Ethics Management Plan
Work Package 8 Task 8.1 Deliverable 8.1

Authors

Nikos Giatrakos, Antonios Deligiannakis
Athena Research & Innovation Center

Distribution list:

Groups:	Others:
WP Leader: Athena Task Leader: Athena	Internal Reviewer Partner: Barcelona Supercomputing Center (BSC) INFORE Management Board INFORE Project Officer

Document history:

Revision	Date	Section	Page	Modification
0.1	25/02/2019	1,2, Appendices A, B, C	1-5	Creation
0.2	26/02/2019	3,4	6-14, (5,7,8)	Creation, (Figure addition)
0.3	27/02/2019	1,2,3,4	All, (14-16)	Figure addition, Small edits, (Creation)
0.4	28/02/2019	All	All	Self-review
0.5	01/03/2019	-	-	Submitted for Internal Review
0.6	12/03/2019	All	All	Internal Review Comments' Incorporation
0.7	13/03/2019	3	10	DPO Info and contact details added
1.0	26/03/2019	-	-	Final Version

Approvals:

First Author: Nikos Giatrakos (Athena) Date: 28/02/2019

Internal Reviewer: Arnau Montagud (BSC) Date: 12/03/2019

Coordinator: Antonios Deligiannakis (Athena) on behalf of the Management Board Date: 26/03/2019

Table of contents:

1 Executive Summary

This deliverable elaborates on ethical issues in INFORE and explains the way they are addressed. It details the procedures and criteria that will be used to identify/recruit expert users throughout the various stages of the project's workplan. More specifically, information on how the consortium will reach these participants, and detailed information on the informed consent procedures that will be implemented are included. Moreover, it provides templates of project related information.

This deliverable validates that INFORE remains in accordance with the ethics screening procedures in H2020, as summarized in Figure 1, also discussed in Section 3.8. In particular, the ethics self-assessment that preceded the submission of the INFORE proposal was followed by an ethics review during the evaluation and grant preparation for the proposal. The current deliverable performs necessary checks for ethics appraisal throughout the lifespan of INFORE, as marked on Figure 1.

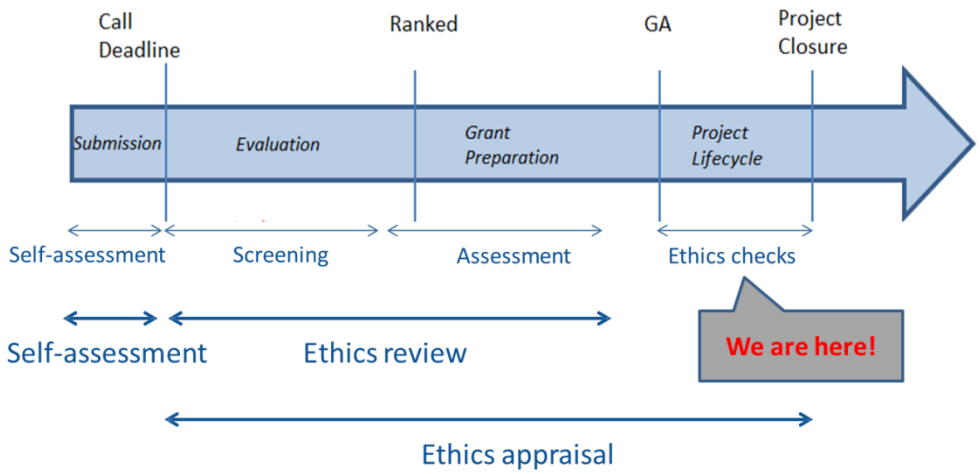


Figure 1: Where this document stands in H2020 Ethics Appraisal Procedure. (Adapted from https://ec.europa.eu/info/sites/info/files/14_rea_ethics_evaluation_afr.pdf) . GA stands for Grant Agreement.

2 Introduction

INFORE aims at integrating within a novel pluggable architecture (a) innovative real-time, interactive machine learning and data mining tools, (b) distributed complex event forecasting to support proactive decision making, (c) a framework for supporting non-programmer data analysts to specify processing workflows and data analytics tasks, (d) data summarization and approximate query processing techniques over Big Data platforms and/or High Performance Computing (HPC) and Cloud infrastructures.

In order to achieve its goals and evaluate the success of the developed technologies, INFORE employs both objective and subjective means. Objective means mainly involve IT benchmarking, which is destined to measure the success of INFORE's innovations from a technological perspective, with respect to quantitative Key Performance Indicators (KPIs). For instance, developing an event forecasting technology that provides alerts about maritime surveillance related events 15 minutes before their actual occurrence, compared to an existing situation that detects these events only upon they occur, is an objective mean of technological success. In the same spirit, cutting down the time required to set up a new workflow for data processing by an order of magnitude is another KPI for the technological advancements measured in the scope of the project.

On the other hand, subjective means involve qualitative measures and requires the engagement of expert users in the development and the evaluation of the developed technologies to make sure INFORE substantially contributes in meeting their needs and in solving everyday business problems the way they would expect. Therefore, expert user input and human factor evaluation is entailed in key project workplan phases.

2.1 How Humans Interact with INFORE

Use case partners, first identify candidate expert users to be interviewed for each use case. To aid this process and better determine candidate expert users, the Management Board provides the form included in Appendix B to be filled by partners, prior to approaching potential candidates. Then expert user engagement comes after receiving all the necessary information about the project, its purposes and the aim of the interview/questionnaire/survey and having provided their explicit consent on a voluntary basis as discussed in Section 4, in all three use cases and in three phases of the project's workplan (before the Finalization Phase which incorporates final expert users feedback in the developed technologies).

In particular:

- **Initialization Phase**: use case needs, and requirement analysis is performed for each of the Life Science, Financial and Maritime use cases studied within the scope of the project. Expert user input comes for knowledge elicitation and understanding of domain expertise. Interviewing expert users at the early stage of the project creates descriptions of their work system and captures both the programmable (automated) and human oriented (manual, semi-automated) parts of the workflow followed to accomplish everyday business tasks such as setting up new Big Data processing pipelines, processing the output of existing tools and reaching business decisions. The number of interviewed users is at least two per use case and interviews ideally take place on their working environment, a time of the day they declare their availability. During the interview, an INFORE researcher for each use case poses questions and derives answers based on a formulated questionnaire which is provided along with Deliverables D1.1, D2.1, D3.1.
- **Core Research and Architecture Development Phase**: The midterm report of the project presents INFORE's Initial Prototype. For each use case, at least three domain experts will be interviewed in order to assess the practical utility and usability of the developed prototype. However, additional expert users from various economic sectors may be recruited taking advantage of own network of connections of industrial partners of the project (such as RapidMiner partners). The aim of this process is to incorporate expert user feedback in advancements of INFORE's technology and apply course correction where appropriate and if needed. Expert users will be given access to the prototype for a sufficient amount of time and they will be surveyed right after. All relevant questionnaires will be included in evaluation reports with Deliverables D1.3, D2.2, D3.2.
- **Advanced Research and Architecture Development Phase**: Towards the final report of the project INFORE's Final Prototype will be presented. For receiving expert user feedback, if possible, the prototype will be integrated into their work practice (e.g., running in parallel to their current platforms). Evaluation will

involve observation and interview with expert users in their workplace (or mimic equivalent wherever this is not possible). This step will involve in total at least 20 experts, with a minimum of 3 per use case. Again, additional expert users from various economic sectors may be recruited taking advantage of own network of connections of industrial partners of the project or volunteers being informed via web announcements and press releases. All relevant questionnaires will be included in evaluation reports with Deliverables D1.5, D2.3, D3.3.

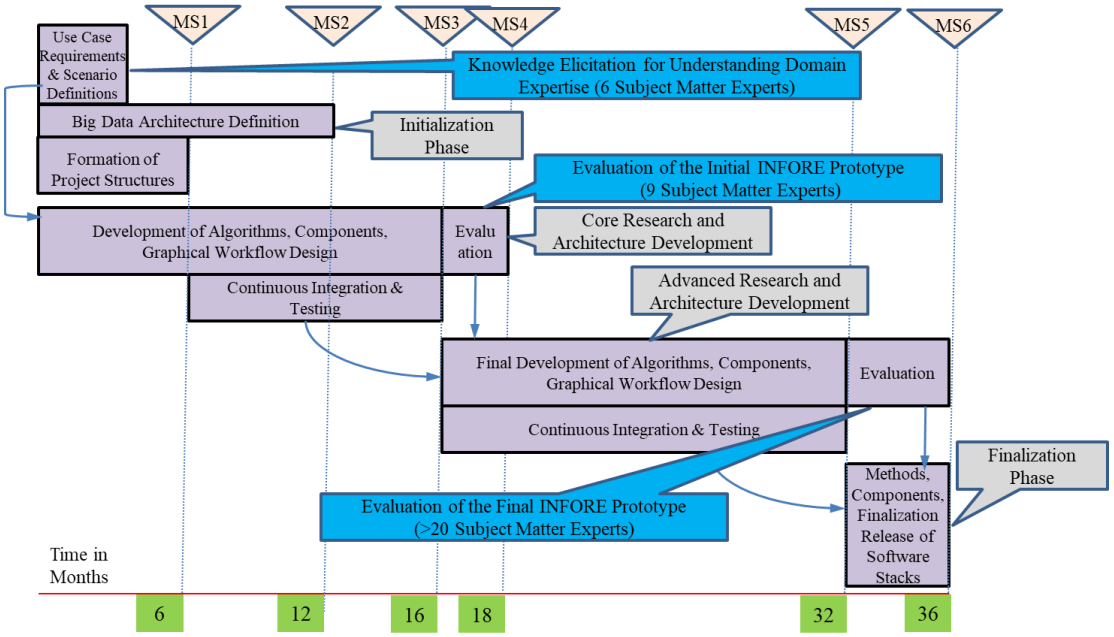


Figure 2: How humans interact with INFORE. Light blue boxes refer to expert user engagement in the various phases of INFORE's Workplan.

Figure 2 provides an illustration of the above discussion combining INFORE's workplan along with expert user participation throughout its lifespan.

2.2 No Ethical Issues in INFORE's Use Cases

Apart from the expert users mentioned in the previous section, humans are not engaged as study and/or data collection subjects in any of the INFORE use cases. Furthermore, INFORE's use cases do not involve sensitive data that may require additional individual consent.

2.2.1 Life Science Use Case

The purpose of this use-case is to provide a "virtual laboratory" for studying tumor growth and evolution. In particular, the goal is to use **in-silico** models of cell systems found in in-vivo tumors, aiming to facilitate the design, test and optimization of cancer treatments based on combinations of different drugs. As the term "in-silico'' suggests per se, biological processes of in-vivo tumors are simulated in a controlled environment. Prior biological knowledge or experimental data from heterogeneous sources are incorporated in the simulation models [1][2][3] and will stem from publicly available and published research results. The starting point will be to incorporate experimental evidence on drug synergies, collected from public available resources such as the Drug Combination DataBase (DCDB) [4] and published works. No new human cells are to be employed in INFORE's research whatsoever. INFORE performs IT research aiming at providing the necessary architectural and algorithmic apparatus for speeding up the repetitive procedures employed by biologists and data scientists upon modelling, setting up and running the simulations. Figure 3 presents an example of the basic structure of a data tuple holding cell information during in-silico simulations. Please refer to Deliverable D1.1 for the details of simulations and relevant data.

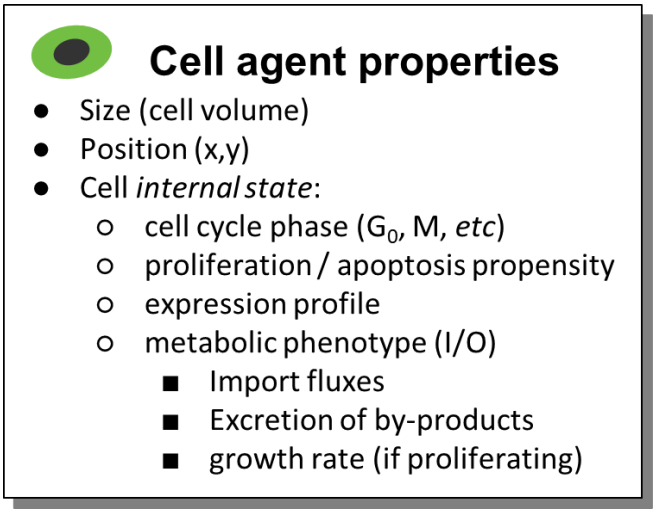


Figure 3: Exemplary data tuple structure during in-silico simulations in the Life Science use case.

2.2.2 Financial Use Case

This use-case will use massive amounts of historical and near real-time market data to allow for forecasting price swings of stocks, currencies, commodities and industrial goods, predict systemic risk, decision support for investment opportunities via the identification of relative strengths and weaknesses of assets and goods at a given time. All information within Level 1,2,3 stock data are shared according to procedures that abide by the rules created by stock exchanges and approved by securities regulators[1]. The project does not incorporate or study market makers' or brokers' behavioral data. Table 1 presents an example of the basic structure of a data tuple holding Level 1 data for a particular stock. Please refer to Deliverable D2.1 for the details of relevant stock data.

Date	Time	Volume	Quantity
01/08/2019	00:00:01	6.18177	1

Table 1: Exemplary data tuple for Level 1 stock data for a particular stock

2.2.3 Maritime Use Case

This use-case relies on existing maritime surveillance systems such as the Automatic Identification System (AIS) and Sentinel-1 Synthetic Aperture Radar (SAR) satellite imagery, with "local" views, obtained from autonomous maritime vessels, acting as on-site sensing devices. The goal is to extract in real-time valuable insights from the incoming data, forecasting critical events improving Maritime Surveillance. AIS data are provided by vessels on a voluntary basis and vessel location tracking devices cannot be remotely switched on by any authority. Because AIS transmissions take place on the basis of open frequencies a number of private companies have set up their own coastal monitoring stations to provide AIS information on a commercial basis over the internet usually on a subscription basis. In particular, MarineTraffic provides AIS data on a subscription basis[2]. SAR imagery and on-site devices provide data deprived of any vessel identification information. No personal information of individuals or groups of individuals are identified by SAR or on-site devices per se, but only vessel information at abstract level such as the vessel type, size etc. The interest is to classify single vessels of different size, their activities and possible interactions. Please refer to D3.1 for particulars of data collection and format.

[1] https://ec.europa.eu/info/business-economy-euro/banking-and-finance/financial-markets/securities-markets/investment-services-and-regulated-markets-markets-financial-instruments-directive-mifid_en

[2] https://www.marinetraffic.com/en/p/apply-for-free-ais-receiver

Figure 4: Radar silhouette of a ship, produced with the ISAR-Processor of the Ocean Master (http://www.radartutorial.eu/20.airborne/ab07.en.html)

3 Regulatory Frameworks and Initiatives

The INFORE Consortium is fully aware of ethical principles and data protection regulations at national, European and international level. In particular, INFORE has extensively reviewed and will assure its compliance with the regulatory frameworks in all interactions with humans during interviews and surveys as described in the previous section. Data protection and pseudonymization of expert users will also be accounted for in INFORE, in all versions of our Data Management Plan (Deliverables D8.3, D8.4 and D8.6).

Although no ethical or issues regarding personal data arise in the project use cases, the consortium is fully aware about and keeps screening relevant regulatory frameworks and legal aspects governing respective sectors per use case. Despite the absence of ethical issues, ATHENA (as the project coordinator) has appointed a Data Protection Officer (DPO), who is fully aware of the project, and the contact details (Name: Prodromos Tsiavos, email: dpo@athena-innovation.gr) of the DPO have been made available to all data subjects involved in the research for posing their questions and exercising their rights. Moreover, ATHENA (as the project coordinator) may use its Committee of Ethics of Athena RC (see https://www.athenarc.gr/el/ethics_committee – page currently available only in Greek) for consulting purposes regarding future ethical issues that may arise.

3.1 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016[3]

on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. This repeals the Data Protection Directive 95/46/EC mentioned in Section 5 – Part B of INFORE's Description of Action. The following guidelines are accounted for in ethics management procedures in INFORE.

(26) The principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments. The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.
(33) It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose.
(50) The processing of personal data for purposes other than those for which the personal data were initially collected should be allowed only where the processing is compatible with the purposes for which the personal data were initially collected. In such a case, no legal basis separate from that which allowed the collection of the personal data is required. If the processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller, Union or Member State law may determine and specify the tasks and purposes for which the further processing should be regarded as compatible and lawful. Further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be considered to be compatible lawful processing operations. The legal basis provided by Union or Member State law for the processing of personal data may also provide a legal basis for further

[3] http://data.europa.eu/eli/reg/2016/679/oj

processing. In order to ascertain whether a purpose of further processing is compatible with the purpose for which the personal data are initially collected, the controller, after having met all the requirements for the lawfulness of the original processing, should take into account, inter alia: any link between those purposes and the purposes of the intended further processing; the context in which the personal data have been collected, in particular the reasonable expectations of data subjects based on their relationship with the controller as to their further use; the nature of the personal data; the consequences of the intended further processing for data subjects; and the existence of appropriate safeguards in both the original and intended further processing operations.

Where the data subject has given consent or the processing is based on Union or Member State law which constitutes a necessary and proportionate measure in a democratic society to safeguard, in particular, important objectives of general public interest, the controller should be allowed to further process the personal data irrespective of the compatibility of the purposes. In any case, the application of the principles set out in this Regulation and in particular the information of the data subject on those other purposes and on his or her rights including the right to object, should be ensured. Indicating possible criminal acts or threats to public security by the controller and transmitting the relevant personal data in individual cases or in several cases relating to the same criminal act or threats to public security to a competent authority should be regarded as being in the legitimate interest pursued by the controller. However, such transmission in the legitimate interest of the controller or further processing of personal data should be prohibited if the processing is not compatible with a legal, professional or other binding obligation of secrecy.

(52) Derogating from the prohibition on processing special categories of personal data should also be allowed when provided for in Union or Member State law and subject to suitable safeguards, so as to protect personal data and other fundamental rights, where it is in the public interest to do so, in particular processing personal data in the field of employment law, social protection law including pensions and for health security, monitoring and alert purposes, the prevention or control of communicable diseases and other serious threats to health. Such a derogation may be made for health purposes, including public health and the management of health-care services, especially in order to ensure the quality and cost-effectiveness of the procedures used for settling claims for benefits and services in the health insurance system, or for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. A derogation should also allow the processing of such personal data where necessary for the establishment, exercise or defence of legal claims, whether in court proceedings or in an administrative or out-of-court procedure.

(53) Special categories of personal data which merit higher protection should be processed for health-related purposes only where necessary to achieve those purposes for the benefit of natural persons and society as a whole, in particular in the context of the management of health or social care services and systems, including processing by the management and central national health authorities of such data for the purpose of quality control, management information and the general national and local supervision of the health or social care system, and ensuring continuity of health or social care and cross-border healthcare or health security, monitoring and alert purposes, or for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, based on Union or Member State law which has to meet an objective of public interest, as well as for studies conducted in the public interest in the area of public health. Therefore, this Regulation should provide for harmonised conditions for the processing of special categories of personal data concerning health, in respect of specific needs, in particular where the processing of such data is carried out for certain health-related purposes by persons subject to a legal obligation of professional secrecy. Union or Member State law should provide for specific and suitable measures so as to protect the fundamental rights and the personal data of natural persons. Member States should be allowed to maintain or introduce further conditions, including limitations, with regard to the processing of genetic data, biometric data or data concerning health. However, this should not hamper the free flow of personal data within the Union when those conditions apply to cross-border processing of such data.

(62) However, it is not necessary to impose the obligation to provide information where the data subject already possesses the information, where the recording or disclosure of the personal data is expressly laid down by law or where the provision of information to the data subject proves to be impossible or would involve a disproportionate effort. The latter could in particular be the case where processing is carried out for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. In that regard, the number of data subjects, the age of the data and any appropriate safeguards adopted should be taken into consideration.

(65) A data subject should have the right to have personal data concerning him or her rectified and a 'right to be

forgotten' where the retention of such data infringes this Regulation or Union or Member State law to which the controller is subject. In particular, a data subject should have the right to have his or her personal data erased and no longer processed where the personal data are no longer necessary in relation to the purposes for which they are collected or otherwise processed, where a data subject has withdrawn his or her consent or objects to the processing of personal data concerning him or her, or where the processing of his or her personal data does not otherwise comply with this Regulation. That right is relevant in particular where the data subject has given his or her consent as a child and is not fully aware of the risks involved by the processing, and later wants to remove such personal data, especially on the internet. The data subject should be able to exercise that right notwithstanding the fact that he or she is no longer a child. However, the further retention of the personal data should be lawful where it is necessary, for exercising the right of freedom of expression and information, for compliance with a legal obligation, for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller, on the grounds of public interest in the area of public health, for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, or for the establishment, exercise or defence of legal claims.

(156) The processing of personal data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be subject to appropriate safeguards for the rights and freedoms of the data subject pursuant to this Regulation. Those safeguards should ensure that technical and organisational measures are in place in order to ensure, in particular, the principle of data minimisation. The further processing of personal data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes is to be carried out when the controller has assessed the feasibility to fulfil those purposes by processing data which do not permit or no longer permit the identification of data subjects, provided that appropriate safeguards exist (such as, for instance, pseudonymisation of the data). Member States should provide for appropriate safeguards for the processing of personal data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. Member States should be authorised to provide, under specific conditions and subject to appropriate safeguards for data subjects, specifications and derogations with regard to the information requirements and rights to rectification, to erasure, to be forgotten, to restriction of processing, to data portability, and to object when processing personal data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. The conditions and safeguards in question may entail specific procedures for data subjects to exercise those rights if this is appropriate in the light of the purposes sought by the specific processing along with technical and organisational measures aimed at minimising the processing of personal data in pursuance of the proportionality and necessity principles. The processing of personal data for scientific purposes should also comply with other relevant legislation such as on clinical trials.

(157) By coupling information from registries, researchers can obtain new knowledge of great value with regard to widespread medical conditions such as cardiovascular disease, cancer and depression. On the basis of registries, research results can be enhanced, as they draw on a larger population. Within social science, research on the basis of registries enables researchers to obtain essential knowledge about the long-term correlation of a number of social conditions such as unemployment and education with other life conditions. Research results obtained through registries provide solid, high-quality knowledge which can provide the basis for the formulation and implementation of knowledge-based policy, improve the quality of life for a number of people and improve the efficiency of social services. In order to facilitate scientific research, personal data can be processed for scientific research purposes, subject to appropriate conditions and safeguards set out in Union or Member State law.

(159) Where personal data are processed for scientific research purposes, this Regulation should also apply to that processing. For the purposes of this Regulation, the processing of personal data for scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research. In addition, it should take into account the Union's objective under Article 179(1) TFEU of achieving a European Research Area. Scientific research purposes should also include studies conducted in the public interest in the area of public health. To meet the specificities of processing personal data for scientific research purposes, specific conditions should apply in particular as regards the publication or otherwise disclosure of personal data in the context of scientific research purposes. If the result of scientific research in particular in the health context gives reason for further measures in the interest of the data subject, the general rules of this Regulation should apply in view of those measures.

(160) Where personal data are processed for historical research purposes, this Regulation should also apply to that processing. This should also include historical research and research for genealogical purposes, bearing in

mind that this Regulation should not apply to deceased persons.
(161) For the purpose of consenting to the participation in scientific research activities in clinical trials, the relevant provisions of Regulation (EU) No 536/2014 of the European Parliament and of the Council (15) should apply.
(162) Where personal data are processed for statistical purposes, this Regulation should apply to that processing. Union or Member State law should, within the limits of this Regulation, determine statistical content, control of access, specifications for the processing of personal data for statistical purposes and appropriate measures to safeguard the rights and freedoms of the data subject and for ensuring statistical confidentiality. Statistical purposes mean any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results. Those statistical results may further be used for different purposes, including a scientific research purpose. The statistical purpose implies that the result of processing for statistical purposes is not personal data, but aggregate data, and that this result or the personal data are not used in support of measures or decisions regarding any particular natural person.
Article 5 Principles relating to processing of personal data: *(b) collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation');* *(e) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject ('storage limitation');*
Article 9 Processing of special categories of personal data: *(j) processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.*
Article 14 Information to be provided where personal data have not been obtained from the data subject: *5. Paragraphs 1 to 4 shall not apply where and insofar as:* *(a) the data subject already has the information;* *(b) the provision of such information proves impossible or would involve a disproportionate effort, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, subject to the conditions and safeguards referred to in Article 89(1) or in so far as the obligation referred to in paragraph 1 of this Article is likely to render impossible or seriously impair the achievement of the objectives of that processing. In such cases the controller shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available;*
Article 17 Right to erasure ('right to be forgotten') *1. The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies:* *(a) the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed;* *(b) the data subject withdraws consent on which the processing is based according to point (a) of Article 6(1), or point (a) of Article 9(2), and where there is no other legal ground for the processing;* *(c) the data subject objects to the processing pursuant to Article 21(1) and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing pursuant to Article 21(2);* *(d) the personal data have been unlawfully processed;* *(e) the personal data have to be erased for compliance with a legal obligation in Union or Member State law to which the controller is subject;* *(f) the personal data have been collected in relation to the offer of information society services referred to in Article 8(1).*
Article 21 Right to object *1. The data subject shall have the right to object, on grounds relating to his or her particular situation, at any*

time to processing of personal data concerning him or her which is based on point (e) or (f) of Article 6(1), including profiling based on those provisions. The controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims.

6. Where personal data are processed for scientific or historical research purposes or statistical purposes pursuant to Article 89(1), the data subject, on grounds relating to his or her particular situation, shall have the right to object to processing of personal data concerning him or her, unless the processing is necessary for the performance of a task carried out for reasons of public interest.

Article 89 Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.

1. Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner.

2. Where personal data are processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.

3.2 Charter of Fundamental Rights of the European Union[4]

The Charter of Fundamental Rights dedicates a separate article to the protection of personal data. Article 8 of the Charter refers to the legitimate processing of personal data which must be fair and for specified purposes based on the consent of the data subject or other legitimate basis laid down by law. INFORE will further assure that no kind of discrimination as described below will be performed while choosing candidate expert users.

Article 8 Protection of personal data:

1. Everyone has the right to the protection of personal data concerning him or her.

2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.

3. Compliance with these rules shall be subject to control by an independent authority.

Article 21 Non-discrimination:

1. Any discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited.

2. Within the scope of application of the Treaties and without prejudice to any of their specific provisions, any discrimination on grounds of nationality shall be prohibited.

3.3 European Convention on Human Rights

Expert users will be treated with dignity and respect according to Article 8 of the European Convention on Human Rights.

[4] http://data.europa.eu/eli/treaty/char_2012/oj

Article 8 Right to respect for private and family life:

1. Everyone has the right to respect for his private and family life, his home and his correspondence.

2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

3.4 Declaration of Helsinki[5]

Although INFORE does not engage medical research on human subjects, the Consortium is aware of the generic principles stemming from the Declaration of Helsinki.

3.5 Legal Aspects on Financial Markets

- Regulation (EU) No 600/2014 of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Regulation (EU) No 648/2012 Text with EEA[6].

(23) Market data should be easily and readily available to users in a format as disaggregated as possible to allow investors, and data service providers serving their needs, to customise data solutions to the furthest possible degree. Therefore, pre-trade and post-trade transparency data should be made available to the public in an 'unbundled' fashion in order to reduce costs for market participants when purchasing data.

- Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU Text with EEA relevance[7].

(96) In order to enhance the conditions under which investment firms comply with their obligation to execute orders on terms most favourable to their clients in accordance with this Directive, it is appropriate to require that for financial instruments subject to the trading obligation in Articles 23 and 28 of Regulation (EU) No 600/2014 that each trading venue and systematic internaliser and for other financial instruments that each execution venue to make available to the public data relating to the quality of execution of transactions on each venue.

3.6 Legal Aspects of Maritime Monitoring & Surveillance Data[8]

INFORE remains aware of the legal aspects concerning maritime monitoring and surveillance including the Law of the Sea, reporting regimes, surveillance systems, legal data sharing mechanisms and restriction on data sharing. All such aspects are directly issued from MarineTraffic and CMRE INFORE partners' legal existence.

3.7 European Group on Ethics (EGE) in Science and New Technologies[9]

Beyond the recruitment of expert users are described in Section 2, INFORE Consortium ensures compliance with Opinions[10] and Statements[11] of EGE on the formulation of a code of conduct for research integrity for projects funded by the European Commission.

[5] https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/
[6] http://data.europa.eu/eli/reg/2014/600/oj
[7] http://data.europa.eu/eli/dir/2014/65/oj
[8] https://ec.europa.eu/maritimeaffairs/documentation/studies/study_monitoring_en
[9] https://ec.europa.eu/info/research-and-innovation/strategy/support-policy-making/scientific-support-eu-policies/european-group-ethics-science-and-new-technologies-ege_en#ege-opinions-and-statements
[10] https://ec.europa.eu/info/publications/ege-opinions_en
[11] https://ec.europa.eu/info/publications/ege-statements_en

3.8 H2020 Ethics Appraisal Procedures

The process to assess and address the ethical dimension of activities funded under Horizon 2020 is called Ethics Appraisal Procedure[12]. The phases of this procedure are included in Figure 1. INFORE has already completed the three first stages and is now in the implementation phase as marked on the figures.

[12] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting- issues/ethics_en.htm

4 Expert User Recruitment Protocol

For the purposes outlined in Section 2 and in compliance with the Regulatory Frameworks and Initiatives discussed in Section 3, expert users will be recruited using different means throughout the three phases of the project (see Figure 2) through own network of organizations and previously established contacts, during the first two phases of our workplan, as well as web announcements and press releases while deriving feedback for INFORE's final prototype. None of the studies involve any invasive procedures.

The procedure will proceed in steps summarized below and analyzed in the upcoming sections:

1) Expert User Identification: Each use case or industrial partner (depending on the project's phase) will make a list of potential expert users relevant to the technologies developed in INFORE.
2) Approach Procedure: Each use case or other partner (depending on the project's phase) will follow a generic recruitment protocol based on which expert users will be approached.
3) Declaration of Consent: Having acquired information about the project, the research and conducts and the purpose of the study, expert users will be asked to provide their consent by signing a corresponding form.
4) Interviews, Questionnaires, Surveys: Details in Deliverables D1.1, D1.3, D1.5, D2.1, D2.2, D2.3, D3.1, D3.2, D3.3, D8.5, D8.7.
5) Keeping Evidence of Accomplishment: Briefly summarized in Section 4.4 below and analyzed in Deliverables D8.3, D8.4, D8.6.

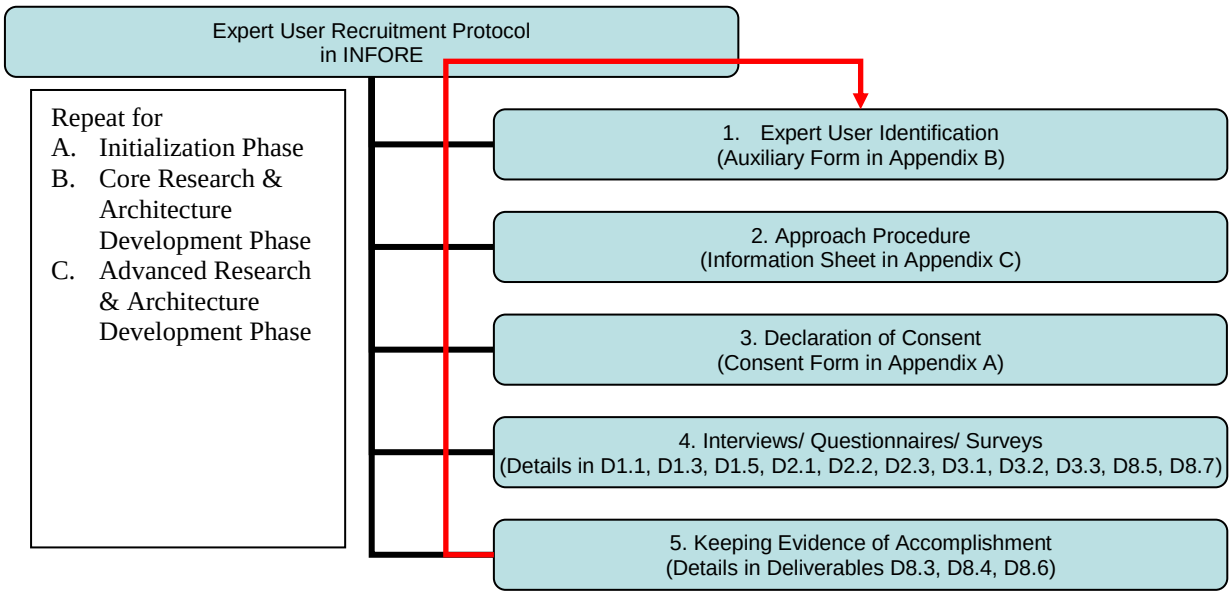


Figure 5: Ethics Management with Respect to Expert User Recruitement in INFORE.

4.1 Expert User Identification

In order to facilitate partners in identifying potential expert users to be engaged in INFORE from their own network of connections or to prioritize volunteers attracted via web announcements and press releases, the Management Board provides an auxiliary form to be filled with information about candidate expert users. The corresponding form is provided in this deliverable in Appendix B. The generic idea is that members of the consortium are to answer themselves these questions about each candidate expert user so that they can distinguish their level of expertise and relevance and potential interest in the results of the project. The form gathers information about:

- The position a candidate expert user possesses in an organization.
- The type of the organization they work for.
- The economic sector to which the organization puts in.

- The component of INFORE that may be of interest to the candidate expert user.
- The phase(s) of the workplan that the expert user could contribute the most.

Expert users are to be recruited so that the overall INFORE's approach can benefit from feedback coming from heterogeneous domain knowledge in different types of organizations and a variety of economic sectors.

4.2 Approach Procedure

Each potential expert user coming from a partner's own network will be contacted via e-mail or phone and asked if they would be interested in participating in the project. In case of expert users recruited via press releases and web announcements they are supposed to declare their interest by sending an email to INFORE partners included in the press release or web announcement, or get in touch via the contact details at INFORE's website.

Candidate participants will be provided an Information Sheet about the project which is included in Appendix C. If agreement is reached, the candidate participant will be asked to officially provide their consent participating in an interview or a survey.

4.3 Informed Consent

Expert users in interviews or surveys will be asked to sign a consent form, which will be provided prior to the interview. Participants will be provided both an Information Sheet (Appendix C) on the research conducted in INFORE as well as a Consent Form and will be given enough day(s) to respond to or ask questions about the form and the process itself prior to participating in the research. If the participants consent to the terms indicated in the Consent Form, they can proceed to participate.

Consent forms will be held securely by the INFORE researcher(s) taking part in the process. A Sample Consent Form is included in Appendix A.

Expert users will again receive written and/or oral information on:

- Overall purpose of the project
- Purpose of the concrete data collection
- Data retention period
- Withdrawal options
- Contact details of the researcher collecting the data
- Contact details of project coordinator which is also the data controller organization
- Contact details of the Data Protection Officer (DPO)
- Procedure of the study
- Potential risks and benefits
- Inclusion/exclusion criteria (Auxiliary Form Appendix B)
- Reimbursements to participants (No reimbursement)
- Planned use of the data

and they will be able to pose as many questions they need to feel completely aware about the details of the questionnaire, interview or survey.

Notably our Consent Form takes into account all relevant regulatory framework especially GDPR principles (Section 3.1) such as "the right to be forgotten" (Paragraph (65) and Article 17 of Section 3.1) when expert users that have already participated in all or some of use case elicitation or human factor evaluation phases change their mind with respect to the data they provided during interviews, surveys and respective questionnaires.

4.4 Keeping Evidence of Accomplishment

All data collected during interviews, through questionnaires or surveys will initially be kept securely by the INFORE researcher who collected the data. It is the responsibility of the INFORE partner, affiliated with the corresponding researcher to ensure abidance by the regulatory framework of Section 3 including paragraphs and

articles related to pseudonymization, personal data protection, right to object and right to be forgotten. This stands so long as the data remain with them, i.e., before being received by the data controller organization (Athena) via secure communication channels or in person during INFORE's meetings Such procedures indicatively involve:

- Each participant's data (Consent Form, questionnaires, recorded interview or responses to surveys) will be assigned a code to identify his/her respective data. To enable participants to withdraw their data at any state of the project, even upon their data have been de-identified, the participant should simply mention that code.
- The key-file containing identity information will be kept separate from the de-identified, pseudonymized parts of the data. The data containing information about the participants' identity will be stored in a secure file to which only the data controller will have access. The encryption will use the AES 256 algorithm or equivalent. Any data acquisition or communication will be performed using asymmetric algorithms using keys of at least 2048 bits.
- Only de-identified, anonymized data and aggregative statistic reports extracted from the database can be used in publications, project presentations and websites for academic and educational purposes.

We will also ensure that all partners will confirm that they have received these ethics requirements and that they will adhere to the expert user recruitment protocol. We may store partners' confirmation in our database for future reference. Throughout the project reminder emails on the key ethic requirements will be periodically send to all partners. We will use both teleconferences and Management Board discussions to remind all partners of the need to adhere to the ethics requirements and the protocol formulated in this deliverable. Athena, as coordinator, may also use face-to-face meetings with other consortium members as an opportunity to perform 'light touch' audits on the partner's procedures related to ethics.

5 Summary and Conclusive Remarks

This deliverable confirms that all partners involved in INFORE are aware of relevant regulatory frameworks and properly account for ethical issues in the scope of the project. It provides an overview of ethical issues that arise due to INFORE's approach in engaging expert users during use case elicitation and in performing human factor evaluation analysis on its prototypes. It analyzes that no ethical issues arise by the nature of the project use cases. It formally presents all the relevant regulatory framework screened by the consortium and finally described in detail the protocol followed for selecting and recruiting expert users.

The development of the latter protocol will aid in following a simple, uniform and easy to follow procedure to reduce the risk of misunderstandings within the consortium while recruiting expert users. Furthermore, technical aspects of data security have been addressed which results in mandatory rules for safekeeping of sensitive or personal data of expert users. The results of this deliverable are the ethical guidelines regarding INFORE and procedures in order to ensure the safety of the data included in the project. The first version of the INFORE's Data Management Plan is provided on Month 6.

6 References

[1] L. Calzone et al., “Mathematical Modelling of Cell-Fate Decision in Response to Death Receptor Engagement,” PLOS Comput. Biol., vol. 6, no. 3, p. e1000702, Mar. 2010.

[2] P. Bloomingdale, V. A. Nguyen, J. Niu, and D. E. Mager, “Boolean network modeling in systems pharmacology,” J. Pharmacokinet. Pharmacodyn., vol. 45, no. 1, pp. 159–180, Feb. 2018.

[3] A. Ghaffarizadeh, R. Heiland, S. H. Friedman, S. M. Mumenthaler, and P. Macklin, “PhysiCell: An open source physics-based cell simulator for 3-D multicellular systems,” PLOS Comput. Biol., vol. 14, no. 2, p. e1005991, Feb. 2018.

[4] Y. Liu, Q. Wei, G. Yu, W. Gai, Y. Li, and X. Chen, “DCDB 2.0: a major update of the drug combination database,” Database J. Biol. Databases Curation, vol. 2014, Dec. 2014.

Appendix

Appendix A: Sample Consent Form

Project name	Interactive Extreme Scale Analytics and Forecasting
Project acronym	INFORE
Lead institution	ATHENA RESEARCH AND INNOVATION CENTER
Funder	European Union Horizon 2020
Duration	02/01/2019 to 31/12/2021 (data retention period)
Website	www.infore-project.eu
Grant No.	825070
Name of the researcher for this task	
Contact details to responsible researcher	
Controller of personal data	Athena-Research and Innovation Center in Information, Communication and Knowledge Technologies - Athena RC (Research Organisation private law legal body) Artemidos 6 and Epidavrou 15125, Maroussi, Athens tel: + 302106875300, fax: 2106854270
Contact details of Data Protection Officer (DPO) of ATHENA RC	Name: Prodromos Tsiavos Email: dpo@athena-innovation.gr

1. Introduction

You have been invited to take part in an IT requirement elicitation questionnaire. Before deciding on whether you want to participate or not, please read this document carefully. Make sure you ask all the questions you may have so you can be completely confident you understand all the details of the study.

2. Purpose of the project

INFORE aims at integrating in a novel pluggable architecture innovative real-time, interactive machine learning and data mining tools, distributed complex event forecasting to support proactive decision making, a framework for supporting non-programmer data analysts to specify processing workflows and data analytics tasks, data summarization and approximate query processing techniques over High Performance Computing and Cloud infrastructures.

3. Risks or inconveniences

No risk is foreseen. You are only requested to be available to participate and answer questions in a simple questionnaire. The questionnaire relates to the IT challenges you face upon carrying out the data analytics tasks required in your everyday working life.

4. Benefits

There are no personal benefits besides the fact that the outcome of INFORE will provide readily available solutions for problems that are ubiquitous in today's analytics tasks. In that participants and their colleagues can then use the open source solutions developed by INFORE and evaluate that the issues they mention in the questionnaire have been addressed to the major extent.

5. Privacy and confidentiality

Controller of your personal data is ATHENA RC (see above). Responses you give in the questionnaires and interviews may be recorded. The questionnaires and the interviews will be stored by the Controller in its premises until 31/12/2021; then they will be deleted. You have the right to request from the controller access to and rectification or erasure of your personal data or restriction of processing or to object to processing as well as the right to data portability. If the controller cannot satisfy your demand, you can always lodge a complaint with the supervisory authority (http://www.dpa.gr). Your recorded data will be de-identified for further processing by the parties of the consortium; hence it will not be possible to identify your identity afterwards. Information will be processed and aggregative results may be presented in project reports. It will not be possible to identify the individual source of the information. The aggregative results of this investigation may be published in scientific journals or conferences and may be used in further studies. No personal data will be handed out to third parties and be used for other reason than the scope of this project.

6. Withdrawal options

You can request your information be deleted at any time, without giving a reason. If you should decide to deny your consent, please contact the leading investigator and let her/him know of your intention of leaving the study.

7. Contact persons

In case of any issue involving you in your role of participant of this study, please inform the Project Coordinator, Prof. Antonios Deligiannakis (adeli@imis.athena-innovation.gr).

8. Confirmation

Your participation in this study is only possible if you freely and independently sign this consent to authorize the INFORE consortium to use the data you provide.

Consent *Below you give your consent to voluntarily participate in the INFORE H2020 EU project. Read the information in this form carefully and consent with your signature below.*
• I have read and understood the information sheet, or it has been read to me. I have been given the opportunity and enough time to ask questions about the study and my questions have been fully answered.

• I agree voluntarily to be a participant in this study and understand that I can decide to withdraw from the study at any time.
• I understand that taking part in the study may involve audio-recorded interviews and I am willing to devote enough time to respond to the questionnaire.
• I give permission for the data that I provide to be de-identified and deposited in a repository so it can be used for INFORE's purposes as outlined above.
• I understand that personal information collected about me, will not be shared and only de-identified or aggregative information will be used for INFORE purposes.
• I understand that de-identified or aggregative information I provide may be used for reports and research outputs, possibly made available to websites and social media.
Signatures
I hereby declare that I have accurately read and fully understood the information given in this form, and I give my consent to participate in the INFORE H2020 EU project. _____ _____ _____ Name of participant Signature Date

Appendix B: Auxiliary Form for Expert User Identification

AA	Name	Organization	Position in Organization	Organization Type				Economic sector								INFORE's Phase. Request Expert User for			INFORE's Component. Expert User Interested in:							Provides Consent (Yes/No)	Expert User Contact Info	Responsible INFORE Researcher
				Government (national / EU / Other)	Research/ Academic	Business	Other (write)	Commerce	Transport	Financial / Trader	Education	Banking	Telecommunications	Manufacturing	Other	Initialization Phase	Core Research & Architecture Development Phase	Advanced Research & Architecture Development Phase	Data Summarization / Stream Processing	Distributed Learning & Data Mining	Distributed Complex Event Forecasting	Optimization & Runtime Adaptation	Workflow Design Facilities	Other Use Case Specific			
1																											
2																											
3																											
4																											

Appendix C: Sample Information Sheet

(to be substituted by project flyers)

INFORE
Interactive Extreme-Scale
Analytics and Forecasting

Goal

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 an 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.

Consortium

ATHENA IMIS (Coordinator) Greece	**NCSR Demokritos** Greece	**RapidMiner** Germany	**Barcelona Supercomputing Center** Spain
Center for Genomic Regulation Spain	**Spring Techno** Germany	**MarineTraffic** Greece	**NATO STO CMRE** Belgium

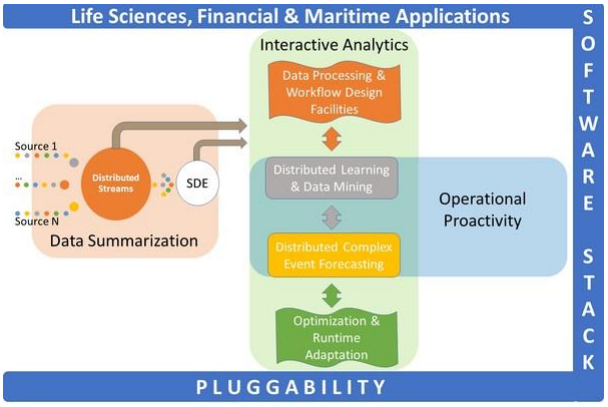
Why invest on INFORE?

The era of Big Data is upon us. A deluge of business data flows into corporate data centers at any given time, faster than anyone can go through it. This is the case in all (life science, financial, maritime) use cases served by INFORE. A common vision and challenge in all such use case scenarios is to couple the proliferation of a vast amount of distributed streaming data sources with the deployment of extreme-scale computing platforms for gaining unprecedented insights and knowledge about interesting phenomena deeply hidden in the data oceans. However, simply devoting HPC resources to incoming data is not good enough. Without clever algorithms capable of exploiting novel computing paradigms, such as HPC and the Cloud, for knowledge extraction and detection of interesting phenomena, we are essentially searching for a needle in a haystack. Challenges related to a seamless integration of streaming data management, predictive (learning) analytics, prescriptive (forecasting) analytics and their convergence with High-performance and Cloud computing, need to be overcome for realizing that vision.

INFORE aims at pioneering in confronting the aforementioned challenges. Imagine an architectural framework and analytics platform that can be easily installed as a whole or in parts, work in parallel with other adopted solutions if needed and support the whole processing pipeline from: a) digesting rapid distributed data streams from various sources, b) identifying the most important parts of the data constructing concise summaries, c) allowing for fast development, parameterization and testing of machine learning models for knowledge extraction at the same time when already developed learning models extract knowledge both out of the whole set of streams or their summaries, d) enabling automatic deployment of learned business rules, and e) enabling proactive acquisition of valuable knowledge about imminent events of interest. Further imagine that all the above are assisted by integrated graphical tools that enable the design of analytics workflows with minimal programming effort. Finally, add an optimization module that bears the characteristics of the available HPC or other resources, automatically devising the optimal plan for executing the whole data processing pipeline and, when needed, generates a new plan.

In that, INFORE manages to materialize goals no existing academic platform or commercial competitor has reached. It diminishes the time to set up online processing pipelines, it enables the dynamic online discovery and monitoring of business events and reduces time for searching, communicating and processing new business ideas. It thus cuts time to market and improves quality, with all the above formulating a set of breakthroughs of considerable magnitude.

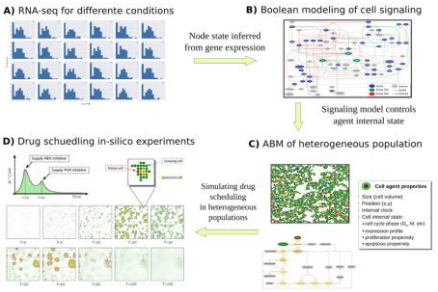
Overview



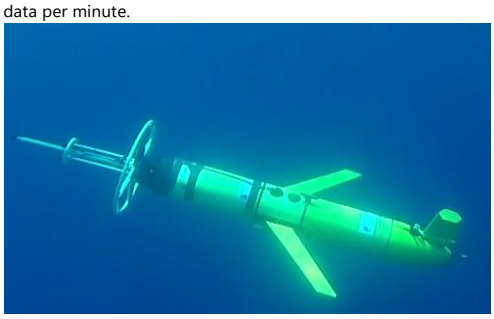
The above figure illustrates our envisioned approach. First, extreme-scale data streams are continuously acquired from various sources. Second, a dedicated, lightweight engine relying on approximate query processing techniques **extracts compact synopses of the data**, to allow for interactive response times while exploring the massive, high-speed input (in the figure this corresponds to SDE, standing for "Synopses Data Engine"). Third, the extracted data synopses are forwarded to INFORE's toolboxes, responsible for **interactive analytics** and **operational proactivity**. INFORE's toolboxes consists of four main components for: **(a) distributed machine learning & data mining; (b) data processing and workflow specification; (c) distributed complex event forecasting and (d) optimization and runtime adaptation of big data processing workflows**. In many of their tasks, these toolboxes try to maximize their performance through the use of these synopses in their operation.

INFORE Use Cases and Data

The approach of INFORE is applicable to a wide range of domains where stakeholders need to extract rapid insights from a multitude of correlated, extreme-scale data streams. During the project, the INFORE technology will be tested in three such domains, challenging our full set of objectives:



The life sciences use case, paving the way for the development of new cancer treatments, by identifying personalized 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 use 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.

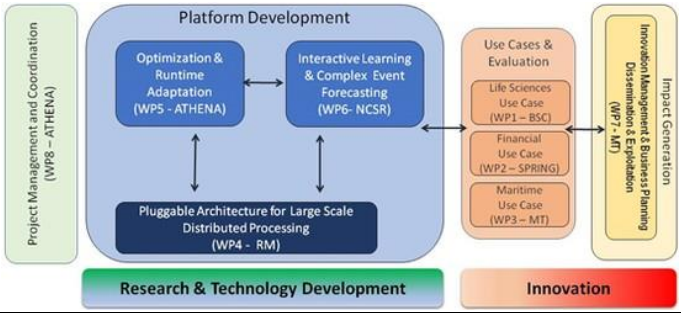


The maritime use case, improving maritime situational awareness, i.e. the ability to perceive and forecast activities and threats in maritime environments. For a complete picture of maritime activities, we will combine global maritime surveillance systems, such as the AIS (Automatic Identification System), 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.

Workplan

Strategy of the Workplan

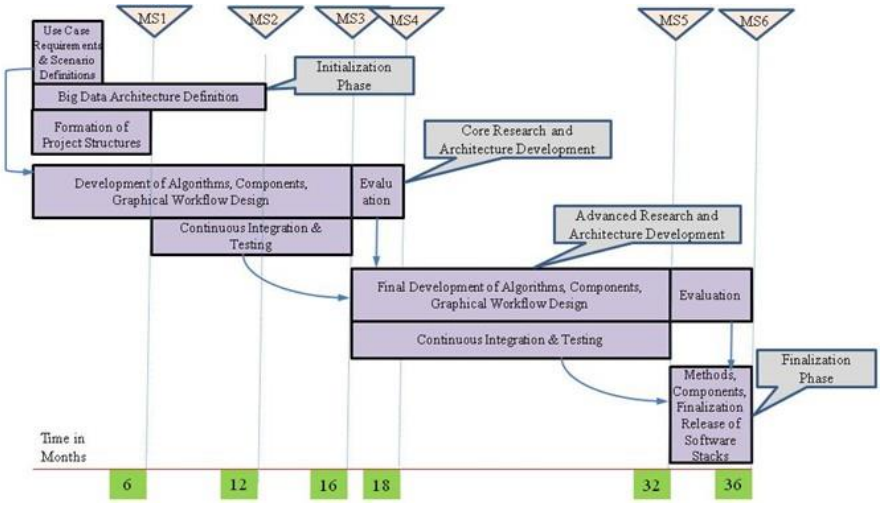


The rationale of the work plan structure is summarized as follows and depicted above. Work in INFORE is organized into 8 work packages: "Life Sciences Use Case" (WP1), "Financial Use Case" (WP2), "Maritime Use Case" (WP3), "Pluggable Architecture for Extreme-Scale Analytics" (WP4), "Optimization and Runtime Adaptation" (WP5), "Interactive Learning and Complex Event Forecasting" (WP6), "Innovation Management, Dissemination, Exploitation & Business Planning" (WP7) and "Project Management and Coordination" (WP8).

Timing of Work Packages

In more detail, the figure below depicts the timeline of the research and development phases in INFORE, the integration of the results from the individual work packages and the overall project milestones. The work in INFORE will be carried out in four main phases:

- The initialization phase (associated with Milestones MS1, MS2),
- Two iteration phases for the research and architecture development: the core (associated with Milestones MS3, MS4) and the advanced (associated with Milestones MS5, MS6) phases, and
- The finalization phase (also closed by MS6).



The Initialization Phase contains the **Initial Use Case Requirements and Formation of Project Structures** (MS1) Milestone, along with the **Architecture Design** (MS2) Milestone. Work for the Milestone MSI (month 6) includes the initial documentation of the three use cases and their requirements, along with the setup of an initial data management and dissemination plan. Also, an Ethics Management Plan is developed. At the same time the research in WP5-6 and the design of the system architecture in WP4 starts. Milestone MS2 is characterized by the design of the system architecture.

The **Core Research and Architecture Development** Phase will lead to the evaluation of the first version of the integrated INFORE infrastructure including all components scheduled for the first release – **Final Use Case Requirements, First Version of Scenarios and User Evaluation** (MS4), in month 18. Within this phase, Milestone MS3 (in month 16) **Initial Theoretical Framework and First Architecture Prototype**) defines, as an important intermediate goal, the completion of the first INFORE architecture prototype, to be used in the initial evaluation of the use cases. Although each work package will deliver intermediary stable versions to a common repository for continuous integration and automated regression testing based on testing data sets, months 17 to 18 will be dedicated to stabilize the first integrated version of the INFORE infrastructure. This will include integration and performance tests based on the use case data sets. The first version of the INFORE architecture will then be the basis for the first use case evaluation **using expert users**. The evaluation results will be documented and fed back into research and development for improvement. In the mean time, research and development on improved approaches and the remaining components will continue.

The **Advanced Research and Architecture Development** Phase, closed by the Milestone MS6 (**Use Case Final Evaluation and Project End**) focuses on improving the results of the previous phase, including research and development of the remaining components (starting already at month 16), integration of the feedback of the first evaluation, and on integration, release of the second architecture prototype (month 32) and evaluation (months 33-36) of the second version of the INFORE infrastructure and the application components. Again, an additional milestone, in this case Milestones MS5 (**Final Architecture Prototype)** will define the intermediate goal of completing major building blocks for the final version of the infrastructure in month 32. Akin to the previous phase, continuous integration and automated regression testing will be carried out using the extended testing data set. The second version of the INFORE infrastructure will be subject to an overall experimental and expert evaluation.

In parallel, as part of the **Finalization Phase**, WP4-WP6 will continue working on the final versions of the methods and components, i.e., integrating intermediary feedback from the second evaluation as well as internal/external suggestions collected during the refinement phase and improving documentation. This phase is also closed by Milestone MS6 in month 36.

A crucial factor for the success of the project involves the Human Factors evaluation. The project will make best use of experts by focusing on collection of qualitative data in real work settings. **In all three use cases we will perform an evaluation of the INFORE prototype by expert users.**