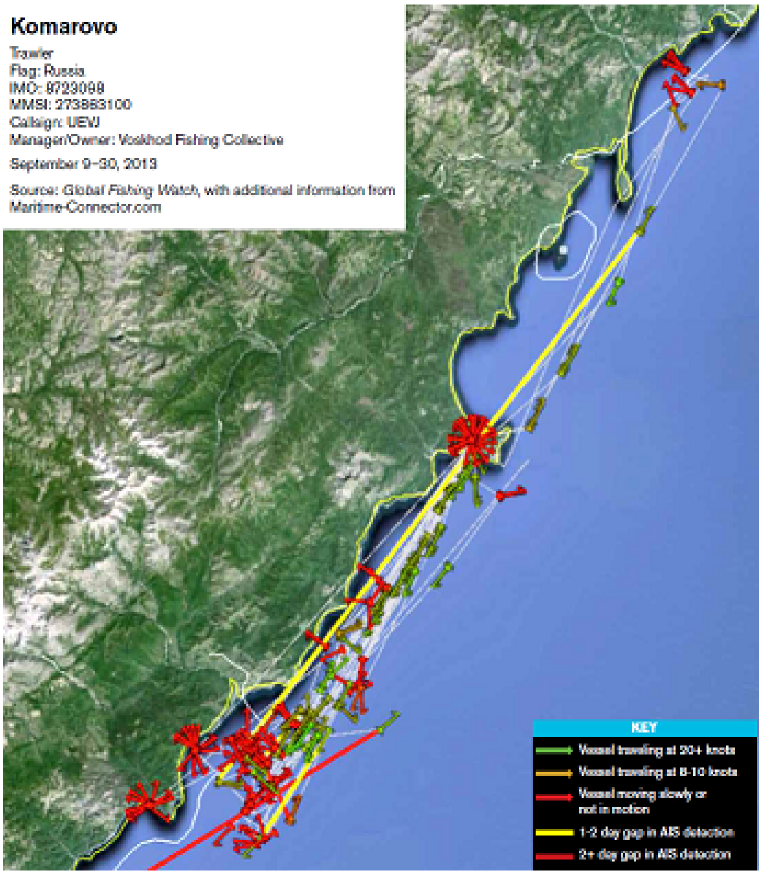


Figure 8: Track of the Komarovo in September 2013 (source *Global Fishing Watch*)

Objective - In order to identify overfishing areas, the user wants to identify fishing zones from positioning data. He would like to visualize and evaluate changes over time/seasons and see the cumulative impact.

Actions - Upon detection of intensive fishing areas and regarding concerned species, local, national and European authorities provide new regulations and update the list of protected areas.

3.4.3 Maritime security

Fishing vessels, either real or fake ones, are often used to conceal illegal activities that affect maritime security. In particular, two scenarios, regularly reported by authorities, can be explored by datAcron . The first concerns illegal immigration and human trafficking (SC31), the second one concerns more generically illicit activities occurring at sea (SC32).

SC31 - Migrants/refugees and human trafficking Human trafficking and migrant smuggling dramatically increased in the last years, in particular in the Mediterranean because of the refugees crisis, challenging the human dignity and the security and sovereignty of nations that have to tackle the situation. While the majority of the vessels engaged in these activities are small vessels that do not transmit AIS, fishing vessels have been known to engage in these activities as well. Accordingly, a fishing vessel detected along a common smuggling route and not engaged in fishing may be identified by the Maritime Security authorities as a "suspicious vessel", potentially engaged in trafficking activities.

For example, on May 1-4, 2015, a vessel was reported loitering in a common human trafficking/migrant route. The fact the vessel was lacking the fishing gear, suggested it might have been

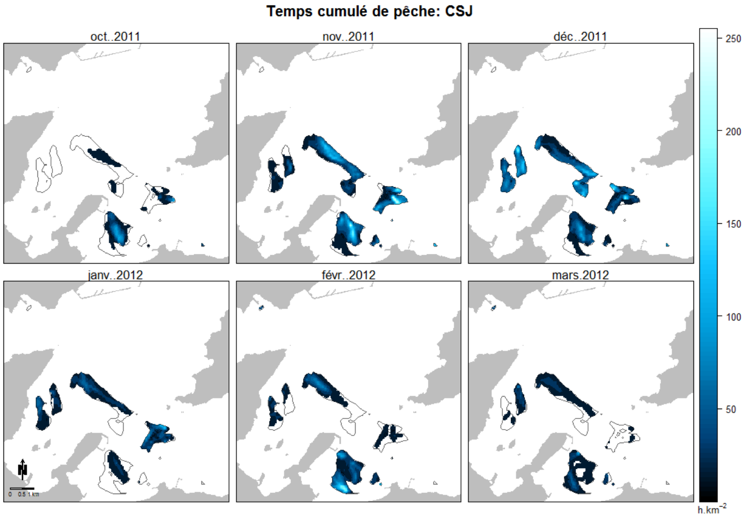


Figure 9: Estimated fishing activity for scallop dredging vessels per month [2]

involved in illicit activities and further investigation was necessary. In some cases, the presence of a vessel in the area can be licit, but it can still raise an alarm. For instance, the Phoenix, the vessel in Figure 10, is a former trawler (fishing vessel) reconditioned by an ONG to help migrants. It presence in an area under surveillance may be an indicator of migrant vessels.

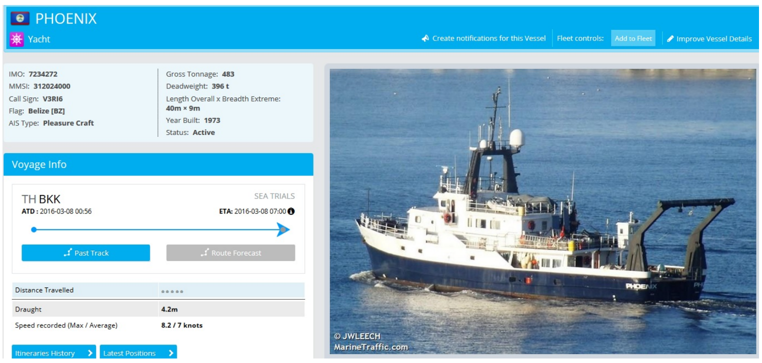


Figure 10: The vessel Phoenix (screen shot from MarineTraffic.com).

A system supporting the monitoring of migrants routes, able to alert when fishing vessels travel or loiter within such routes, and should be able to proved additional information on such vessels (do they have a fishing gear or not?), crosschecking information with lists of black listed vessels and of vessels that have been reported being involved in human trafficking, will be a valuable resource for maritime security operations. The system should also be able to predict the position of the suspicious vessels, to optimize helicopter and coast guards intervention.

Objective - In order to detect and rescue illegal migrants, refugees and identify human trafficking activities, the user wants to detect fishing vessels that are heading within migrants routes or loitering in these areas. Additional information on such vessels (fishing gear or previous negative records) should be at hand to help identify and predict movements of these vessels.

Actions - Once a vessel is detected traveling within a human trafficking route, the user will further investigate on the vessel records and decide to launch a mission for intercepting the suspicious vessel with helicopters or by sea, using the prediction route capability of the system.

SC32 - Illicit activities A number of criminal activities conducted at sea such as piracy, environmental pollution and drug smuggling are often linked to transnational crime organizations and can be difficult to detect and thwart, as well as prosecute. Predicted and real-time notification of *rendez-vous* of group of vessels in areas that are known to be used for trafficking, or any other *anomalous* behaviour as defined by the user, and handy historical records concerning suspicious or already blacklisted vessels, could enhance the ability to detect and intervene in real-time and to prosecute illicit activities.

In Figure 11, the path the Dona Liberta, a rusty refrigerated cargo vessel, followed from 2011 to 2014 along the coasts of Africa and Europe is shown. During that period, the Dona Liberta was reported to abandon crew members, abuse stowaways, dump oil and commit other crimes along the way[11]. However, port calls were often the main mean of locating the ship, which frequently turned off the satellite tracking signal.

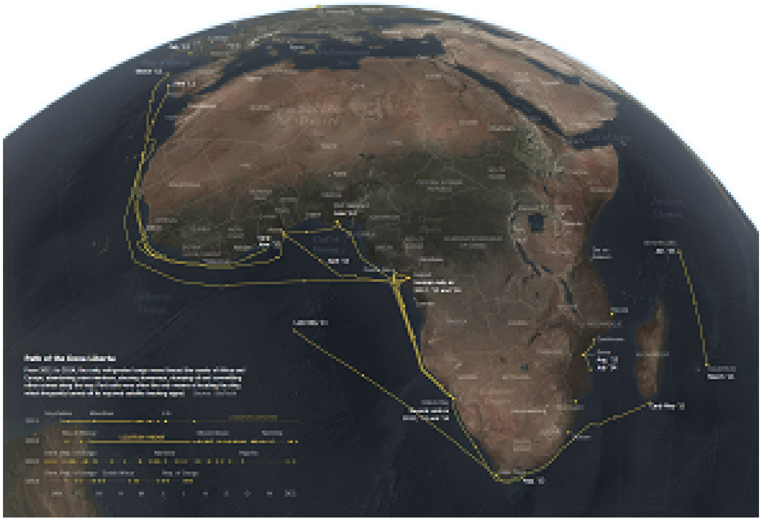


Figure 11: Path of the Dona Liberta (Source SkyTruth)

A system supporting the detection of illicit activities should provide support for monitoring of trafficking routes and areas. It should provide for instance alerts when groups of vessels, including fishing ones, are predicted to head in one of these areas, and to reach the area in the same period of time. The system should be able to let the user define a suspicious behaviour to be monitored and to be alerted for once the behaviour is detected by the system. The system should support the user to identify what vessels were reported to be involved in past illicit activities. The system should also be able to predict the future positions of the suspicious vessels, to optimize helicopter and coast guards intervention.

Objective - In order to detect illicit activities at sea, the user wants to detect group of vessels following a predefined or customized suspicious behaviour, such as heading or loitering in group in an area that is known to be used for trafficking (*rendez-vous*), identify the vessels involved and crosscheck them against existing records of illicit activities. The user wants also to predict the next positions of the vessels involved.

Actions - Once a *rendez-vous* or another suspicious behaviour is detected, the user may decide to launch a mission for intercepting the suspicious vessels with helicopters or by sea, using the prediction route capability of the system.

[11] http://www.nytimes.com/2015/07/19/world/stowaway-crime-scofflaw-ship.html?_r=0

3.4.4 Summary

Figure 12 summarises the scenarios as related to the datAcron work packages and big data challenges. Some examples of relevant information needs are provided, as well as some relevant data sources which could be used. Datasets and sources are further described in Section 4.

Table 1 summarises the scenarios, their objective, and user's possible actions.

Table 1: Scenarios' objectives, user's role and associated performance criteria

Scenario			Objective	Actions	Scenario-events	MSIs	Performance criteria
Secure fishing	SC11	Collision prevention	Protect fishing vessels from collision with big vessels (cargos, tankers, ferries)	Warn fishing vessels at risk, warn vessels heading to fishing areas	Real collisions	See Table 2	MSIs detected and predicted along the 5 dimensions: Completeness
	SC12	Vessel in distress / MOB (SAR)	Provide early assistance to a vessel in distress	Warn the closest vessels for early assistance, provide precise location of the vessel for the SAR team	Real vessels in distress		User detects specific scenario-events (through MSI queries)
Sustainable development	SC21	Protection of ecological areas	Protect specific areas from illegal fishing activities	Send control patrol boat to suspicious vessels location	Events of illegal fishing		
	SC22	Fishing pressure	Estimate and predict fishing pressure, identify areas at risk	Control future fishing activities	-		
Maritime security	SC31	Migrants and human trafficking	Detect possible human trafficking involving fishing vessels (or the like)	Communicate to security boarder control authorities, provide possible assistance (see SC12)	Real suspicious events		
	SC32	Illicit activities	Detect suspicious activities involving fishing vessels	Send control boats for further checking	Real suspicious events		

3.5 Maritime Situational Indicators for fishing vessels monitoring

The maritime use case and associated scenarios describe possible uses of a "datAcron software" with its different functions corresponding to the ones to be developed in the project. Given these high-level goals, Table 2 lists the expected outputs from a datAcron system fulfilling the user information needs in support to his specific task.

This list of Maritime Situational Indicators (MSIs) is extracted and abstracted from the outcomes of different workshops held in Sweden [1], Canada [20], NATO [22]. The global list of MSIs has been filtered out considering only (1) the ones that datAcron can provide, (2) the ones that are relevant to the scenarios considered.

In this table, the MSIs have been numbered for convenience and each of them should be split into an estimation at the current instant, denoted as MSI$\#n(0)$ and its predicted counter-part, denoted as MSI$\#n(t)$. By convention, we use 0 to denote the current instant, while t can be any instant in a near or farther future. As a consequence, t will be a parameter to be selected by the user upon his query.

The user would proceed by queries, parametrised according to his current and contextual needs (driven by the scenario). Parameters can include the speed, a given area, a given period of time, etc. Note that this parametrisation may require a prior processing such as route extraction,

		User information needs				Big data challenges			
	Use Case: Monitoring fishing activities	**[WP.1] Scalable integration and management of data from disparate and**	**[WP.2] Real-time detection and forecasting of trajectories**	**[WP.3] Real-time event recognition and forecasting**	**[WP.4] Real-time interactive visual analytics**	**Variety**	**Veracity**	**Volume**	**Velocity**
1	**Secured fishing**								
1.1	***Collision prevention*** : Protect fishing vessels from possible collision with other large vessels	*See the Big data challenges columns - Veracity: data cleaning*	Fishing vessels current position	Detection of fishing areas	Adaptive selection of the area of interest	AIS, possibly from different sources (terrestrial, satellite), different types of AIS (type B, c), with partially overlapping coverage	No AIS emission when fishing, AIS turning off or duping	AIS current streams, historcial streams	Varying temporal resolution of AIS contacts: e.g., fishing vessel every 30', cargo every 5'
	Possible actions (decisions) : Warn the fishing vessel at risk, Warn the vessel heading to the fishing vessel, etc.		Other vessels' position	Prediction of risk of collision	Selection of the time window		Intermittency of the AIS signal	Depends on the number of vessels in the area	
	Some rules: The merchant (power driven) vessels have to give way to fishing vessels while they are fishing			Recognition of fishing patterns			Lack of AIS coverage	Depends on the size of the area considered	
			Estimation of routes for large vessels and fishing vessels	Detection of maneuvers	Filtered display of all vessels with the risk of collision with the fishing vessels	METOC data: Low visibility determines safe speed	Imperfect METOC model, uncertain predictions	Depends on the time window considered	
				Constant bearing decreasing range detection		Legal navigation rules (for fishing and large vessels) IMO COLREGS			
			Prediction of large vessels position	Prediction the rendez-vous/coliision time with the vessel	Adaptive display of corresponding MSIs (mulitple spatial and temporal resolutions)	Known fishing areas	Variability of fish location according for instance to season		
1.2	***Vessel in distress:*** Rescue any vessel in distress as soon as possible	*See the Big data challenges columns - Veracity: data cleaning*	Current location of vessel in distress	Recognition of distress pattern	Filtered display of all vessels possibly meeting the fishing vessel in distress area	AIS streams, history of the vessel		Possible sparcity of the data (lack of AIS signal), to be completed by other sources (e.g., radar, SAR)	*Real-time processing contraints*
			Prediction of the vessel in case of drift	Reconstruct vessel in distress' trajectory for position prediction		METOC data (wind and currents)	Vessels can change direction		
	Possible actions (decisions) : Warn the closest vessels to the vessel in distress for early assistance, launch the SAR operation		Prediction of any vessel's position in the "vicinity" of the vessel in distress	Prediction the rendez-vous time with the vessel		Radar or SAR in case the vessel in distress' AIS signal is off	Measurement noise on radar images		
							SAR imagery (satellite) needs to be extrapolate to current time		
						VHF communication, distress signal (COLREGS)	Poor communication, weak unreadable message		
						EPIRB emergency beacon			
2	**Sustainable development**								
2.1	***Protection of marine protected areas(MPAs)*** : Protect specific areas from fishing activities (and from maritime traffic???)		Estimation of fishing vessels location	Recognition of fishing patterns	Visualisation of MPAs	AIS	No AIS emission when fishing, AIS turning off or duping		
			Prediction of future fishing vessels location	Detection of fishing areas	Visualisation of common fishing grounds from historical data,	Map of the protected areas	Variability of fish location according for instance to season		
2.2	***Fishing pressure on areas*** : Predict and prevent the pressure on fishing areas		Prediction of fishing vessels location	Detection of fishing areas	Visualisation of fishing areas: fishing density vs. seasons vs. types of fish vs. time	AIS	No AIS emission when fishing, AIS turning off or duping		
					Visualisation of preferred MPA management options (adding or removing protected areas)	Map of the known fishing areas	Variability of fish location according for instance to season		
						SAR imagery			
3	**Illicit activities**								
3.1	***Detect and locate fishing vessels engaged in illicit activities (fish trafficking, good trafficking, fishing spoofing, …)***		Analysis of historical pattern (e.g., Dona Liberta)	AIS spoofing detection (position, changing name, etc)			Deception of AIS signal		
				Recognition of fishing patterns			Deception in behaviour		
3.2				Rendez-vous detection or prediction					
3.3				Any inconsistency between type and behaviour					

Figure 12: Use case and scenario description as related to the datAcron work packages and big data challenges.

Table 2: Maritime Situational Indicators needs for each scenario. 0 means that the information is required at the current instant in time, while t means that it is required as a prediction for within a time to be set by the user.

			Scenarios					
			Collision prevention	Vessel in distress / MOB	Protection of ecological areas	Fishing pressure	Migrants / human trafficking	illicit activities
Maritime Situational Indicators		**Query parameters**	**SC11**	**SC12**	**SC21**	**SC22**	**SC31**	**SC32**
Close to a critical infrastructure (CI)	MSI#1	CI locations, range					×	×
Within a given area	MSI#2	Areas locations	×		×	×	×	×
On a maritime route	MSI#3	Route locations	×	×			×	×
Proximity of other vessels	MSI#4	Range of distance	×	×	×	×	×	×
In stationary area (ports or offshore platforms)	MSI#5	Specific ports locations					×	×
Null speed	MSI#6		×	×	×	×	×	×
Change of speed	MSI#7	Change rate		0			×	×
Not compatible with range values from:								
- the current area	MSI#8				0	0	0	0
- the type of vessel	MSI#9				0	0	0	0
- the vessel's history	MSI#10	Time window			0	0	0	0
- user defined	MSI#11	Range of values	×	×	×	×	×	×
Change of course	MSI#12	Change rate	×				×	×
Not compatible with:								
- the vessel's expected destination	MSI#13		×	×	×		×	×
- user defined	MSI#14	Range of values	×	×	×	×	×	×
No AIS emission/reception	MSI#15		×		×		×	×
AIS emission interrupted	MSI#16		×	×	×	×	×	×
Change in AIS static information	MSI#17	AIS field			×		×	×
AIS error detection	MSI#18	Type of error			×		×	×
Under way (using engine or sailing)	MSI#19		×		×		×	×
At anchor or moored	MSI#20							
Movement ability affected	MSI#21			×				
Aground	MSI#22			×				
Engaged in fishing	MSI#23		×	×	×	×	×	×
Tugging (tugged or tugging)	MSI#24							
In SAR operation	MSI#25			×				
Loitering	MSI#26		×	×	×	×	×	×
Dead in water, drifting	MSI#27		×	×				
Rendez-vous	MSI#28				×	×	×	×

loitering or fishing areas detection, the estimation of vessel's destination. It can also require some external databases such as the World Port Index, the maps of known restricted or ecological areas. Consequently, the list of MSIs should be understood as the user's information needs regarding specific vessels' state (current or predicted). It does not describe the information needs in terms of synthesised information such as the maritime routes, fishing or loitering areas.

The information needs depend on the task and goal of the user. Although it may be argued that the more information the best, and that all MSIs would be useful to the user regardless his task, we refine below the user's needs for each goal in each specific scenario. For example, the user whose goal is to prevent collision between fishing vessels and big vessels would need to know which vessels are currently in a known fishing area (MSI#10), or will be in the next 3 hours (MSI#1t). He may need to know which fishing vessels are actually located on a maritime route (MSI#31), which big vessel may be in a fishing areas in the next 3 hours (MSI#2t), but also which fishing vessels are engaged in fishing (MSI#230) or will be in 3 hours (MSI#23t). Table 2 summarises the MSIs needs for each scenario.

Note

So far, there is no standard list of MSIs. A consolidated list of MSIs requires certainly a huge amount of work well beyond the datAcron objectives. It is thus expected that this list will evolve and be refined during the project, as the different use case partners and member interest groups are involved. However, the list provided here complies with some basic criteria of simplicity, generality and operational relevance which make it suitable as a starting point for datAcron activities.

Thus, during the project, it is expected that:

- a collaborative work among scientists and end-users will provide an updated and validated list of MSIs;
- the datAcron partners comment on their ability to design the supporting algorithms for these MSIs.

3.6 Operational performance criteria and datAcron assessment

From the end-user standpoint, an algorithm or system will be successful if:

1. it provides the end-user with the requested information (*i.e.* , the MSIs) within a adequate time, with a high clarity and accuracy, and that the information is maintained in time.
2. the user is actually able through his interaction with the system to perform his task and in our particular use case, to discover the scenario-events.

We see thus two levels of assessment, *the MSI level* and *the scenario-event level*.

The quality of the Recognized Maritime Picture (RMP)[12] which can be assessed according to the five criteria of *Completeness*, *Accuracy*, *Clarity*, *Continuity* and *Timeliness* as presented in Table3 Each criterion may be defined relatively to a given area, a given period of time, and a given MSI. Hence, for a given scenario, the user expects datAcron algorithms (from WP2, WP3 and WP4) to provide answers to his MSI queries with a quality defined by the five dimensions in Table 3.

The visual analytics (WP4 functions) would help him in creating the queries and refining them so that he is able to detect the scenario-related events. These *scenario-events* are the few events that the user is looking for and expecting to find. Another layer of performance criteria

[12]The Recognized Maritime Picture (RMP) is defined as a compiled plot to depict maritime activities, as provided by different sensors or platforms. It includes thus the detection of vessels, their classification, the classification of any behaviour of interest.

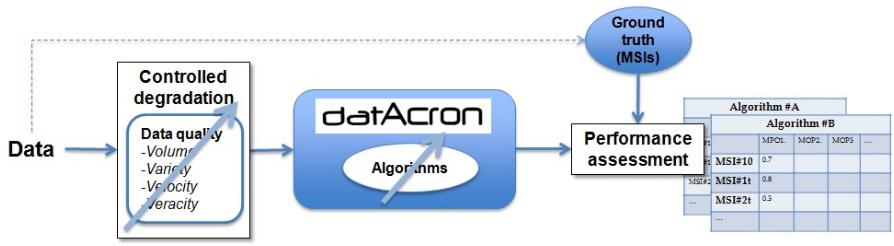
Table 3: Performance criteria for the maritime picture, at the MSI level

Criterion	Definition	Example of measure
Completeness	Degree to which the trajectories are processed to detect MSIs	Nb of trajectories processed over the available ones
Accuracy	Degree to which the detection of each MSI agrees with the corresponding true value	True positive rate & False positive rate
Clarity	Degree of confidence to which the detection of each MSI is provided	Probability value before decision
Continuity	Degree to which the detection of MSIs is maintained in time	
Timeliness	Time to provide the detection result for each MSI	

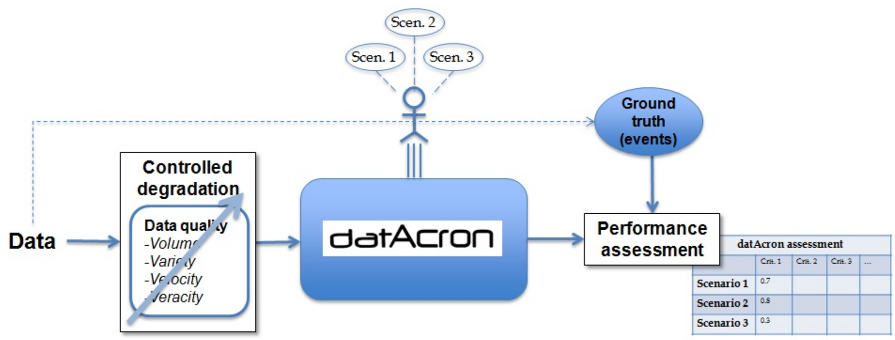
would then be defined as related to human factor tasks. Since these scenario-events (collision, vessel in distress, smuggling, etc) are rare events, the statistics would be less relevant.

Sketch of the experimental plan

Figures 13 sketch the general idea to be followed in developing the experimental plan (datAcron task 5.2). Mainly, datAcron algorithms will be evaluated along both the operational (Table 3) and technical (given in the GA) criteria, although some may overlap. The data would be degraded to study the impact of the different Big Data dimensions on the algorithms outputs. Also, several algorithms or techniques would be compared, including the new ones developed by the different teams as well as some state-of the art approaches from the literature.



(a) Performance assessment during the training phase (in laboratory) - MSI level assessment



(b) Performance assessment during the testing phase (operational) - Scenario level assessment

Figure 13: Sketch of datAcron evaluation

The dataset would be controlled to provide some ground truth information to be able to assess some robustness to the *veracity* of data. The *volume* and *velocity* of data may be made varying to observe the impact on the timeliness of the algorithms. The *variety* of the data would vary depending on the sources selected to feed the algorithms and its impact on the results would be assessed.

4 MARITIME DOMAIN DATA IN SUPPORT OF THE datAcron MARITIME USE CASE

Heterogeneous and independently maintained data sources, including quasi-real time data streams as well as archival European datasets, will support the datAcron maritime use case described in Section 2.3. Different sensor technologies are being developed and only if properly combined and integrated these data can provide useful information and knowledge in support to maritime situational awareness. Moreover, data coming from multiple sources need to be cleaned up from inconsistencies, converted into standard formats to be harmonized and summarised. To meaningfully implement datAcron 's algorithms, we need also to process real (or synthetic) data.

Thus, the maritime use case should be supported by datasets such as:

- *Automatic Identification System (AIS)* [13] messages broadcasted by ships for collision avoidance;
- *Maritime regulations*, specifying the legislation and the rules for navigation and fishing;
- *Marine protected/closed areas*, where fishing and sea traffic may be (temporarily) forbidden;
- *Traffic separation schemes* and *Nautical charts*, useful to define vessel routes;
- *Vessel routes* and *Fishing areas* estimated from historical traffic data;
- *Registry data* on vessels and ports;
- *Records* of past events, such as incidents and illegal activities reports;
- *Meteorological and oceanographic data (METOC)* on atmospheric and sea state conditions and currents;
- *Long-range identification and tracking (LRIT)* [14] data for global identification and tracking of ships;
- *Marine and costal Radar* data, used for supporting navigation and collision avoidance;
- *Synthetic aperture radar (SAR) or ISAR* images.

AIS data will be the main data source for generating the maritime moving objects trajectories that will be analysed by the algorithms developed by WP2, WP3 and WP4 and visualised by WP4. In particular, historical sea traffic data from AIS may be a powerful source of information to infer actual vessel routes and fishing areas, which are usually not available. Marine protected areas and maritime regulations will be used to define the context of the use case, for example to define the area to monitor, the time range of interest and the potential illegal activities to prevent. Historical and METOC data may be important to correctly set up the analysis parameters and to validate the results, for example to filter out false alarms. Registry data and records of past events may be used to define ground truth information useful to refine the analysis algorithms. Other types of information, in particular RADAR and Synthetic Aperture Radar (SAR) data, could be considered in a later stage to enlarge available ground truth data. Moreover, ground truth can be further complemented with and synthetic datasets.

A potential list of datasets that can be used in support of the datAcron 's Maritime Use Case for testing and validating the datAcron developments of WP1, WP2, WP3 and WP4 has been identified and summarized in Tables 4, 5 and 6 included at the end of this section. These datasets are described in the rest of the section, discussing how they may support the use case and their availability.

[13] AIS: www.navcen.uscg.gov/?pageName=AISmain
[14] LRIT: www.imo.org/en/OurWork/Safety/Navigation/Pages/LRIT.aspx

This list of datasets is preliminary and will be consolidated in Deliverable D.5.2 and D5.3, which will discuss data preparation and the experimental evaluation. Not all these datasets are necessary to the use case, and it's possible that other datasets will be identified in the next months.

4.1 Spatial coverage

A preliminary area of interest for the specification of the maritime use case is specified as depicted in Figure 14. This area can be further enlarged, to include events of interest for the validation of datAcron 's developments.

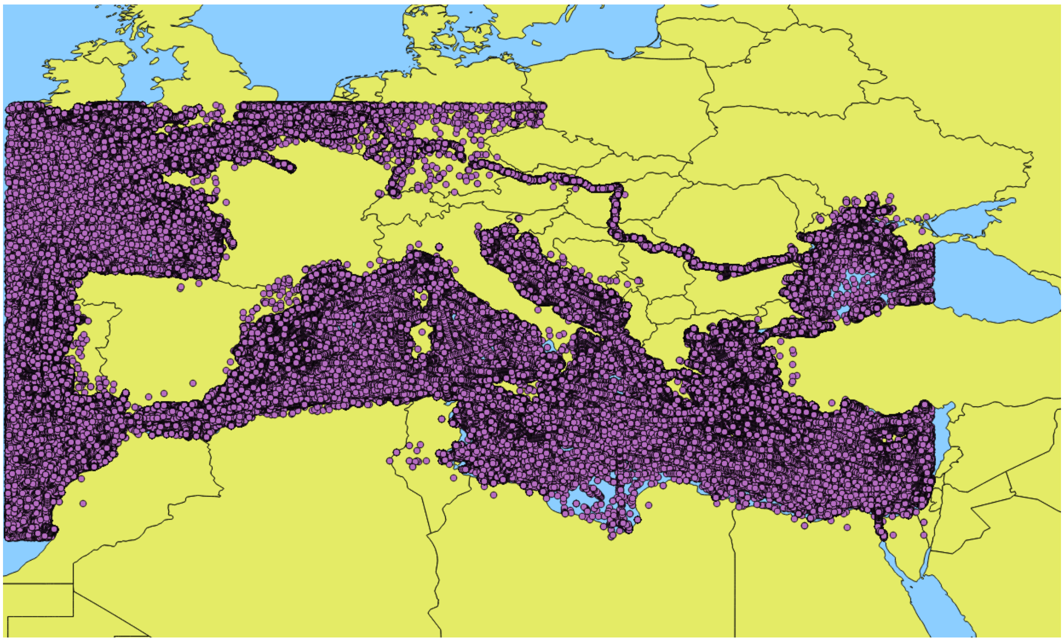


Figure 14: Geographical coverage of the maritime use case datasets (IMISG)

4.2 Automatic Identification System (AIS) datasets

According to the European Commission regulations in the Mediterranean Sea, several types of ships are obliged to broadcast AIS messages, including: ships of 300 gross tonnage and upwards in international voyages; 500 and upwards for cargoes not in international waters and passenger vessels; and, more recently, smaller fishing vessels. Raw AIS messages comply with ITU-R.M 1371-4 and NMEA 4.0 standards, and are differentiated in 27 type of messages. Two main classes of messages are distinguished:

kinematic messages from which 2D vessel routes can be derived, and which include information on position, speed, heading, course over ground (COG), Rate of Turn (ROT);

static messages providing ship meta-information such as ship identifiers (MMSI and IMO number), name, type, and dimension of vessel, and route-based information, such as destination (Port of Call), danger, Estimated Time of Arrival (ETA), draught.

AIS messages are collected by coastal and satellite networks of receivers. Terrestrial AIS (T-AIS) messages from coastal receivers are characterized by high persistence but limited coverage, while Satellite AIS (S-AIS) messages collected by satellite receivers can pick up messages in the open sea, far away from the coastline, and have an larger coverage than T-AIS. Moreover, AIS messages from different data providers might have complementary spatial and temporal coverages. It is therefore advisable, depending on the area or on the period of interest for the use case, to have an harmonized dataset including AIS messages from different providers and from different types of receivers. datAcron will exploit both T-AIS and S-AIS messages, and might consider AIS messages from a variety of sources. For example, the following open data providers, together with commercial data providers, could be exploited:

- AISHub[15], an open network of AIS receivers that make available to partners a merged TCP stream of raw AIS messages according to the NMEA format. Data can be freely distributed. Historical data may also be available from partners.
- MarineTraffic [16], an open network of AIS receivers for community sharing real-time AIS data, made available in a *quasi* real-time stream, together with vessels and ports information.

Historical T-AIS and S-AIS data will be provided by IMISG and DCNS Research in the form of formatted flat files that can be used to simulate quasi-real time TCP streams. Additional datasets on limited areas of interest can be made available by other partners. For instance, Figure 15 illustrates the coverage of NARI's terrestrial receiver, which is located in Brest bay, and that could be made accessible as historical flat files. Other data can be simulated or derived from real ones for particular applications, *e.g.* e, to include simulated alarms or events of interest for algorithms testing.

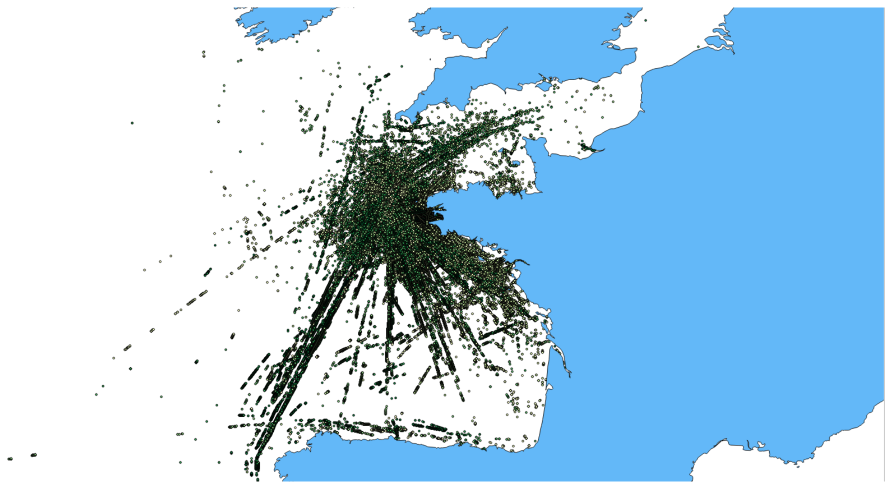


Figure 15: Coverage of NARI dataset

4.3 Contextual and archival data sources

Several data sources can be used to contextualize the Maritime Sustainable Development scenario, as well as to further characterize the other scenarios included in the fishing activities monitoring use case, including:

[15] AISHub website: www.aishub.net
[16] MarineTraffic website: www.marinetraffic.com

- EU regulated fishing areas included in the recent European proposal on the conservation of fishery resources and the protection of marine ecosystems through technical measures. These are given as thematic as well as spatial information, together with the coordinates of the regulated fishing areas. This dataset can be used as a contextual information for the IUU fishing scenario.
- FAO[17] and ICES[18] Fishery Statistical Areas, and Fish Catches by FAO [19].
- The Community Fishing fleet register, a European register of offial fishing vessels maintained by the European Commission [20]. This dataset can be used as a contextual information for the IUU fishing scenario.
- The European Marine Observation and Data Network (EMODnet) datasets[21], that include a series of standard and harmonized datasets including European coastal maps, human activities such as ports and fishing areas, biological datasets, a digital terrain model for bathymetry, etc. The EMODnet portal is an entry point for harmonized marine data generated by 100 organisations, free of restrictions on use. These datasets can help contextualizing the illegal fishing use case, as well as other use cases that will be developed.
- The marine protected areas in Europe defined by the *NATURA2000* ecological network of protected areas and freely downloadable from the European Environmental Agency (EEA) website [22].
- Environmental biodiversity datasets (*e.g.* *e*, marine biodiversity, waste, sediments) harmonized at European level and freely downloadable in different formats from the EEA website [23] that can be useful to further develop the maritime sustainable development scenario.
- A high resolution coast line map for analysis from EEA [24]
- The *World Port Index (WPI)*[25], an open, freely available and distributable port database maintained by the National Geospatial-Intelligence Agency that contains the locations, the physical characteristics and the facilities and services offered by major ports and terminals world-wide. It contains approximately 3700 entries. It can be useful to interpret events and activities done by vessels closeby ports (e.g., vessels loitering in a loitering area of a port that never enter the port). It is distributed as a PDF report, as an access database and as ESRI shape file.
- The *Open Sea Map*[26], an open and free nautical chart, including beacons, buoys and other navigation aids as well as port information, repair shops and chandlerys. It can integrate the World Port Index supporting the maritime use case with information on port facilities and provide also information on vessel routes close to ports. It is an Open Street Map project and data are freely available, usable and distributable according to the ODBL open data common licence [27].
- Official nautical charts (IHO-S-57, format for Electronic Navication Charts, ENC), cartographic information that could help understand sea traffic. Designed under the umbrella

[17] http://www.emodnet-humanactivities.eu/search-results.php?dataname=FAO+Fishery+Statistical+Areas
[18] http://www.emodnet-humanactivities.eu/search-results.php?dataname=ICES+Statistical+Area
[19] http://www.emodnet-humanactivities.eu/search-results.php?dataname=Fish+Catches+by+FAO+Fishery+Statistical+Areas
[20] http://ec.europa.eu/fisheries/fleet/index.cfm
[21] EMODnet portal: emodnet.eu
[22] NATURA2000: natura2000.eea.europa.eu
[23] EEA biodiversity data: biodiversity.europa.eu/data
[24] EEA European coastline: www.eea.europa.eu/data-and-maps/data/eea-coastline-for-analysis-1
[25] WPI:msi.nga.mil/NGAPortal/MSI.portal
[26] Open Sea Map: www.openseamap.org, Downloadable from: planet.openstreetmap.org
[27] opendatacommons.org/licenses/odbl/CC Share Alike 2.0

of the International Hydrographic Organisation, these vector maps (not free) define navigation features (*e.g.* *e*buoys, traffic separation schemes, regulated and restricted areas, accurate bathymetry for coastal areas. Data scheme IHO-S-57 is documented [28]. A GDAL driver[29] is available for converting data in other vector formats.

4.4 METOC data

Weather data and ocean data from forecast models and from observations (e.g., in-situ sensor data), which are openly available from several providers, can help validate analysis results and reduce false alarm rate, for example identifying sea and weather conditions that force vessels to change direction or modify their normal behaviour. They can also be used to characterize seasonal trends in traffic routes, and to contextualise movement parameters such as speed of vessels.

The reference source of harmonized oceanographic data in Europe is by far the Copernicus Marine Environment and Monitoring Service (CMEMS) [30], developed by the EU as part of the European Programme for the establishment of a European capacity for Earth Observation and Monitoring. This operative service provides an interactive catalogue of updated oceanographic products produced by the network of oceanographic centres in Europe. 140 data products world wide can be downloaded for free after registration, including data on the Mediterranean, the North West Coast of Europe and the Iberian Peninsula. Depending on the dataset, historical forecast data may be available, as well as reanalysis of past forecast.

In particular, weather and ocean datasets that can be of helpful to the maritime use case include:

- The *Mediterranean Sea Physics Analysis and Forecast (MFS)* model, an hydrodynamic-wave model for the Mediterranean basin that forecast physical ocean variables including the sea currents (sea water velocity) and the sea surface height above the sea level. MFS is one of the products available in the CMEMS catalogue[31]. It is updated daily, and data are available from 2013-01-01, and variable values are available as daily mean and hourly mean. Its geographical coverage is Latitude North 45.937, Latitude South 30.187, Longitude East 36.25, Longitude West -15, with a resolution of 1/16 degree (*i.e.* , ca. 6-7 km).

- *Global Ocean Wind* observations available through CMEMS service, with horizontal resolution of 0.25x0.25 degrees and 6 hours in time[32], as well as daily and monthly mean from 2007;

- *Ocean Wave model data* from the European Centre for Medium-Range Weather Forecasts (ECMWF) for example from the ERA-Interim model [33], which offers many other atmospheric variables, and is available daily until 2015; the ERA-20C model[34] available until 2010. Other open ECMWF datasets are available [35].

- Model data for the Adriatic from the *Adriatic Forecasting System* [36], which makes available one week of data.

[28] IHO-S-57 www.s-57.com
[29] http://www.gdal.org/drv_s57.html
[30] Copernicus Marine Environment and Monitoring Service: marine.copernicus.eu
[31] MFS: marine.copernicus.eu/web/69-interactive-catalogue.php?option=com$_$csw&view=details&product$_$id=MEDSEA$_$ANALYSIS$_$FORECAST$_$PHYS$_$006$_$001$_$a
[32] marine.copernicus.eu/web/69-interactive-catalogue.php?option=com_csw&view=details&product_id=WIND_GLO_WIND_L4_NRT_OBSERVATIONS_012_004
[33] apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/
[34] ERA-20C apps.ecmwf.int/datasets/data/era20c-daily/levtype=sfc/type=an/)
[35] apps.ecmwf.int/datasets/
[36] www.ionioproject.eu

- Model data for the Ionian Sea, from the *Ionian Forecasting System* [37], which makes available three months of data.
- In-situ Observation on the Adriatic sea, from the *IONIO in-situ database* [38], with historical data from 1986.
- The *Sea Conditions* [39] dataset, which includes Significant Wave Hight (SWH) (Data are not available for download).
- The *National Oceanic and Atmospheric Administration (NOAA) datasets* [40], including global meteorological and oceanographic datasets from cooperating networks of ships and buoys.
- Energy data from *Seabed Habitat from EMODnet* include wind, waves and currents harmonized at European Level [41].
- *Physics data from EMODnet*, including Sea water temperature, salinity or density, water currents, level, Waves and winds, Light attenuation, Atmospheric parameters at sea level, HF radar data [42]

4.5 Events reports

Several sources of information are available for events reports that can be used to define ground truth datasets. Most of these sources are unstructured, and may include fisheries facts from news, media and social medias. A non-exaustive list of datasets of interest for the illegal fishing scenario is as follows:

- The *Blacklists of IUU fishing vessels*. A list of fishing vessels involved in IUU is compiled by Greenpeace from publicly available official registries of IUU vessels and companies[43].
- The *Worldwide Threat to Shipping (WTS)* reports.
- *Interpol reports on illegal fishing events* [44].
- Several reports and datasets on piracy events are compiled by ASAM, IBM, Interpol.

4.6 Available datasets

Table 4, 5 and 6 summarize the main characteristics of the available datasets that can be potentially exploited for the validation of datAcron results through the datAcron 's Maritime Use Case. This list overview the datasets and their main characteristics, such as coverage, access interface, formats, to help WP1, WP2, WP3 and WP4 partners identify the most suitable datasets for testing and validation of the datAcron developments.

[37] ionioproject.hcmr.gr
[38] www.mediterraneanmarinedata.eu/ionio/home.htm
[39] sea-conditions.com
[40] www.ndbc.noaa.gov/data
[41] http://www.emodnet-seabedhabitats.eu/default.aspx?page=1934
[42] http://www.emodnet-physics.eu/Map/service/Catalogue.aspx
[43] http://blacklist.greenpeace.org/0/vessel/list
[44] http://www.interpol.int/INTERPOL-expertise/Notices/Purple-notices-%E2%80%93-public-versions/2014

Table 4: Inventory of available datasets

Type	Dataset	Description	Data source	Spatial coverage	Temporal coverage	Availability	Licence/Ter for use	Available formats
Positioning and static vessels' information	Consortium AIS data	Terrestrial (T-AIS) and Satellite (S-AIS) information on 2D vessel routes, including ship position (kinematic messages: position, speed, heading, COG, ROT), ship meta-information (static messages: ship ids, name, type, dimension), route-based information (destination, danger, ETA, draught)	Consortium partners	Mediterranean Sea	Historical data available	1 month sample	Consortium only	Text files and TCP streams according to ITU-R.M 1371-4, NMEA 4.0[45]
	Global AIS - AISHUB		www.aishub.net	Global	Historical data may be available	1 month sample	Data received from AISHUB can be shared freely	TCP/UDP quasi-RT streams according to NMEA AIS format
	Global AIS - Marine Traffic		www.marinetraffic.com	Global	Historical data may be available	1 month sample	Data received from AISHUB can be shared freely	TCP/UDP quasi-RT streams according to NMEA AIS format
	Brest AIS		NARI's AIS receiver	Brest area	Historical data		Consortium only	Flat files in NMEA AIS format
Contextual Information	EU regulated fishing areas	Fishing areas	European Commission	European		WebGIS European Atlas of the Sea http://ec.europa.eu/maritimeaffairs/atlas/maritime_atlas/#lang=EN;p=w;pos=33.156:50.493:4;bkgd=5:1;gra=0;mode=1;theme=52:0.8:1:0;time=2013;	Open Data	spatial data
	Community Fishing fleet register	register of declared fishing vessels	European Commission	European		http://ec.europa.eu/fisheries/fleet/index.cfm	Open Data	Web service, text file
	Fisheries areas	FAO Fishery Statistical Areas, Fish Catches by FAO Fishery Statistical Areas, and ICES Statistical Areas through Human activities datasets from EMODnet, harmonized at European level	FAO, ICES	European		http://www.emodnet-humanactivities.eu/search-results.php?dataname=FAO+Fishery+Statistical+Areas, http://www.emodnet-humanactivities.eu/search-results.php?dataname=Fish+Catches+by+FAO+Fishery+Statistical+Areas, http://www.emodnet-humanactivities.eu/search-results.php?dataname=ICES+Statistical+Areas	Open Data	WMS, WFS services
	EMODnet Main Ports statistics	Main European ports data from EMODnet Human activities dataset, with information on datasets on passengers, vessels and goods (EUROSTAT), traffic in main ports, harmonized at European level	EUROSTAT, ***Eurofish International Organisation***	European	Updated up to 2013	http://www.emodnet-humanactivities.eu/search-results.php?dataname=Main+Ports	Open Data	
	EEA datasets	Environmental datasets harmonized at European level, including biodiversity, waste, sediments, high resolution coastline		European		http://www.eea.europa.eu/data-and-maps/find/global?c12=marine&search=Search#c1=Data&c1=Graph&c1=Indicator&c1=Interactive+data&c1=Interactive+map&c1=Map&c6=&c0=12&b_start=0&c12=marine	Open Data	
	Natura2000	Database of European Marine Protected Areas (MPAs) and MPAs designated under the Regional Sea Conventions		European		http://www.eea.europa.eu/data-and-maps/data/natura-6 http://www.eea.europa.eu/data-and-maps/figures/mpas-designated-under-the-regional	Public, Free, distributable	Access database (MDB), CSV
	World Port Index	Port database. Location and physical characteristics of, and the facilities and services offered by major ports and terminals world-wide (approximately 3700 entries)	National Geospatial-Intelligence Agency	Global		http://msi.nga.mil/NGAPortal/MSI.portal?_nfpb=true&_pageLabel=msi_portal_page_62&pubCode=0015	Public, free for use, redistributable	PDF, Access database (MDB), shape file (SHP)
	OpenSeaMap	Nautical chart, including beacons, buoys and other navigation aids as well as port information, repair shops and chandlerys	www.openseamap.org/index.php?id=openseamap&L=1	Global		OSM http://planet.openstreetmap.org/	Open data. Free. ODBL open data common[46] Share Alike 2.0. Data freely available, usable and distributable	Data scheme IHO-S-57 for Electronic Navication Charts (ENC)[47]
	Electronic Navication Charts	Nautical charts and navigation features, buoys, traffic separation schemes, regulated and restricted areas, accurate bathymetry for coastal areas	International Hydrographic Organisation	Global			Not free	Data scheme IHO-S-57 for Electronic Navication Charts (ENC)[48]

Table 5: Inventory of available datasets (cont)

Type	Dataset	Description	Data source	Spatial coverage	Temporal coverage	Availability	Licence/Ter for use	Available formats
METOC data	Mediterranean Sea Physics Analysis and Forecast	Model data (forecast and re-analysis) from Mediterranean Sea from Mediterranean Forecasting System, a coupled hydrodynamic-wave model with data assimilation components. The Nucleus for European Modelling of the Ocean (NEMO) www.nemo-ocean.eu/ is used for the hydrodynamics and WaveWatch-III polar.ncep.noaa.gov/waves/index2.shtml for the waves. Regional and Global data available from other models from marine.copernicus.eu	medforecast.bo.ingv.it	Mediterranean Sea, 1/16° of resolution (6-7 km, approximately) Regional and Global data available from other models. Cf. marine.copernicus.eu	Daily updates. Historical data available from 2013.	marine.copernicus.eu/web/69-interactive-catalogue.php?option=com_cswview=details&product_id=MEDSEA_ANALYSIS_FORECAST_PHYS_006_001_a. FTP dowload from catalogue; Phyton scripting download; Catalogue Service for the Web (CSW)	Open data. Registration required (marine.copernicus.eu)	NetCDF. Web Map Service (WMS) visualization
	Model data from Copernicus Marine Service	Regional and Global model data (forecast and re-analysis) and observations	marine.copernicus.eu	Global. Regional: Mediterranean, North Sea, Baltic Sea, Iberian Sea, Black Sea, Artic Sea	Daily updates. Historical data available.	FTP dowload from marine.copernicus.eu catalogue; Phyton scripting download; Catalogue Service for the Web (CSW)	Open data. Registration required (marine.copernicus.eu)	NetCDF. Web Map Service (WMS) visualization
	Global Ocean Wind		marine.copernicus.eu		Hor. res of .25x.25°, 6 hours. daily and monthly mean from 2007.	FTP dowload marine.copernicus.eu/web/69-interactive-catalogue.php?option=com_csw&view=details&product_id=WIND_GLO_WIND_L4_NRT_OBSERVATIONS_012_004;	Open data. Registration required (marine.copernicus.eu)	NetCDF.
	ECMWF datasets		European Centre for Medium-Range Weather Forecasts (ECMWF)			apps.ecmwf.int/datasets/	Open data.	
	ECMWF Ocean Wave model data	ERA-Interim model (ocean waves, atmospheric variables), ERA-20C model	European Centre for Medium-Range Weather Forecasts (ECMWF) 2010	Global	ERA-Interim: daily until 2015. ERA-20C model: daily until	apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/, apps.ecmwf.int/datasets/data/era20c-daily/levtype=sfc/type=an/	Open data.	
	Adriatic Forecasting System (AFS) data	Forecast and simulation data on sea conditions based on Princeton Ocean model (POM) model	www.ionioproject.eu	Adriatic sea	Last 7 days	OpenDAP (THREDDS Data Server) http://tds.tessa.cmcc.it/thredds/catalog.html; FTP (needs account from ocean-labcmcc.it)		NetCDF [49]
	Ionian Forecasting System (IFS)	Forecast data from SANI2 (Southern Adriatic âĂŞ Northern Ionian Sea 2) model, which was developed on the basis of the Princeton Ocean model (POM)	ionioproject.hcmr.gr	Ionian Sea	Last 3 months	OpenDAP (Hyrax data server) tethys.hcmr.gr/opendap/hyrax/ionio/	Public	NetCDF
	Observations data from IONIO In-situ database	Observations from buoys and ship of opportunities	www.mediterranean-marinedata.eu/ionio/home.htm	Adriatic Sea	From 1986	Web interface for download http://www.mediterranean-marinedata.eu/ionio/home.htm	Public releasable and Confidential data. Authorization is required for download	NetCDF, Medatlas
	Sea conditions	Significant Wave Height (SWH)	sea-conditions.com				Data not available for download from web	
	Energy data from Seabed Habitat EMODnet	Seabed Habitat datasets, including Energy data: wind, waves and currents harmonized at European Level	Copernicus Marine Service, ECMWF, IOWAGA	European		http://www.emodnet-seabedhabitats.eu/default.aspx?page=1934	Open data	WMS service
	Physics data from EMODnet	Sea water temperature, salinity or density, water currents, level, Waves and winds, Light attenuation, Atmospheric parameters at sea level, HF radar data	EuroGOOS associates and their regional operational systems (ROOSs), CMEMS (Copernicus) in-situ TAC system and the SeaDataNet consortia	European	last 60 days	Web GIS: http://www.emodnet-physics.eu/map/ Catalogue: http://www.emodnet-physics.eu/Map/service/Catalogue.aspx; THREDDS catalogue http://thredds.emodnet-physics.eu:8080/thredds/catalog.html	Open data	WMS, Web Feature Service (WFS), Web Sensors service (SOS)
	National Oceanic and Atmospheric Administration (NOAA) datasets	Global meteorological and oceanographic datasets from cooperating networks of ships and buoys	Global	Historical data available	Catalogue www.ndbc.noaa.gov/data	Open Data		

Table 6: Inventory of available datasets (cont)

Type	Dataset	Description	Data source	Spatial coverage	Temporal coverage	Availability	Licence/Ter for use	Available formats
Event reports	Blacklists of IUU fishing vessels	Official Blacklist of IUU fishing vessels compiled by Greenpeace from publicly available official registries of IUU vessels and companies		Global		http://blacklist.greenpeace.org/0/vessel/list	Open Data	online resource
	NGA WTS	Worldwide Threat to Shipping (WTS) reports						
	Interpol reports	Reports on illegal fishing events		Global		http://www.interpol.int/INTERPOL-expertise/Notices/Purple-notices-%E2%80%93-public-versions/2014	Public	online resource, PDF
	Fisheries facts from news, social medias							Unstructured
	ASAM datasets	reports on piracy events						
	IMB pirate map	map of piracy events						
	pirate attacks 2007							
	pirate attaks no date							
	Global Maritime Piracy Database from Interpol							

5 CONCLUSIONS AND FUTURE STEPS

This report proposes a description of a maritime surveillance use case as well as several relevant scenarios addressing challenging problems for the maritime community.

- We plan to hold a meeting in M05 which would serve as a second-round validation of the maritime use case with the help of the User Case Interest Group (UCIG);
- The list of Maritime Situational Indicators (MSI) is an initial proposal issued from the survey of several national and NATO workshops aiming at capturing the operational information needs for maritime surveillance purposes. However, it is expected that the list or structure may evolve during the project;
- Equivalently, the list of scenarios proposed may be enlarged or refined during the project to eventually reflect specific operational needs or interests. However, this should not affect the generic scientific work to be performed within.

References

[1] Sten Andler, M. Fredin, F. Gustafsson, J. va Laere, M. Nilsson, and P. Svenson. SMARTracIn - a concept for spoof resistant tracking of vessel and detection of adverse intentions. In *SPIE Defense, Security, and Sensing*, Orlando, FL, 2009.

[2] David Brosset Damien Le Guyader, Cyril Ray. Defining fishing grounds variability with automatic identification system (ais) data. In *2nd International Workshop on Maritime Flows and Networks (WIMAKS'16)*, 2016.

[3] Mica R. Endsley and Daniel J. Garland. *Situation Awareness Analysis and Measurement*. Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey, 2000.

[4] Benjamin S Halpern, Melanie Frazier, John Potapenko, Kenneth S Casey, Kellee Koenig, Catherine Longo, Julia Stewart Lowndes, R Cotton Rockwood, Elizabeth R Selig, Kimberly A Selkoe, et al. Spatial and temporal changes in cumulative human impacts on the world/'s ocean. *Nature communications*, 6, 2015.

[5] Niels T Hintzen, Francois Bastardie, Doug Beare, Gerjan J Piet, Clara Ulrich, Nicolas Deporte, Josefine Egekvist, and Henrik Degel. Vmstools: open-source software for the processing, analysis and visualisation of fisheries logbook and vms data. *Fisheries Research*, 115:31–43, 2012.

[6] Anne-Laure Jousselme, Karna Bryan, Michael Ilteris, and Eric Reardon. Selection of use cases to support Maritime Situation Awareness algorithmic development. Technical memorendum, NATO Centre for Maritime Research and Experimentation (CMRE), La Spezia (IT), 2016. To be published.

[7] R. O. Lane, D. A. Nevell, S. D. Hayward, and T. W. Beaney. Maritime anomaly detection and threat assessment. In *Proc. of the 13th Int. Conference on Information Fusion*, Edinburgh, UK, 2010.

[8] Rikard Laxhammar. Anomaly detection for sea surveillance. In *Proc. of the Int. Conference on Information Fusion*, Firenze, Italy, July 2008.

[9] Janette Lee, Andy B South, and Simon Jennings. Developing reliable, repeatable, and accessible methods to provide high-resolution estimates of fishing-effort distributions from vessel monitoring system (vms) data. *ICES Journal of Marine Science: Journal du Conseil*, page fsq010, 2010.

[10] James Llinas, Christopher Bowman, Galina Rogova, Alan Steinberg, Ed Waltz, and Frank White. Revisiting the JDL data fusion model II. In *Proceedings of the 7^{th} International Conference on Information Fusion*, pages 1218–1230, Stockholm, Sweeden, June-July 2004.

[11] Jim Llinas, Anne-Laure Jousselme, and Geoff Gross. Context as an uncertain source. In L. Snidaro, J. Garcia, J. Llinas, and E. Blasch, editors, *Context Enhanced Information Fusion: Improving real world performance with domain knowledge*. Springer, 2015.

[12] Fabio Mazzarella, Michele Vespe, Dimitrios Damalas, and Giacomo Osio. Discovering vessel activities at sea using AIS data: Mapping of fishing footprints. In *17th International Conference on Information Fusion, FUSION 2014, Salamanca, Spain, July 7-10, 2014*, pages 1–7, 2014.

[13] Pete McBreen. Using use cases for requirements capture. www.mcbreen.ab.ca, McBreen.Consulting, 1998.

[14] M. Monperrus, B. Long, J. Champeau, B. Hoeltzener, G. Marchalot, and J.-M. Jézéquel. Model-driven architecture of a maritime surveillance system simulator. *Systems Engineering*, 13(3):290–297, 2010.

[15] Fabrizio Natale, Maurizio Gibin, Alfredo Alessandrini, Michele Vespe, and Anton Paulrud. Mapping fishing effort through ais data. *PloS one*, 10(6):e0130746, 2015.

[16] Giuliana Pallotta, Michele Vespe, and Karna Bryan. Vessel pattern knowledge discovery from AIS data - A framework for anomaly detection and route prediction. *Entropy*, 5(6):2218–2245, 2013.

[17] GJ Piet, FJ Quirijns, L Robinson, and SPR Greenstreet. Potential pressure indicators for fishing, and their data requirements. *ICES Journal of Marine Science: Journal du Conseil*, 64(1):110–121, 2007.

[18] M. A. Razzaque, S. Dobson, and P. Nixon. Classification and modeling of the quality of contextual information. *Journal of Autonomic and Trusted Computing*, 2006.

[19] Council Regulation. No 2371/2002 of 20 december 2002 on the conservation and sustainable exploitation of fisheries resources under the common fisheries policy. *OJ L*, 358(31.12):2002, 2002.

[20] Jean Roy and Mike Davenport. Exploitation of maritime domain ontologies for anomaly detection and threat analysis. In *Proc. of the IEEE int. Waterside Security Conference (WSS)*, 2010.

[21] Lauro Snidaro, L. Vaci, J. Garcia, E. D. Marti, A.-L. Jousselme, K. Bryan, D.D. Bloisi, and D. Nardi. A framework for dynamic context exploitation. In *Proceedings of the 18th International Conference on Information Fusion*, Washington, D. C. (USA), July 2015.

[22] STANAG 4162. Identification Data Combining Process (edition 3) - AIDPP-01 Edition 1. Technical report, NATO Standardization Agency, 2015. NATO UNCLASSIFIED.

[23] Alan N. Steinberg and Christopher L. Bowman. Rethinking the jdl data fusion levels. In *National Symposium on Sesnor and Data Fusion*, 2004.