Dissemination Plan
Work Package 7 Task 7.2 Deliverable 7.2

Authors
Konstantinos Chatzikokolakis, Konstantina Bereta,
Giannis Spiliopoulos, Dimitris Zissis
MarineTraffic

Distribution list:

Group:	Others:
WP Leader: MarineTraffic Task Leader: MarineTraffic	Internal Reviewer: Center for Genomic Regulation (CRG) INFORE Management Team INFORE Project Officer

Document history:

Revision	Date	Section	Page	Modification
0.5	05/06/2019	All	All	Creation
0.7	13/06/2019	All	All	Athena's internal review comments incorporated
0.8	15/06/2019	All	All	CRG's internal review comments incorporated
1.0	28/06/2019	All	All	Self-review

Approvals:

First Author: Konstantinos Chatzikokolakis (MT) Date: 05/06/2019

Internal Reviewer: Gian Gaetano Tartaglia (CRG) Date: 15/06/2019

Coordinator: Antonios Deligiannakis (Athena) Date: 28/06/2019

Table of contents:

1 Executive Summary

The purpose of this document is to describe the dissemination plan of INFORE. The dissemination plan belongs to the context of the WP7, which is responsible for the coordination of activities with the aim of communicating the novel ideas and techniques developed in the context of the project to the wider public, especially to the research communities and industrial organisations and companies that are relevant to the project and could potentially benefit from the outcomes of the project. In this context, different communication means will be used to disseminate the project results (e.g., scientific publications, presentations, participation in industrial events, social media posts, etc.) and different user communities will be targeted. The communication means and methodologies, the target audience, as well as the different dissemination channels that will be used are described in detail in this document. The dissemination activities that will be carried out in the course of the project are reported in deliverable D7.3 in two phases, one covering the first 18 months of the project which is due M18 and the second covers the rest of the project and is due M36.

2 Introduction

The purpose of the dissemination and communication plan is to set out the key steps and activities that will ensure maximum stakeholder engagement and project awareness. Throughout this document, dissemination is referred to as the public disclosure of the results of a project across any communication means. It is the process of promoting and generating awareness about the project from the very beginning. Its aim is to make the project results known to various stakeholder groups in a targeted way. Communication refers to taking measures for promoting the action itself and its results to a multitude of audiences including the broad public, while engaging in a two-way exchange of information (European Commission, 2016).

In the context of this document we identify the targeted audience that could be interested in using the outcomes of the project. Then, we define specific communication channels to reach out to the various target groups and we also define the mechanisms for efficient communications between the project and with the targeted audience and the wider public.

The structure of the document is described as follows. In the rest of this section we describe the project overview and objectives. In Section 3 we present the dissemination objectives, the targeted audience, and we decribe in detail the dissemination plan, including the dissemination activities that we plan to perform throughout the project in accordance with the dissemination objectives. In Section 0 we describe how the dissemination activities carried out in the context of the project are monitored. Last, in Section 0 we conclude this report.

2.1 Project overview

At an increasing rate, industrial and scientific institutions need to deal with massive data flows streaming in from a multitude of sources. For instance, maritime surveillance applications combine high-velocity data streams, including vessel position signals emitted from hundreds of thousands of vessels across the world and acoustic signals of autonomous, unmanned vessels; in the financial domain, stock price forecasting and portfolio management rely on stock tick data combined with real-time information sources on various pricing indicators; at the fight against cancer, complex simulations of multi-cellular systems are used, producing extreme-scale data streams in an effort to predict the effects of drug synergies on cancer cells. In these applications, the data volumes are expected to dramatically grow in the future.

Processing this data often requires not only using a High Performance Computing (HPC) infrastructure, but also having data scientists, who are typically not expert programmers, program complex workflows, with a vast number of parameters to tune through time-consuming repeated programming and testing. INFORE will address these challenges and pave the way for real-time, interactive extreme-scale analytics and forecasting. The ability to forecast, as early as possible, a good approximation to the outcome of a time-consuming and resource-demanding computational task allows to quickly identify undesired outcomes and save valuable amount of time, effort and computational resources, which would otherwise be spent in vain. Consider, for example, the ability to forecast the outcome of a complex multi-cellular system simulation for tumor evolution, without the need to wait for the simulation to be completed. INFORE will also design and develop a flexible, pluggable, distributed software architecture that is programmable and set up by graphical data processing workflows. The INFORE prototype will be tested on massive real-world data from the life sciences, financial and maritime domains.

2.2 Project objectives

The scientific and technological objectives of INFORE are described as follows:

O1. Real-time, interactive machine learning and data mining tools, supporting the interactive construction of highly accurate models from extreme-scale data streams and massive data volumes, allowing users to gain fast insights on the properties of the data they are dealing with.

O2. Distributed complex event forecasting to further support interactive analytics. This involves innovative techniques for adaptive, low-latency, complex event processing, allowing not only for the timely detection of critical events as they occur, but also, forecast of their future occurrences. These techniques will exhibit the necessary expressivity for capturing the complex phenomena of real-world applications, and they will be supported by machine learning tools for learning event patterns from data.

O3. A flexible, pluggable and extendable architecture, supported by corresponding software stacks, which will make the interactive data analytics library easy-to-use and increase its effectiveness, e.g. by allowing users to easily compose data analytics pipelines from the existing tools in the library, while supporting the real-time incorporation of new knowledge by machine learning algorithms.

O4. Construction of a framework for supporting non-programmer data analysts to specify processing workflows and data analytics tasks. The framework will consist of a family of data processing operators that can be interconnected to provide a family of complex data processing tasks with minimal or no programming overhead.

O5. Data summarization and approximate query processing techniques to allow for real-time response times while exploring representative views of massive, high-velocity input. These techniques involve novel, highly optimized algorithms for generating data synopses, ranging from one-pass synopses over single-source data to more elaborate techniques that monitor cross-source correlations, joins, etc.

O6. Rigorous testing and evaluation, involving controlled experiments and reviews by domain experts, with real life data from the financial, the maritime and the life sciences domains.

The technological advancements of INFORE, as described above, will be showcased in the following three use cases:

- The life sciences use case, paving the way for the development of new cancer treatments, by identifying personalised drug combinations, which, by acting synergistically, will be able to fight cell resistance in target therapies and ultimately increase the patients' life expectancy. To this end, the use case will employ in-silico models of multi-cellular systems found in in-vivo tumors, to identify promising combinations of drugs. The simulations, which will be carried-out on BSC's High Performance Computing infrastructures, are extremely complex; they involve billions of cells and generate approximately 100GB of data per minute.
- The maritime use case, improving maritime situational awareness, i.e. the ability to perceive and forecast activities and threats in maritime environments. For a complete picture of maritime activities, we will combine global maritime surveillance systems, such as the AIS (Automatic Identification System), with local autonomous unmanned vehicles, such as Wavegliders. This way, we will advance the state-of practice by identifying and forecasting the activities of "dark targets" that (intentionally) hide from traditional monitoring systems. The maritime data available in INFORE generate streams of approximately 1TB of data per day.
- The financial use case, forecasting price swings of stocks, currencies, commodities and systemic risk, and offering decision support for investment opportunities. The financial data available in the project include a variety of market data, including stock market and crypto-currencies market data, arriving in tens of thousands of correlated, high-velocity streams, for a total of more than 450GB of data per day.

3 Dissemination and Communication Plan

3.1 Dissemination Objectives

The dissemination objectives of INFORE are described as follows:

- Identify the target audience, i.e., the different categories of organisations, communities and individuals that could benefit from the outcomes of the project.
- Identify the dissemination and communication channels available
- Specify and coordinate the communication channels that are more appropriate in order to communicate the project results to the differet target groups in order to maximise impact and user engagement.
- Plan and monitor the dissemination activities in accordance with the KPIs that have been defined in the grant agreement.

3.2 Dissemination and communication roles and points of contact

Role	Responsibility	Assignee
Project Coordinator	Constant monitoring of the project progress and results. Ensuring that the progress is made to ensure that the project is on track with respect to the project objectives	Antonios Deligiannakis (ATHENA)
Dissemination Manager	Coordination of the communication and dissemination activities of the project. Monitoring the dissemination activities with respect to the KPIs defined in the grant agreement.	Konstantinos Chatzikokolakis (MarineTraffic)
Exploitation and Innovation Manager	Coordination of all exploitation activities of the project, guaranteeing consistency between technical and market choices.	Konstantinos Chatzikokolakis (MarineTraffic)
Work Package 7 Leader	Coordination of the WP7 tasks and monitoring their progress, ensuring that the work package objectives are met and that all WP7 deliverables are of high quality and delivered on time.	Dimitrios Zissis (MarineTraffic)

3.3 The Dissemination, Exploitation and Innovation Board (DEIB)

During the INFORE Kick off meeting the consortium decided to set up a Dissemination, Exploitation and Innovation Board so as to improve the dissemination, exploitation and innovation activities of each of the three use cases of the project. Each consortium member has identified one representative in this board, responsible to communicate its activities to rest of the consortium members. The following table lists the selected representatives of the INFORE consortium.

Partner	Assignee
ATHENA	Antonios Deligiannakis
NCSR	Alexander Artikis
RM	David Arnu
BSC	Arnau Montagud
CRG	Michele Monti
SPRING	Holger Arndt
MT	Konstantinos Chatzikokolakis
CMRE	Elena Camossi

3.4 Dissemination Channels

In this section we present the main dissemination channels that will be used for the dissemination and communication activities of the project.

3.4.1 Theme and key messages

Themes and key messages give a brief overview of the vision and objectives of the project and can be used in the context of various communication activities (e.g., they can be integrated in talks and presentations about the project). The messages have been agreed upon by all partners and provide the framework for any promotional speeches, presentations, and interactions. The summary, vision, scientific and technical objectives, as well as acknowledgement and key message are detailed below.

Table 1: Theme and key messages

Summary	The aim of the INFORE project is to address the challenges posed by huge datasets and pave the way for applying real-time, extreme-scale analytics, deriving hidden information and forecasting to support interactive analytics. Today, at an increasing rate, industrial and scientific institutions need to deal with massive data flows, streaming-in from maritime surveillance applications, financial forecasting applications or cancel cells growth simulations as well as a multitude of other sources. The ability to forecast, as early as possible, a good approximation to the outcome of a time-consuming and resource demanding computational task allows to quickly identify undesired outcomes and save valuable amount of time, effort and computational resources.
Vision	INFORE envisions to provide a flexible, pluggable, distributed software architecture that will pave the way for real-time, interactive extreme-scale analytics and forecasting and help data scientists reduce the time consumed to program complex workflows.
Scientific and Technical Objectives	1. Real-time, interactive machine learning and data mining tools, supporting the interactive construction of highly accurate models from extreme-scale data streams and massive data volumes, allowing users to gain fast insights on the properties of the data they are dealing with. 2. Distributed complex event forecasting to further support interactive analytics. This involves innovative techniques for adaptive, low-latency, complex event processing, allowing not only for the timely detection of critical events as they occur, but also, forecast of their future occurrences. These techniques will exhibit the necessary expressivity for capturing the complex phenomena of real-world applications, and they will be supported by machine learning tools for learning event patterns from data. 3. A flexible, pluggable and extendable architecture, supported by corresponding software stacks, which will make the interactive data analytics library easy-to-use and increase its effectiveness, e.g. by allowing users to easily compose data analytics pipelines from the existing tools in the library, while supporting the real-time incorporation of new knowledge by machine learning algorithms. 4. Construction of a framework for supporting non-programmer data analysts to specify processing workflows and data analytics tasks. The framework will consist of a family of data processing operators that can be interconnected to provide a family of complex data processing tasks with minimal or no programming overhead. 5. Data summarization and approximate query processing techniques to allow for real-time response times while exploring representative views of massive, high-velocity input. These techniques involve novel, highly-optimized algorithms for generating data synopses, ranging from one-pass synopses over single-source data to more elaborate techniques that monitor cross-source correlations, joins, etc. 6. Rigorous testing and evaluation, involving controlled experiments and reviews by domain experts, with real life data from the financial, the maritime and the life sciences domains.

Acknowledgement	This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825070.
Key message	Interactive Extreme-Scale Analytics and Forecasting

3.4.2 Project identity

One of the most important elements that identify a project is the project logo. The marketing team of MarineTraffic and SPRING designed several candidate logos so that the consortium would then decide which one was the best representative of the project. The final selection that was made by the consortium is depicted in Figure 1 (and the rest of the candidate logos are listed in the Appendix A). The name of the project occupies the most central part of the logo so that when the logo is used in public documents (e.g., posters, flyers, presentations), the credits to the project are clearly stated.

The right side of the logo describes the main technological breakthrough of the project, which is addressing the challenge of providing a platform or performing interactive extreme-scale analytics and forecasting. The rationale behind this design choice is twofold: First, we wanted to link the project name with its technological breakthrough. This enables someone with no background knowledge of the project that sees the logo for first time, to get a highly representative view of the main objective of the project. Second, since providing solutions for interactive big data analytics and forecasting is a very hot topic in both academia and industry, we wanted to make the logo, and as a result, the project, more approachable to our target audience.

On the left side of the logo there is a group of four icons comprising one icon representing each use case and an icon of a binary representation which is commonly used to represent computing. In this way, it is easy for someone to link the main technical challenge that the project addresses, which is mentioned in the right part of the logo, to its application to the use cases, which is portrayed in the left side of the logo. Since our goal is to enable someone that sees the project logo for first time to understand what the project is about, we believe that the project logo that is depicted in Figure 1 is a visual manifestation of the project outcomes.

Figure 1: Project logo

3.4.3 Media kit

Our media kit consists of a presentation presenting the project and a factsheet. The presentation gives an overview of the project, including the technical challenges, the technological contributions to the state-of-the-art, and explains how the solutions that are being developed apply to the three use cases. The factsheet summarises the aforementioned topics. The presentation and the factsheet will be updated throughout the project duration in order to reflect its current status. The initial version of the project presentation and factsheet have been reported in D7.1.

3.4.4 Online dissemination and use of Web 2.0 channels

The most important online dissemination channel is the project website (http://infore-project.eu/). The website of the project gets regularly updated and contains the following information:

- It presents an overview of the project, describing the challenges, main technological advancements and the use cases.
- It includes news items to inform the general public about news and events.
- It serves as a gateway to all other dissemination channels, as it provides links to social media platforms, publications, and will constantly get updated to accommodate all dissemination material (white papers, brochures, videos, etc.) developed throughout the project.
- It will include links to public tangible results of the project, such as open source code, demos and public technical reports.

- It includes contact details[1] so that people that are interested to learn further information about the project, request access for more resources or seek possible collaboration can get in touch with the consortium.

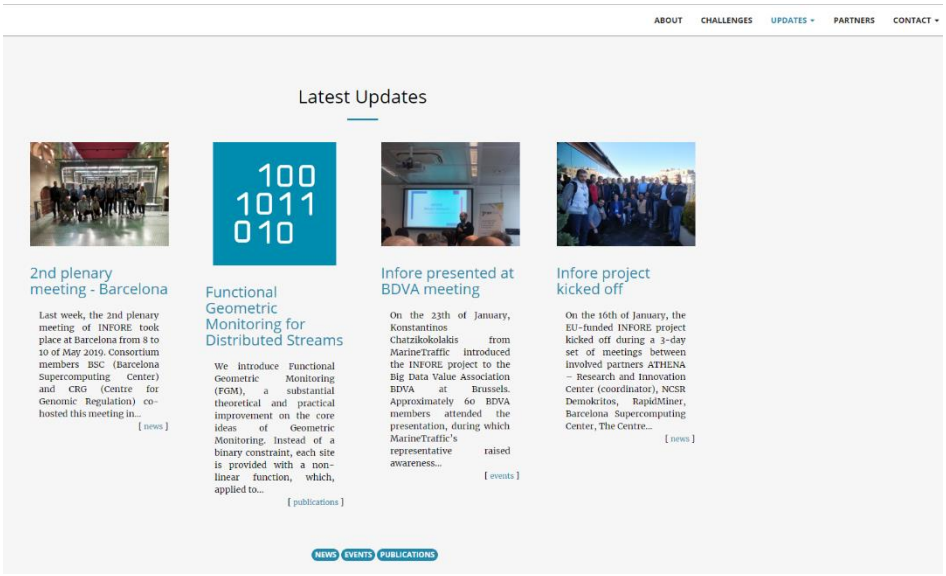

Figure 2: Website updates section

Another important dissemination channel are the social media platforms. We have created a dedicated account for the project in each one of the platforms LinkedIn[2], Facebook[3], Twitter[4] and YouTube[5], where we publish latest news about the project (e.g., publications, events, milestones, etc.). Furthermore, we have set up a ResearchGate project[6] and a Zenodo community[7], where we intend to share our scientific publications (in accordance with the original publisher's terms of use).

Figure 3 is a snapshot of the INFORE twitter account. The real-time nature of Twitter is used to inform the wider public about current project developments and events that take place. For example, we post tweets about presentations that are about to take place so that participants of the respective events can attend. INFORE Facebook account is used in a similar way, and a snapshot of the Facebook account is provided in Figure 4. LinkedIn is a business-oriented platform in which we have created a dedicated account and a group for the project. A snapshot of the group page of INFORE is depicted in Figure 5.

[1] infore-contact@googlegroups.com
[2] https://www.linkedin.com/groups/13683100
[3] https://facebook.com/infore.project
[4] https://twitter.com/InforeProject
[5] http://www.youtube.com/channel/UCQRj2coemjbc1UUnkxcSpYw
[6] https://www.researchgate.net/project/INFORE-H2020-EU-grant-No-825070
[7] https://zenodo.org/communities/infore-project/

Figure 3: INFORE Twitter account

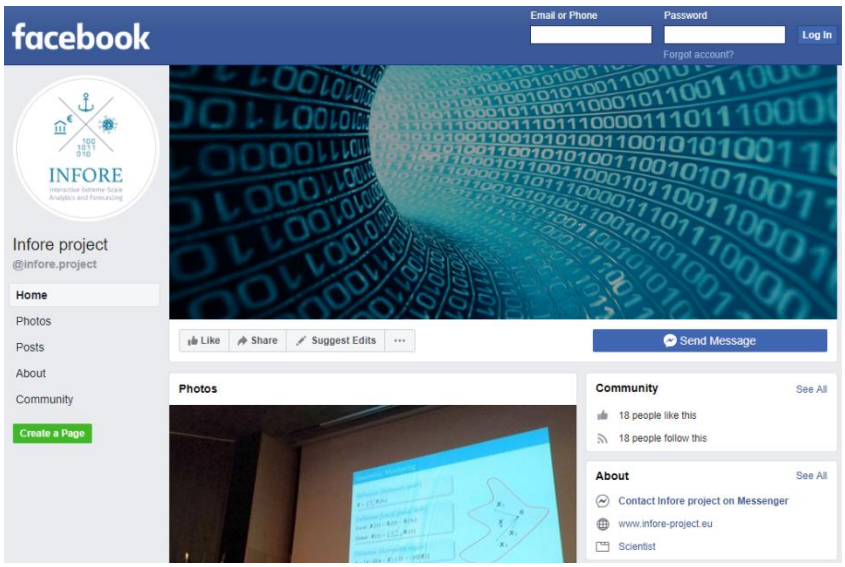

Figure 4: INFORE Facebook account

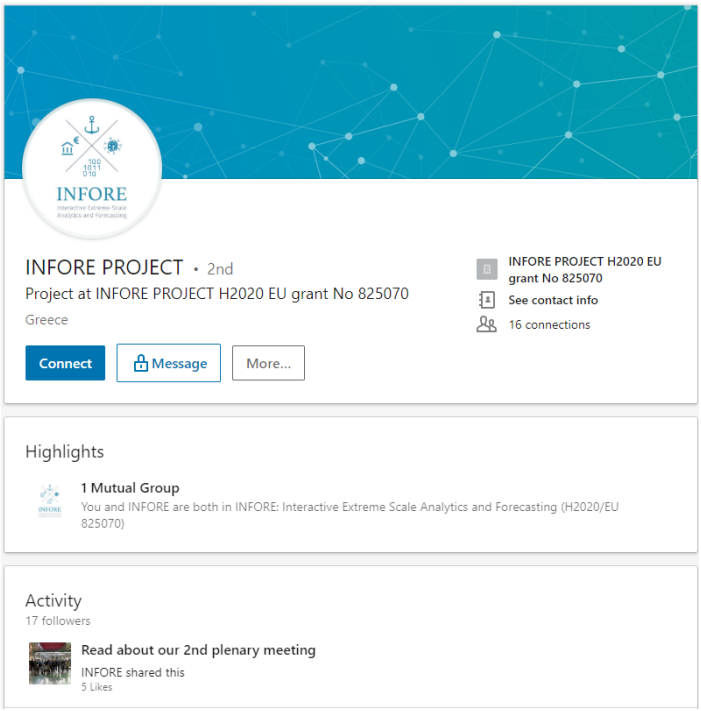

Figure 5: INFORE LinkedIn account

3.4.5 Research and Academic oriented dissemination

Research and academic oriented dissemination refers to the commitment of the consortium in organising scientific and industry workshops in conjunction with such events, especially dedicated to raise awareness of potential service providers/users and researchers. Furthermore, members of the consortium are frequently invited to give Keynote speeches and talks in reputable related events and it is a common practice of INFORE consortium members to provide talks in tutorial sessions of the above events. All such opportunities will be exploited to raise awareness about the visions, progress and outcomes of INFORE.

Scientific dissemination will take place through publications of research results in top tier and/or reputable related conferences and journals, the inclusion of research results as state-of-the-art in lecture materials, internal seminars organised by the consortium members, training of students in the respective technologies and encouraging them to join the research efforts as Masters/PhD students and increasing awareness of the project achievements within national and regional public institutions.

Members of the consortium are currently editors and authors of highly cited papers and journals in their field generating confidence regarding research/academic dissemination. Table 2 shows a list of target conferences were the members of the consortium intend to publish the outcomes of the projects according to the area of expertise of each partner.

Table 2: Research venues

RTD events
ACM International Conference, Special Interest Group on Management of Data (SIGMOD) International Conference on Very Large Data Bases (VLDB) IEEE International Conference on Data Engineering (ICDE) International Conference on Extending Database Technology (EDBT) International Conference on Database Theory (ICDT) ACM International Conference on Distributed and Event-Based Systems (DEBS) European Conference on Machine Learning (ECML)

RTD events
International Conference on Machine Learning (ICML) European Conference on Artificial Intelligence (ECAI) International Conference on Artificial Intelligence (IJCAI) Oceans IEEE/MTS (OCEANS) International Conference on Information Fusion (FUSION) Intelligent Systems for Molecular Biology Conference (ISMB-ISCB) European Conference on Computational Biology (ECCB) International Conference on Research in Computational Molecular Biology (RECOMB) International Conference on Genome Informatics (GIW/BIOINFO)

3.4.6 Commercially oriented dissemination

Commercially oriented dissemination activities aim at increasing visibility to stakeholders that might be interested in the take-up of the technology. Commercially oriented dissemination, thus, includes presentation of results at fairs and exhibitions, on user forums, to industrial consortia and to regulatory bodies and in workshops and seminars to inform professionals about the services developed in the context of the project. Table 3 presents a list of the most important commercially-oriented venues that we plan to target with respect to the area of expertise of each partner.

Table 3: Commercially oriented dissemination activities

Commercial-oriented events
Hannover Messe is the world's most important technology show, a powerful driver of investment in new technology and automation. Hannover Messe is where businesses showcase their latest products and solutions. **European Big Data Value Forum** The EBDVF is a meeting place for industry, research, policymakers and community initiatives to discuss the challenges of Big Data and the emerging Data Economy and to develop suitable action plans for addressing these challenges. **CeBIT** is the world's largest computer expo held yearly in Hannover, Germany, a recognized meeting place for disseminating the latest research results in ICT. **World of Trading** offers a trading exhibition revolving around the theme of TRADING with companies representing banks, brokers, issuers, professional associations, software, hardware, further education, specialist publishers and many others. **Poseidonia** is an international shipping exhibition, which takes place every two years, and which is the largest meeting place for the Greek shipping industry and international transportation experts. The fair offers direct access to the Greek fleet, but also to international industry leaders. National and international exhibitors show the latest innovations in the boat market and present spare parts, boat accessories and all services related to the boat sector. **Shipping2030 Europe** is a brand-new disruptive technology global series presenting the future of shipping: Going digital, new business models, and new partnerships to create digital ecosystems. **SeaFuture** is the must attend event on the sector of Maritime Technologies, involving Navies and companies and including highly specialized conferences with most qualified speakers in the Maritime sector. **BitKom Big-Data.AI Summit** is the largest Business IT interest group in Germany during which a lot of smaller conferences and events take place, where we can hop on if we have a good presentation. **RapidMiner Wisdom** is the RapidMiner's user and customer trade show and conference. **IDS** (Industrial Data Science conference) is co-organized by RapidMiner and is traditionally use-case and industry oriented. **Predictive Analytics World** is a business-oriented conference series on various topics (industry, business, deep learning) with high level speaker and audience. **AI4U** is held in parallel with TDWI, a very big trade show/conference for business intelligence, typically held on June in Munich, Germany.

3.5 Target Audience

In this section we specify the different target groups that will be addressed by the dissemination and communication activities of the project.

Table 4: Description of the different target groups

Target groups	Stakeholders per group
End users	Enterprises, entrepreneurs, NGOs, policy makers, researches that are interested in Big Data processing, especially in the Finance, Maritime and Health industry domains
IT Industry Players	IT companies, entrepreneurs, software engineers undertaking tasks requiring big data processing, streaming data processing and/or involving workflows containing machine learning tasks
Industry Associations & Technology Clusters	European initiatives and clusters (like BDVA and FIWARE), research communities, associations, federations (like IMS, EFFRA, IFIP, IEEE, NEM)
H2020 Programme Stakeholders	Participants, project partners and relevant stakeholders active in the H2020 projects
Researchers and Academia	Research communities in the areas of Big Data management, stream processing, complex event detection and forecasting, anomaly detection
Policy Makers & Standardisation Organisations	Policymakers at any level like EC Directorates and Units, Ministries and Governments, Regulatory Agencies, Standardisation Organisations (CEN, ISO, ETSI, etc.)
General Public	Civil society representatives, general public and anyone interested in the project

3.6 Dissemination and Communication Activities

In this section we describe the dissemination and communication activities that we have performed so far in the project.

3.6.1 Dissemination Activities per Target Group

Since INFORE is a Research and Innovation project, both research and industrial communities should be targeted. Different types of dissemination activities are targeted to different communities. Table 5 shows the dissemination activities that will be carried out per target group.

Table 5: Dissemination Activities grouped by Targeted audience

Target group	Dissemination Activities
Regulatory bodies	• Information via existing channels of the industrial/scientific project partners • Involvement in requirements collection and expert evaluation • Presentation of project results on associated events
Service and technology providers	• Contributions to special-interest groups and magazines • Participating in and organizing workshops, exhibitions, and fairs • Contributions to relevant industrial consortia • White papers
Potential End Users	• Dissemination via channels to related industry and SMEs via their associations, specific meetings, and fairs • Contribution to on-line forums, and mailing lists
Researchers	• Conference presentations • Contribution to special-interest newsletters • Publications of papers on high-quality conferences and in renowned scientific journals within relevant fields of expertise • Contact to other EU projects especially in the field of (Big) Data Analytics

Target group	Dissemination Activities
	• Sharing open source software on publicly available software repositories (e.g., http://github.com/)
General Public and Interested Communities	• Web Page • Social Networking • Press Releases

3.6.2 Dissemination Plan and Calendar

Table 6 describes the main dissemination activities that have already been planned for the following year.

Table 6: Short-term planned dissemination activities

Event Name	Type of Event	Date	Location	Participation	Type of Audience	Partners
BDVA Activity Group Meeting	Workshop	22-23/01/2019	Brussels	Presentation	Research Industry Governmental organisations	MT
BDVA Activity Group Meeting	Workshop	27/02/2019	Brussels	Presentation	Research Industry Governmental organisations	ATHENA
EDBT/ICDT 2019	Conference	26-29/03/2019	Lisbon	Presentation	Research	ATHENA
DEBS	Conference	25-28/6/2019	Darmstadt	3 Papers, Demo	Research	NCSR, CMRE, ATHENA
BDV PPP Meetup	Workshop	26-28/06/2019	Riga	Presentation	Research Industry Governmental organisations	MT
FUSION 2019	Conference	02-05/07/2019	Ottawa	Paper	Research	MT
ICT Proposers Day		19-20/09/2019	Helsinki	Presentation	Research Industry Governmental organisations	MT
Minds Mastering Machines	Conference	30/09/2019	London	Presentation	Research Industry	RapidMiner
KDML	Workshop	30/09-02/10/2019	Berlin	Paper	Research	RapidMiner
1st Maritime Situational Awareness Workshop (MSAW 2019)	Workshop	08-10/10/2019	La Spezia	Paper	Research, Industry Governmental organisations	CMRE, MT
European Big Data Value Forum	Conference / Exhibition	14-16/10/2019	Helsinki	Presentation / workshop organisation	Research, Industry Governmental organisations	MT, ATHENA
Papis.io	Conference	10/2019	Boston	Presentation	Research, Industry Governmental organisations	RapidMiner
Dagstuhl	Seminar	9-14/2/2020	Saarbrücken	Organisation	Research	NCSR

Event Name	Type of Event	Date	Location	Participation	Type of Audience	Partners
seminar						
EDBT/ICDT 2020	Conference	30/03-02/04/2020	Copenhagen	Paper	Research	All
SIGMOD 2020	Conference	14-19/06/2020	Portland	Paper	Research	All
VLDB 2020	Conference /Demo	31/08-04/09/2020	Tokyo	Paper or Demo	Research	All

3.6.3 Linked Projects and Established Synergies

In this section we describe the projects that are closely related to INFORE in the sense that it builds upon their technological advancements. These projects, as well as their relation to INFORE, are described as follows.

BigDataOcean[8] is an H2020 Innovation Action project with the aim of proposing and validating maritime big data scenarios for the benefit of EU-based companies, organisations and scientists. The Maritime Security and Anomaly Detection pilot of BigDataOcean is closely related to the Maritime Situational Awareness use case of INFORE. MarineTraffic, which leads the Maritime Security and Anomaly Detection pilot of BigDataOcean, developed the Anomaly Detection tool that is able to detect real-time anomalies in vessel trajectories (e.g., proximity events, route deviations). This tool will form the basis of the Maritime Situational Awareness use case of INFORE which is also led by MarineTraffic. The MSA and Anomaly Detection tool of MarineTraffic was selected among the 10 winners of the NITEC innovation challenge 2019[9].

In the context of INFORE, the anomaly detection tool will be enriched with support for complex event forecasting and will be able to early detect complex events such as piracy, smuggling, etc.

QualiMaster- A configurable real-time data processing infrastructure mastering autonomous quality adaptation.
QualiMaster[10] was a STREP project, funded under FP7-ICT-2013-11, which developed a scalable, adaptive, quality-aware platform for processing of financial time series and combining them with streams from social networks. To achieve the desired scalability, the project combined traditional cloud-based computation with hardware accelerators (FPGAs). ATHENA members and SPRING collaborated in that project.

INFORE will exploit the relevant scientific results of QualiMaster related to computation of cross-stream correlations and financial data handling. It will further exploit software related to social network data acquisition. An important distinction between the two projects however comes from the fact that processing in QualiMaster was strongly depended on hardware accelerators (FPGS). Additionally, the approach did not exploit complex event processing technologies in its arsenal, while INFORE has both complex event processing, forecasting and machine learning as well as data mining orientation.

FERARI - Flexible Event pRocessing for big dAta aRchItectures.
FERARI[11] was a STREP project funded under FP7-ICT-2013-11 aimed at paving the way for business users to express complex analytic tasks through a declarative language that supports distributed Complex Event Processing and sophisticated machine learning operators. Effective, real-time event processing was achieved by leveraging in-situ data processing as a first choice for realizing distributed Big Data systems. ATHENA members participated in that project.

Since both FERARI and INFORE focus on real-time, distributed event stream processing, INFORE will exploit the relevant experience of the members acquired by FERARI regarding CEP optimization. However, INFORE will push the research boundaries by considering machine learning and data mining outcomes as input to a complex event processing module, introduce event forecasting, enable interactivity in analytics etc. Moreover, the types of queries

[8] http://www.bigdataocean.eu/site/
[9] https://nitec19.com/innovation-challenge-2019/
[10] http://qualimaster.eu/
[11] http://www.ferari-project.eu/

that are targeted by INFORE differ from traditional complex event processing operators.

The Human Brain Project.
The Human Brain Project[12], funded under FP7-ICT-2013- FET-F, is a 10-year project, part of the FET Flagship Programme. The Sub-project of High-Performance Computing (SP7) aimed at providing supercomputing, Big Data and Cloud capabilities at the exascale. Key topics for research also included the development of novel mathematical methods, programming models and tools, in situ analysis of multi-Petabyte datasets, and real-time visualization. ATHENA members participated in SP7 (2013- 2016).

We expect that important experience regarding distributed in-memory storage, Big Data analytics as well as real-time visualization stemming from Human Brain project outcomes will be valuable in the context of INFORE as well.

LEADS - Large-scale Elastic Architecture for Data-as-a-Service (funded under FP7-ICT-2011-08).
LEADS[13] focused on building and showcasing a novel Cloud service model named Data-as-a-Service, an innovative infrastructure based on an elastic collection of geographically distributed micro-clouds. LEADS aims at providing the means to gather, store, and query publicly available data, as well as real-time data processing. ATHENA members were involved in this project.

LEADS focused only on efficient stream processing. INFORE introduces a new world of challenges, both in terms of scalability (due to the large number of streams/devices) and in terms of its focus (distributed complex event processing/forecasting & learning/ mining tasks). We expect that we might exploit some of the LEADS scientific results in the context of INFORE, but we envision that a much more scalable stream processing platform as well as a distinct set of query processing and optimization algorithms tailored to INFORE will be necessary

Track and Know - Big Data for Mobility Tracking Knowledge Extraction in Urban Areas
Track&Know[14] (funded under H2020- ICT-2017-16) will research, develop and exploit a new software framework that aims at increasing the efficiency of Big Data applications in the transport, mobility, motor insurance and health sectors. Stemming from industrial cases, Track&Know will develop user friendly toolboxes that will be readily applicable in the addressed markets.

INFORE will extend the complex event processing technology of Track and Know by making it applicable to a wide variety application domain, as opposed to restricting it to mobility data, and further optimize it for extreme-scale analytics.

SELIS - Towards a Shared European Logistics Intelligent Information Space
SELIS[15] is an H2020 project aimed at delivering a platform for pan-European logistics applications' by: Embracing a wide spectrum of logistics perspectives and creating a unifying operational and strategic business innovation agenda for pan European Green Logistics. MT members are involved in this project.

The SELIS project will provide the INFORE team with essential knowledge of building Big Data tools across the whole transportation/logistics chain.

MARISA - Maritime Integrated Surveillance Awareness
MARISA[16] H2020 project has as its goal to provide security communities operating at sea with a data fusion toolkit, which makes available a suite of methods, techniques and modules to correlate and fuse various heterogeneous and homogeneous data and information from different sources, including Internet and social networks, with the aim to improve information exchange situational awareness, decision-making and reaction capabilities. CMRE is involved in this project.

[12] https://www.humanbrainproject.eu/
[13] http://www.leads-project.eu/
[14] https://www.iit.demokritos.gr/project/track-and-know
[15] http://www.selisproject.eu/
[16] https://www.marisaproject.eu/

INFORE is in line with CMRE past experience and the activity in MARISA to handle heterogeneous, multimodal data sources.

ARGOMARINE[17] FP7 Automatic Oil Spill Recognition and Geopositioning Integrated in a Marine Monitoring Network
The objective of ARGOMARINE (2009-2012 FP7-TRANSPORT FP7-SST-2008-RTD-1). was to develop and test an integrated system for monitoring of the marine traffic and pollution events due to carriers/commercial ships as well as recreational boats through environmental-sensitive sea areas. A distributed network of underwater moored platforms was proposed and demonstrated by CMRE.

INFORE will leverage on the acquired experience on passive acoustic detection and data fusion to integrate the surface autonomous vessels equipped with passive acoustic sensors in the proposed marine surveillance network. Acoustic data are crucial to complement the global view given by AIS and radar sensor in case of AIS spoofing and weak radar signature.

PERSEUS, "Protection of European Seas and borders through the intelligent use of surveillance."
PERSEUS is an FP7 demonstration project supported by the FP7 Security-Research Theme under DG Enterprise Its aim is to build an EU maritime surveillance system integrating existing national and communitarian installations and enhancing them with innovative technologies. CMRE participated in the project, in which it fitted both a Slocum underwater glider and a WaveGlider with a passive sonar system to detect, localize, track and classify surface shipping.

INFORE will leverage on the acquired experience on the integration of passive acoustic sensors on mobile robots to integrate the sensorised surface autonomous vessels in the marine surveillance network. This experience will help handle problems raised by the movement of the sensor during the robot navigation. As a step forward, multiple platforms with acoustic sensors will be integrated in the INFORE network.

ATHLOS – Ageing Trajectories of Health: Longitudinal Opportunities and Synergies.
The objective of the ATHLOS H2020 Project[18] is to achieve a better understanding of ageing by identifying patterns of healthy ageing pathways or trajectories, the determinants of those patterns, the critical points in time when changes in trajectories are produced, and to propose timely clinical and public health interventions to optimise healthy ageing. SPRING is involved in this project.

INFORE will make use of the statistical methods, researched and used in the ATHLOS project. Machine learning techniques tested on ATHLOS data will be transferred and applied in INFORE. INFORE will profit from the methods of gathering, filtering and harmonizing Big Data from distributed sources, which have been tested and applied in the ATHLOS project.

WIISEL – Wireless Insole for Independent and Safe Elderly Living.
WIISEL[19] was an FP7 project developed an open platform network architecture that delivers an unobtrusive, self-learning and wearable prevention and warning system to decrease the incidence of falls in the elderly population. The self-learning system will facilitate the prevention of falls by a gait and activity pattern recognition. SPRING was involved in this project.

INFORE will profit from the system architecture of the WIISEL project, which is designed to handle and analyse Big Data streams. From a huge amount of daily, weekly and monthly data, WIISEL strives to find relevant patterns, which can be used as a basis for predictions and suggestions – just as INFORE does in other fields of applications.

International Cancer Genome Consortium (ICGC), The Cancer Genome Atlas (TCGA) and the PanCancer Project.
ICGC and TCGA are two international consortia dedicated to generating catalogues of genomic abnormalities in tumours of a great variety of cancer types of great clinical and social impact. BSC has played a crucial role within ICGC-CLL, the Spanish project dedicated to the genomic study of Chronic Lymphocytic Leukaemia. Currently the

[17] https://www.argomarine.eu/
[18] http://athlosproject.eu/
[19] http://www.wiisel.eu/

group is actively involved in the PanCancer project[20], which is going beyond the perspective of individual cancers taking advantage of data produced by ICGC and TCGA to develop a holistic view of cancer genomics.

BSC's extensive expertise and in the life sciences domain and its involvement in international organizations for the study of cancer and tumour growth will play a crucial role in the successful outcome of the life sciences use-case in INFORE. Moreover, INFORE will benefit from Rbbt, an integrated Ruby-based bioinformatics toolkit for handling data analytics tasks in the biological domain, developed by BSC in the context of the Pan-Cancer project.

BLUEPRINT - A BLUEPRINT of Haematopoietic Epigenomes.
BLUEPRINT[21] is one of the two first so-called high impact research initiatives funded by the EU. The BLUEPRINT consortium was formed with the aim to further the understanding of how genes are activated or repressed in both healthy and diseased human cells. BSC has developed the Epigenomics comparative cyber-infrastructure (EPICO) and the BLUEPRINT Data Analysis Portal (BDAP), working towards a broadly accessible integrative analysis of epigenomic data across international consortia.

INFORE will benefit from new algorithms developed as part of this project, in order to perform analysis and integration of different kinds of epigenomic data, including definition of reference epigenomes chromatin states, identification of important chromatin regions determining cell type and integration of epigenomic data in 3D chromatin contact networks.

BioExcel: Centre of Excellence for Biomolecular Research (Active).
BioExcel[22] proposes to give significant support to: i) Make e-Infrastructure useable by researchers who are not computing experts, ii) Improve the performance and applicability of key life science applications, iii) Handle large amounts of data in computational workflows. BSC provides HPC support as well as special effort in handle large amounts of data in computational workflows.

All the objectives of BioExcel (enabling nonexperts in computing to be involved in largescale analytics of biological data, improve performance and applicability of research outcomes in the life sciences domain etc.) are highly relevant to INFORE's objectives. INFORE will thus be benefited from the expertise gained in BioExcel.

ASSET: Analysing and Striking the Sensitivities of Embryonal Tumours.
The major goal of ASSET[23] was to identify mechanistically understood network vulnerabilities that can be exploited for new approaches for diagnosis and treatment of highly aggressive and devastating major paediatric tumours, including neuroblastoma, medulloblastoma and Ewing sarcoma family tumours. BSC participated as leader of the integrative analysis Work Package, proposing pathway based Boolean models of the tumour cell lines studied within the project, informed with expression data to make them cell-line specific.

ASSET is highly relevant to INFORE's objectives. The methodologies developed in ASSET for Boolean models of tumour cell lines will be the starting point in the approach that will be followed in INFORE towards the study of the effects of drug synergies in tumour growth. The outcomes of this project are highly relevant to the objectives of the life sciences use-case in INFORE.

ERC H2020. Integration and analysis of heterogeneous big data for precision medicine and suggested treatments for different type of patients.
CRG worked on collecting and integrating data for precise medicine. The outcomes of this project are highly relevant to the objectives of the life sciences use-case in INFORE.

ELIXIR-EXCELERATE: Fast-track ELIXIR implementation and early user exploitation across the life sciences.
ELIXIR[24] coordinates and develops life science resources across Europe so that researchers can more easily find, analyse and share data, exchange expertise, and implement best practices. ELIXIR-EXCELERATE is an EC funded

[20] http://pancancer.info/
[21] http://www.blueprint-epigenome.eu/
[22] http://bioexcel.eu/
[23] http://www.ucd.ie/sbi/asset/
[24] https://www.elixir-europe.org/excelerate

project to help ELIXIR coordinate and extend national and international data resources to ensure the delivery of life-science data services. The Spanish National Bioinformatics Institute, led by BSC, is the Spanish representative in this project. INFORE will benefit from work in ELIXIR-EXCELERATE on the alignment of initiatives, dedicated to information extraction and analysis with the ELIXIR platform, and the project's focus on how researchers can access, store, transfer and analyse large amounts of life science data.

LIFT[25]**:** Local Inference in Distributed Systems
LIFT explored a novel approach for realising sophisticated, large-scale distributed data-stream analysis systems, relying on processing local data in situ. ATHENA participated in this project which demonstrated how to employ novel geometric techniques for intelligently decomposing the monitoring of complex holistic conditions and functions into safe, local constraints that can be tracked independently at each node (without communication), while guaranteeing correctness for the global-monitoring operation. INFORE will exploit the studies conducted to perform distributed Function Monitoring over Massive Data Streams

iPC[26]**:** Paediatric cancer, genomics
This project focuses on identifying effective personalized medicine for paediatric cancers and addresses a multitude of challenges. To meet these challenges, a comprehensive computational effort to combine knowledge base, machine-learning, and mechanistic models to predict optimal standard and experimental therapies for each child is proposed. High-quality multi-disciplinary data are produced, assembled, standardised and harmonised and the potential of Big Data and HPC for the personalized treatments of European citizens is being leveraged. iPC objectives are closely related to INFORE's scope and the project will benefit from BSC's work in iPC.

3.6.4 Societies and Organizations

3.6.4.1 Big Data Value Association

The Big Data Value Association (BDVA) is the private counterpart to the EU Commission to implement the Big Data Value Public Private Partnership (BDV PPP programme). BDVA has over 150 members all over Europe with a well-balanced composition of large and small and medium-sized industries as well as research organisations. INFORE partners will contribute to the BDV PPP activities and participate in relevant events, consultations and matchmaking events organised in the context of the BDV PPP. Since the collection of up-to-date sectorial data is of crucial importance in view of measuring the success of the PPP, INFORE will provide input in surveys and questionnaires about the BDV PPP. The project will also participate in events organised under the umbrella of BDV PPP and boost synergies with other projects through co-organisation of workshops, demos, presentations or lectures.

Figure 6: INFORE presented at BDVA event by the dissemination manager, Konstantinos Chatzikokolakis

[25] https://cordis.europa.eu/project/rcn/95556/factsheet/en
[26] https://ipc-project.eu/

4 Monitoring and Evaluation of Dissemination Activities

The dissemination activities of each partner are closely monitored by the dissemination manager, as they should be reflected on the main dissemination channels of the project (e.g., website, social media platforms). In the context of the regular teleconferences in which all partners participate to report their progress, progress in dissemination activities is also discussed. There is also a dedicated page in the project's internal workspace where partners report project-related publications that have been accepted. The dissemination activities that will be carried out in the context of the project will be included in the dissemination reports (i.e., deliverable D7.3) that will be submitted on months 18 and 36 respectively.

4.1 Dissemination KPIs

Contractually, for every mean of dissemination, there are specific KPIs defined. In the context of monitoring the progress made by the consortium with respect to the dissemination activities, we maintain a table that shows the number of dissemination activities performed in the reported period with respect to the targeted value per category of dissemination activities. The current snapshot of the progress performed with respect to the KPIs is depicted in Table 7. Notably, we already have 5 research papers published, while the target value for the end of the project is 20. On the other hand, we need to increase our social media following.

Table 7: Dissemination KPIs and status (up to 24/06/2019)

Mean of dissemination	KPI description	Targeted value	Actual value
Online dissemination and use of Web 2.0	LinkedIn subscribers	~200/year	17
	Facebook/Twitter followers	~200/year	52
	YouTube video uploads	~2/year	-
	YouTube channel subscribers	~100/year	6
	ResearchGate followers	100/year	22[27]
	White Papers	2/year	-
Event-based Dissemination	Organised industrial workshops	2	-
	Organised scientific workshops	1	-
	Keynote/Invited/Tutorial talks	5	-
Interactive dissemination	Demo papers in top-tier conferences such as SIGMOD, VLDB, etc.	2	1
Research/Academic-oriented dissemination	Publications in journals	~6-8	2[28]
	Publications in conferences	~20	5
	PhD/Master Theses	5	1 (NCSR)
Commercially-oriented dissemination	Participation in fairs	2	-
	Exhibition at events	2	-

[27] This number includes collaborators and followers
[28] 1 of them is in top-tier journal (i.e., VLDB journal)

5 Conclusions

In the context of this document, we described in detail the dissemination and communication plan of INFORE. We identified the different target groups towards which the dissemination activities will be addressed, the main dissemination means and how they will be used to communicate the project outcomes to the targeted audience, ensuring maximum engagement.

At the same time, we explained how the dissemination and communication activities are monitored, as well as the key members of the consortium that are responsible for monitoring the dissemination activities of the project and ensuring that the project is in line with the objectives and KPIs that have been defined in the grant agreement.

In this context, we also mention the dissemination activities that have already been carried out by the consortium, as the dissemination activities started from an early stage in the project. The first report on dissemination and communication activities performed in the first half of the project, in alignment with the dissemination plan presented in this document, is due M18.

Appendix A – Candidate logos

In this section we list the rest of the candidate logos that MarineTraffic and SPRING designed, besides the one selected from the INFORE consortium.


INFORE
Interactive Extreme-Scale
Analytics and Forecasting
INFORE
Interactive Extreme-Scale
Analytics and Forecasting