Horizon 2020

Call: H2020-ICT-2016-2017
(Information and Communication Technologies Call)

Topic: ICT-16-2017

Type of action: RIA
(Research and Innovation action)

Proposal number: 780492

Proposal acronym: DEOS

Deadline Id: H2020-ICT-2017-1
Table of contents

Section	*Title*	*Action*
1	General information	
2	Participants & contacts	
3	Budget	
4	Ethics	
5	Call-specific questions	

How to fill in the forms

The administrative forms must be filled in for each proposal using the templates available in the submission system. Some data fields in the administrative forms are pre-filled based on the previous steps in the submission wizard.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS**

1 - General information

Topic	ICT-16-2017
Call Identifier	H2020-ICT-2016-2017
Type of Action	RIA
Deadline Id	H2020-ICT-2017-1

Acronym	DEOS
Proposal title*	Decentralised Earth Observation Systems

Note that for technical reasons, the following characters are not accepted in the Proposal Title and will be removed: < > " &

Duration in months	*36*		
Fixed keyword 1	*Geo-information and spatial data analysis*	Add	
Fixed keyword 2	*Real time data analytics*	Add	Remove
Fixed keyword 3	*Scalability*	Add	Remove
Fixed keyword 4	*Data curation*	Add	Remove
Fixed keyword 5	*Data visualization*	Add	Remove
Free keywords	*Earth observation, Approximate Computing, Distributed processing, Decentralized systems, Hardware accelerators, Deep memory hierarchies, Near real time services, Mixed precision, Graph Processing.*		

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS**

Abstract

DEOS aims to develop near real time Earth observation (EO) services for the benefit of society. To achieve this, huge datasets produced by the Sentinel satellites under the European Copernicus project must be processed; but such high resolution datasets are noisy and sometimes incomplete due to clouds and differences in resolution and revisit frequencies. With the coming generation of satellites, the volume, velocity and variety of these datasets will increase dramatically. Moreover, processing big data translates into consuming big energy.
Current datacenters consume several megawatts of energy, a large part of it going into the cooling system, which raises questions about the sustainability of those infrastructures. The efficient and sustainable processing of EO data represents a great challenge that cannot be addressed without the right infrastructure, technology and robust scientific methods.
To accomplish this objective, DEOS proposes to push the envelop of big data processing in four different levels: i) DEOS will design innovative EO applications and deliver near real time services ii) DEOS will investigate novel approaches of data reduction, approximate and trans-precision computing and graph processing to speedup the data processing while losing negligible accuracy iii) DEOS will produce middleware to leverage new hardware devices, in order to accelerate the data processing as well as the visualization process iv) DEOS proposes a disruptive sustainable approach in which EO datasets are processed in apartment heaters distributed geographically. This involves the implementation of tools for data streaming and distributed task processing.
The environmental impact of DEOS is without question, both in terms of services delivered, as well as a role model of sustainability.
Nonetheless, the applications can have an incredibly broader impact.
Overall, DEOS aims to be the most sustainable and efficient platform for Earth Observation data processing in the world.

Remaining characters 0

Has this proposal (or a very similar one) been submitted in the past 2 years in response to a call for proposals under Horizon 2020 or any other EU programme(s)? ○ Yes ◉ No

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS**

Declarations

1) The coordinator declares to have the explicit consent of all applicants on their participation and on the content of this proposal.	☒
2) The information contained in this proposal is correct and complete.	☒
3) This proposal complies with ethical principles (including the highest standards of research integrity — as set out, for instance, in the European Code of Conduct for Research Integrity — and including, in particular, avoiding fabrication, falsification, plagiarism or other research misconduct).	☒

4) The coordinator confirms:

- to have carried out the self-check of the financial capacity of the organisation on http://ec.europa.eu/research/participants/portal/desktop/en/organisations/lfv.html or to be covered by a financial viability check in an EU project for the last closed financial year. Where the result was "weak" or "insufficient", the coordinator confirms being aware of the measures that may be imposed in accordance with the H2020 Grants Manual (Chapter on Financial capacity check); or	○
- is exempt from the financial capacity check being a public body including international organisations, higher or secondary education establishment or a legal entity, whose viability is guaranteed by a Member State or associated country, as defined in the H2020 Grants Manual (Chapter on Financial capacity check); or	◉
- as sole participant in the proposal is exempt from the financial capacity check.	○

5) The coordinator hereby declares that each applicant has confirmed:

- they are fully eligible in accordance with the criteria set out in the specific call for proposals; and	☒
- they have the financial and operational capacity to carry out the proposed action.	☒

The coordinator is only responsible for the correctness of the information relating to his/her own organisation. Each applicant remains responsible for the correctness of the information related to him/her and declared above. Where the proposal to be retained for EU funding, the coordinator and each beneficiary applicant will be required to present a formal declaration in this respect.

According to Article 131 of the Financial Regulation of 25 October 2012 on the financial rules applicable to the general budget of the Union (Official Journal L 298 of 26.10.2012, p. 1) and Article 145 of its Rules of Application (Official Journal L 362, 31.12.2012, p.1) applicants found guilty of misrepresentation may be subject to administrative and financial penalties under certain conditions.

Personal data protection

The assessment of your grant application will involve the collection and processing of personal data (such as your name, address and CV), which will be performed pursuant to Regulation (EC) No 45/2001 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data. Unless indicated otherwise, your replies to the questions in this form and any personal data requested are required to assess your grant application in accordance with the specifications of the call for proposals and will be processed solely for that purpose. Details concerning the purposes and means of the processing of your personal data as well as information on how to exercise your rights are available in the privacy statement. Applicants may lodge a complaint about the processing of their personal data with the European Data Protection Supervisor at any time.

Your personal data may be registered in the Early Detection and Exclusion system of the European Commission (EDES), the new system established by the Commission to reinforce the protection of the Union's financial interests and to ensure sound financial management, in accordance with the provisions of articles 105a and 108 of the revised EU Financial Regulation (FR) (Regulation (EU, EURATOM) 2015/1929 of the European Parliament and of the Council of 28 October 2015 amending Regulation (EU, EURATOM) No 966/2012) and articles 143 - 144 of the corresponding Rules of Application (RAP) (COMMISSION DELEGATED REGULATION (EU) 2015/2462 of 30 October 2015 amending Delegated Regulation (EU) No 1268/2012) for more information see the Privacy statement for the EDES Database).

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS**

List of participants

#	Participant Legal Name	Country
1	BARCELONA SUPERCOMPUTING CENTER - CENTRO NACIONAL DE SUPERCOMPUTACION	Spain
2	QARNOT COMPUTING	France
3	UNIVERSITAT DE VALENCIA	Spain
4	IBM RESEARCH GMBH	Switzerland

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

2 - Administrative data of participating organisations

PIC	***Legal name***
999655520	*BARCELONA SUPERCOMPUTING CENTER - CENTRO NACIONAL DE SUPERCOMPUTACION*

Short name: BSC

Address of the organisation

Street Calle Jordi Girona 31

Town BARCELONA

Postcode 08034

Country Spain

Webpage www.bsc.es

Legal Status of your organisation

Research and Innovation legal statuses

Public body ..... yes

Non-profit ..... yes

International organisation ..... no

International organisation of European interest ..... no

Secondary or Higher education establishment ..... no

Research organisation ..... yes

Legal person ..... yes

Enterprise Data

SME self-declared status ..... 01/03/2005 - no

SME self-assessment ..... unknown

SME validation sme ..... unknown

Based on the above details of the Beneficiary Registry the organisation is not an SME (small- and medium-sized enterprise) for the call.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Department(s) carrying out the proposed work

Department 1

Department name: Computer Sciences ☐ not applicable

☒ Same as organisation address

Street: Calle Jordi Girona 31

Town: BARCELONA

Postcode: 08034

Country: Spain

Department 2

Department name: Earth Sciences ☐ not applicable

☒ Same as organisation address

Street: Calle Jordi Girona 31

Town: BARCELONA

Postcode: 08034

Country: Spain

Dependencies with other proposal participants

Character of dependence	***Participant***	

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS** *Short name* **BSC**

Person in charge of the proposal

The name and e-mail of contact persons are read-only in the administrative form, only additional details can be edited here. To give access rights and basic contact details of contact persons, please go back to Step 4 of the submission wizard and save the changes.

Title: Dr. Sex: (•) Male ( ) Female

First name: **Leonardo** Last name: **Bautista**

E-Mail: **leonardo.bautista@bsc.es**

Position in org.: Senior Researcher

Department: Computer Sciences ☐ Same as organisation

☒ Same as organisation address

Street: Calle Jordi Girona 31

Town: BARCELONA Post code: 08034

Country: Spain

Website: www.bsc.es

Phone 1: +34 934017313 Phone 2: +xxx xxxxxxxxx Fax: +xxx xxxxxxxxx

Other contact persons

First Name	*Last Name*	*E-mail*	*Phone*
Sergi	MADONAR	sergi.madonar@bsc.es	
Lucinda	Cash-Gibson	lucinda.cashgibson@bsc.es	
Marc	Casas	marc.casas@bsc.es	
Francisco	Doblas-Reyes	francisco.doblas-reyes@bsc.es	
Pierre-Antoine	Bretonnière	pierre-antoine.bretonniere@bsc.es	
Alicia	Sánchez	alicia.sanchez@bsc.es	
Kim	Serradell	kim.serradell@bsc.es	
Mar	Rodríguez	mar.rodriguez@bsc.es	

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS** *Short name* **QARNOT COMPUTING**

PIC	***Legal name***
938181576	*QARNOT COMPUTING*

Short name: QARNOT COMPUTING

Address of the organisation

Street 10, rue Richer

Town PARIS

Postcode 75009

Country France

Webpage www.qarnot-computing.com

Legal Status of your organisation

Research and Innovation legal statuses

Public body	no	Legal person	yes
Non-profit	no		
International organisation	no		
International organisation of European interest	no		
Secondary or Higher education establishment	no		
Research organisation	no		

Enterprise Data

SME self-declared status	31/12/2015 - yes
SME self-assessment	31/12/2015 - yes
SME validation sme	unknown

Based on the above details of the Beneficiary Registry the organisation is an SME (small- and medium-sized enterprise) for the call.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492**	*Acronym* **DEOS**	*Short name* **QARNOT COMPUTING**

Department(s) carrying out the proposed work

No department involved

Department name		☒ not applicable
	☐ Same as organisation address	
Street	*Please enter street name and number.*	
Town		
Postcode		
Country		

Dependencies with other proposal participants

Character of dependence	***Participant***	

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS** *Short name* **QARNOT COMPUTING**

Person in charge of the proposal

The name and e-mail of contact persons are read-only in the administrative form, only additional details can be edited here. To give access rights and basic contact details of contact persons, please go back to Step 4 of the submission wizard and save the changes.

Title: Mr.

Sex: (•) Male ( ) Female

First name: **Nicolas**

Last name: **SAINTHERANT**

E-Mail: **nicolas.saintherant@qarnot-computing.com**

Position in org.: Innovation Manager

Department: *Please indicate the department of the Contact Point above in the organisation*

☐ Same as organisation

☐ Same as organisation address

Street: 42 rue Barbès

Town: MONTROUGE

Post code: 92120

Country: France

Website: www.qarnot-computing.com

Phone 1: +33681143738

Phone 2: *+xxx xxxxxxxxx*

Fax: *+xxx xxxxxxxxx*

Other contact persons

First Name	***Last Name***	***E-mail***	***Phone***
Yanik	NGOKO	yanik.ngoko@qarnot-computing.com	+33681224484

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS** *Short name* **UVEG**

PIC	***Legal name***
999953019	*UNIVERSITAT DE VALENCIA*

Short name: UVEG

Address of the organisation

Street AVENIDA BLASCO IBANEZ 13

Town VALENCIA

Postcode 46010

Country Spain

Webpage www.uv.es

Legal Status of your organisation

Research and Innovation legal statuses

Public body ..... yes

Non-profit ..... yes

International organisation ..... no

International organisation of European interest ..... no

Secondary or Higher education establishment ..... yes

Research organisation ..... yes

Legal person ..... yes

Enterprise Data

SME self-declared status ..... 12/03/2014 - no

SME self-assessment ..... 12/03/2014 - no

SME validation sme ..... unknown

Based on the above details of the Beneficiary Registry the organisation is not an SME (small- and medium-sized enterprise) for the call.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Department(s) carrying out the proposed work

Department 1

Department name: Image Processing Laboratory (LPI) ☐ not applicable

☐ Same as organisation address

Street: C/Catedrático José Beltran, 2

Town: Paterna

Postcode: 46980

Country: Spain

Dependencies with other proposal participants

Character of dependence	***Participant***	

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS** *Short name* **UVEG**

Person in charge of the proposal

The name and e-mail of contact persons are read-only in the administrative form, only additional details can be edited here. To give access rights and basic contact details of contact persons, please go back to Step 4 of the submission wizard and save the changes.

Title: Dr. Sex: ○ Male ◉ Female

First name: **Maria** Last name: **Piles**

E-Mail: **maria.piles@gmail.com**

Position in org.: Research Scientist

Department: Laboratorio de Procesado de Imágenes (IPL) ☐ Same as organisation

☐ Same as organisation address

Street: C/Catedrático José Beltran, 2

Town: Paterna Post code: 46980

Country: Spain

Website: http://ipl.uv.es/

Phone 1: +34 963543229 Phone 2: +34 963544161 Fax: +xxx xxxxxxxxx

Other contact persons

First Name	***Last Name***	***E-mail***	***Phone***
Ángeles	Sanchis	angeles.sanchis@uv.es	+34963983621

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS** *Short name* **IBM**

PIC	***Legal name***
999909854	*IBM RESEARCH GMBH*

Short name: IBM

Address of the organisation

Street SAEUMERSTRASSE 4

Town RUESCHLIKON

Postcode 8803

Country Switzerland

Webpage www.zurich.ibm.com

Legal Status of your organisation

Research and Innovation legal statuses

Public body ..... no

Non-profit ..... no

International organisation ..... no

International organisation of European interest ..... no

Secondary or Higher education establishment ..... no

Research organisation ..... no

Legal person ..... yes

Enterprise Data

SME self-declared status ..... 08/02/2016 - no

SME self-assessment ..... unknown

SME validation sme ..... 25/11/2008 - no

Based on the above details of the Beneficiary Registry the organisation is not an SME (small- and medium-sized enterprise) for the call.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Department(s) carrying out the proposed work

Department 1

Department name: Cognitive Computing & Industry Solutions ☐ not applicable

☒ Same as organisation address

Street: SAEUMERSTRASSE 4

Town: RUESCHLIKON

Postcode: 8803

Country: Switzerland

Dependencies with other proposal participants

Character of dependence	***Participant***	

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS** *Short name* **IBM**

Person in charge of the proposal

The name and e-mail of contact persons are read-only in the administrative form, only additional details can be edited here. To give access rights and basic contact details of contact persons, please go back to Step 4 of the submission wizard and save the changes.

Title: Dr. Sex: (•) Male ( ) Female

First name: **A. Cristiano I.** Last name: **Malossi**

E-Mail: **acm@zurich.ibm.com**

Position in org.: Research Staff Member

Department: Cognitive Computing & Industry Solutions ☐ Same as organisation

☒ Same as organisation address

Street: SAEUMERSTRASSE 4

Town: RUESCHLIKON Post code: 8803

Country: Switzerland

Website: http://researcher.watson.ibm.com/researcher/view.php?person=zuric

Phone 1: +41 44 724 8616 Phone 2: +xxx xxxxxxxxx Fax: +xxx xxxxxxxxx

Other contact persons

First Name	***Last Name***	***E-mail***	***Phone***
Costas	Bekas	bek@zurich.ibm.com	+41 44 724 8969
Sara	Pittaluga	sar@zurich.ibm.com	+41 44 724 8567
Catherine	Trachsel	ctr@zurich.ibm.com	+41 44 724 8289

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS**

3 - Budget for the proposal

No	Participant	Country	(A) Direct personnel costs/€	(B) Other direct costs/€	(C) Direct costs of sub-contracting/€	(D) Direct costs of providing financial support to third parties/€	(E) Costs of inkind contributions not used on the beneficiary's premises/€	(F) Indirect Costs / € (=0.25(A+B-E))	(G) Special unit costs covering direct & indirect costs / €	(H) Total estimated eligible costs / € (=A+B+C+D+F+G)	(I) Reimburse-ment rate (%)	(J) Max.EU Contribution / € (=H*I)	(K) Requested EU Contribution/ €
1	Bsc	ES	521400	84178	10000	0	0	151394,50	0	766972,50	100	766972,50	766972,50
2	Qarnot Computing	FR	388800	22000	0	0	0	102700,00	0	513500,00	100	513500,00	513500,00
3	Uveg	ES	326400	21000	0	0	0	86850,00	0	434250,00	100	434250,00	434250,00
4	Ibm	CH	405300	21200	0	0	0	106625,00	0	533125,00	100	533125,00	533125,00
	Total		1641900	148378	10000	0	0	447569,50	0	2247847,50		2247847,50	2247847,50

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS**

4 - Ethics issues table

1. HUMAN EMBRYOS/FOETUSES		Page
Does your research involve Human Embryonic Stem Cells (hESCs)?	○ Yes ◉ No	
Does your research involve the use of human embryos?	○ Yes ◉ No	
Does your research involve the use of human foetal tissues / cells?	○ Yes ◉ No	
2. HUMANS		**Page**
Does your research involve human participants?	○ Yes ◉ No	
Does your research involve physical interventions on the study participants?	○ Yes ◉ No	
3. HUMAN CELLS / TISSUES		**Page**
Does your research involve human cells or tissues (other than from Human Embryos/ Foetuses, i.e. section 1)?	○ Yes ◉ No	
4. PERSONAL DATA		**Page**
Does your research involve personal data collection and/or processing?	○ Yes ◉ No	
Does your research involve further processing of previously collected personal data (secondary use)?	○ Yes ◉ No	
5. ANIMALS		**Page**
Does your research involve animals?	○ Yes ◉ No	
6. THIRD COUNTRIES		**Page**
In case non-EU countries are involved, do the research related activities undertaken in these countries raise potential ethics issues?	○ Yes ◉ No	
Do you plan to use local resources (e.g. animal and/or human tissue samples, genetic material, live animals, human remains, materials of historical value, endangered fauna or flora samples, etc.)?	○ Yes ◉ No	
Do you plan to import any material - including personal data - from non-EU countries into the EU?	○ Yes ◉ No	
Do you plan to export any material - including personal data - from the EU to non-EU countries?	○ Yes ◉ No	
In case your research involves low and/or lower middle income countries, are any benefits-sharing actions planned?	○ Yes ◉ No	
Could the situation in the country put the individuals taking part in the research at risk?	○ Yes ◉ No	

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS**

7. ENVIRONMENT & HEALTH and SAFETY		Page
Does your research involve the use of elements that may cause harm to the environment, to animals or plants?	○ Yes ◉ No	
Does your research deal with endangered fauna and/or flora and/or protected areas?	○ Yes ◉ No	
Does your research involve the use of elements that may cause harm to humans, including research staff?	○ Yes ◉ No	
8. DUAL USE		**Page**
Does your research involve dual-use items in the sense of Regulation 428/2009, or other items for which an authorisation is required?	○ Yes ◉ No	
9. EXCLUSIVE FOCUS ON CIVIL APPLICATIONS		**Page**
Could your research raise concerns regarding the exclusive focus on civil applications?	○ Yes ◉ No	
10. MISUSE		**Page**
Does your research have the potential for misuse of research results?	○ Yes ◉ No	
11. OTHER ETHICS ISSUES		**Page**
Are there any other ethics issues that should be taken into consideration? Please specify	○ Yes ◉ No	

I confirm that I have taken into account all ethics issues described above and that, if any ethics issues apply, I will complete the ethics self-assessment and attach the required documents. ☒

How to Complete your Ethics Self-Assessment

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Proposal ID **780492** *Acronym* **DEOS**

5 - Call specific questions

Extended Open Research Data Pilot in Horizon 2020

If selected, applicants will by default participate in the Pilot on Open Research Data in Horizon 2020[1] , which aims to improve and maximise access to and re-use of research data generated by actions.

However, participation in the Pilot is flexible in the sense that it does not mean that all research data needs to be open. After the action has started, participants will formulate a Data Management Plan (DMP), which should address the relevant aspects of making data FAIR – findable, accessible, interoperable and re-usable, including what data the project will generate, whether and how it will be made accessible for verification and re-use, and how it will be curated and preserved. Through this DMP projects can define certain datasets to remain closed according to the principle "as open as possible, as closed as necessary". A Data Management Plan does not have to be submitted at the proposal stage.

Furthermore, applicants also have the possibility to opt out of this Pilot completely at any stage (before or after the grant signature). In this case, applicants must indicate a reason for this choice (see options below).

Please note that participation in this Pilot does not constitute part of the evaluation process. Proposals will not be penalised for opting out.

We wish to opt out of the Pilot on Open Research Data in Horizon 2020.	○ Yes	◉ No

Further guidance on open access and research data management is available on the participant portal: http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm and in general annex L of the Work Programme.

[1] *According to article 43.2 of Regulation (EU) No 1290/2013 of the European Parliament and of the Council, of 11 December 2013, laying down the rules for participation and dissemination in "Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020)" and repealing Regulation (EC) No 1906/2006.*

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

DEOS Proposal

(Technical annex)

Research and Innovation action ICT-16

Decentralised Earth Observation Systems

Acronym: DEOS

List of participants

Participant No	Participant organisation name	Country
1 (Coordinator)	Barcelona Supercomputing Center - Centro Nacional de Supercomputación (BSC)	ES
2	Qarnot Computing (QARNOT)	FR
3	Universitat de Valencia (UVEG)	ES
4	IBM Research GmbH (IBM)	CH

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Table of contents

1. Excellence **5**
1.1 Objectives 7
1.2 Relation to the work programme 8
1.3 Concept and methodology 9
1.3.1 Concept 9
1.3.2 Methodology 16
1.4 Ambition 19

2. Impact **20**
2.1 Expected impacts 20
2.2 Measures to maximise impact 21
2.2.1 Dissemination and exploitation of results 21
2.2.2 Communication activities 26

3. Implementation **29**
3.1 Work plan - Work packages, deliverables and milestones 29
3.1.1 Overall strategy 29
3.1.2 PERT Diagram - Interdependencies among Work Packages 30
3.1.3 Timing of the different work packages and their components 31
3.1.4 Detailed work description 32
3.1.5 List of Work Packages 48
3.1.6 List of deliverables 49
3.2 Management structure and procedures 50
3.2.1 Management structure 50
3.2.2 Management procedures and tools 53
3.2.3 List of milestones 54
3.2.4 Innovation management 55
3.2.5 Risk assessment and contingency strategies 55
3.3 Consortium as a whole 58
3.4 Resources to be committed 59

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Exordium

He placed his phone in the counter next to the mirror, turned on the heater and got in the shower, completely oblivious that his daughter was heading straight to a life-threatening situation. The three processors of his heater started communicating with a server over the internet. The accelerator started warming up and the hybrid memory got flooded with satellite data observations that were done that very morning produced by the Sentinel 2B satellite launched on March 7th 2017. This space jewel that weights over a ton, rotates around the Earth with its twin the satellite 2A at an altitude of 786 km with a speed of 14.3 revolutions per day, and had passed over western Europe just a few hours before. The constellation of the two multispectral satellites produce data with 13 bands in the visible, near infrared and short-wave infrared part of the spectrum. Using a principal component analysis approach, the dataset got reduced to only half of the bands while keeping over 99% of the relevant information. The remaining relevant bands were transferred to the heater and immediately placed in different parts of the hybrid memory according to their access frequency. Additionally, historical data of several precedent years were transferred and stored on the high capacity non-volatile memory.

His phone vibrated announcing a new message in the inbox. It was from his daughter. He didn't notice it. The temperature of the heater began to increase quickly, as the data started to be processed. Some computational kernels were executed in the standard processor, while others more compute-intensive were optimized to run on the accelerator. That morning, the data that was processed in that heater corresponded to the south-west region of the Iberian peninsula. The anomaly detection algorithms discovered an unusual peak in soil surface temperature together with alarming low levels of soil moisture. Making use of historical data, the application reached a conclusion: the likelihood of a wildfire occurring in the coming hours was very high. The process sent the results of this analysis to a server and an automatic alarm report was emailed to the forest reserve in question.

The shower went off and he reached for a towel. A second message arrived. This time he noticed it. It was from his daughter again; she was upset because their visit to the natural forest reserve, which was one of the big attractions of her school trip, had just been cancelled due to a high risk of wildfire. At that precise moment, thousands of kilometers away, the director of the forest reserve protection department ordered to activate the sprinklers as a preventive action. He looked at the report in his hand with a smile. The report was signed “DEOS”.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Abstract

DEOS aims to develop near real time Earth observation (EO) services for the benefit of society. To achieve this, huge datasets produced by the Sentinel satellites under the European Copernicus project must be processed; but such high resolution datasets are noisy and sometimes incomplete due to clouds and differences in resolution and revisit frequencies. With the coming generation of satellites, the volume, velocity and variety of these datasets will increase dramatically. Moreover, processing big data translates into consuming big energy. Current datacenters consume several megawatts of energy, a large part of it going into the cooling system, which raises questions about the sustainability of those infrastructures. The efficient and sustainable processing of EO data represents a great challenge that cannot be addressed without the right infrastructure, technology and robust scientific methods.

To accomplish this objective, DEOS proposes to push the envelop of big data processing in four different levels: i) DEOS will design innovative EO applications and deliver near real time services ii) DEOS will investigate novel approaches of data reduction, approximate and trans-precision computing and graph processing to speedup the data processing while losing negligible accuracy iii) DEOS will produce middleware to leverage new hardware devices, in order to accelerate the data processing as well as the visualization process iv) DEOS proposes a disruptive sustainable approach in which EO datasets are processed in apartment heaters distributed geographically. This involves the implementation of tools for data streaming and distributed task processing.

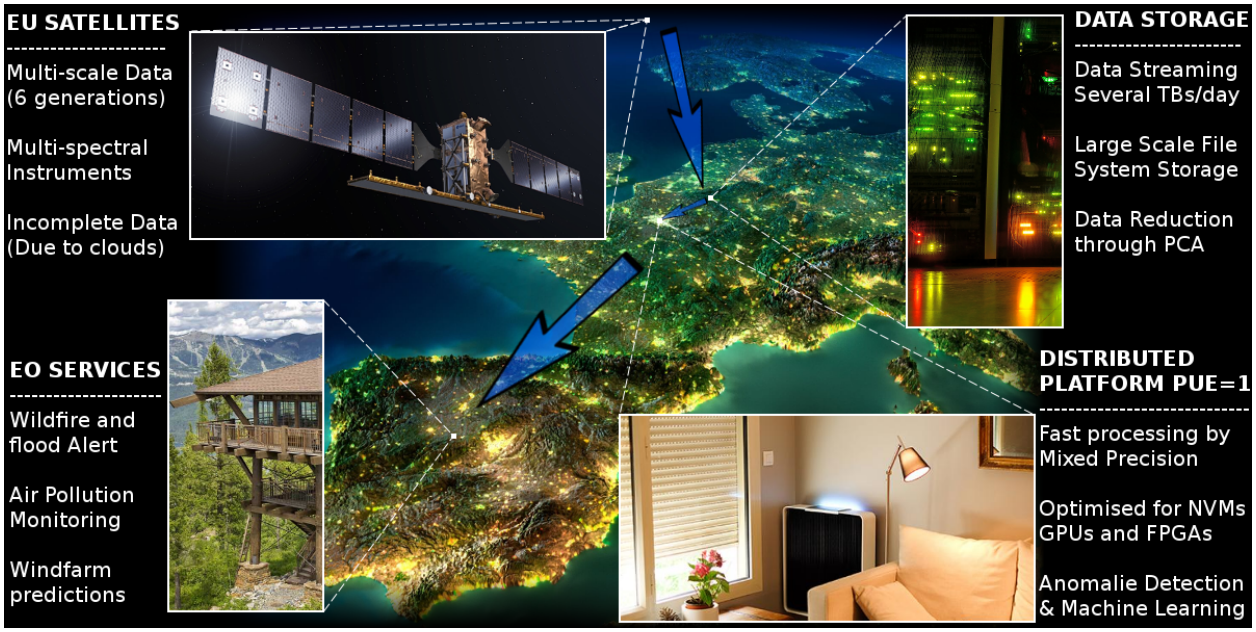
The environmental impact of DEOS is without question, both in terms of services delivered, as well as a role model of sustainability. Nonetheless, the applications can have an incredibly broader impact. Overall, DEOS aims to be the most sustainable and efficient platform for Earth Observation data processing in the world.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Decentralized Earth Observation Services

1. Excellence

Earth Observation (EO) aims to monitor the status of the natural environment of planet Earth and the anthropogenic impact. Given the amount and complexity of the data produced by imaging sensors onboard current satellite platforms, **EO data processing is becoming a fundamental tool for understanding not only our natural environment, but also the global economy**. The applications in EO data analysis vary from wildfire prediction, to cargo tracking systems, passing by crop growth analysis and air quality control. Such vast spectrum of applications touches many fundamental industry sectors, from biodiversity preservation to national security.



The importance of this enterprise is such that a large number of countries around the world have deployed numerous satellites to gather vast amounts of sensory data. Nowadays, EO is producing petabytes of data and the size of EO datasets will continue growing as space technology keeps on improving, and new satellite missions are launched. In Europe, the European Space Agency (ESA) is launching a new constellation of satellites[1], classed in 6 different generations combining numerous instruments for global land, ocean and atmospheric monitoring. The first Sentinel satellites (1A, 1B, 2A, 2B and 3A) from the first three generations

[1] http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Overview4

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

have already been launched and set in polar orbit. The first Sentinel satellites deployed in the last three years have produced several petabytes of data for the Copernicus program. With the recent launch of the second satellite Sentinel-2B, the second generation optical satellites will soon produce over two terabytes of data per day. The processing of EO data is only valuable if it can integrate recent datasets into the processing and deliver near real-time (NRT) services. Current EO service providers can only process a limited amount of data in order to deliver NRT services to the community. **In order to offer truly valuable NRT services with future massive EO datasets, it is critical to make efficient use of the existing and coming computing technologies and resources.**

In addition to the amount of data and the time limitations, future Sentinel generations will have a significantly different set of instruments measuring other aspects of the ocean, the continental surfaces, and the atmosphere. These new instruments acquire observations at multiple spectral bands coming from different sources on Earth, with different spatial resolutions and often polluted with clouds. Just filtering, correcting and decompressing those datasets, often stored in different formats, is a major challenge. However, it is only after preprocessing that the real work of extracting information starts. In order **to produce scientific and societal value out of those datasets, a sophisticated collection of highly efficient machine learning tools needs to be readily available** to discover variable relations, anomalies, extremes and trends.

In parallel to the growth of the space technology, another entirely separate universe has been growing exponentially: the internet of things (IoT). As widely anticipated, the number of devices with access to internet is exponentially growing. This seemingly unrelated trend might have an important impact in future EO services. Indeed, the number of large distributed denial of services (DDoS) attacks has been increasing in the last years and this trend is expected to continue[2]. The large availability of relatively poorly protected devices with access to the internet is making the gathering of massive DDoS botnets dramatically easy for experienced hackers. Thus, any centralized infrastructure providing EO services is vulnerable to be kicked out of service for some period of time, with consequences far more damaging than in the social media industry. If future communities are expected to rely on NRT EO services to take critical decisions, sometimes involving natural disaster response, such systems should have the highest security and reliability standards. Therefore, **it is important to achieve this challenging EO data processing in a completely decentralized fashion.**

The main objective of this project is to design and implement a sustainable low-power platform and software stack for distributed parallel processing of EO data and produce high quality NRT services improving the efficiency of current services by an order of magnitude. By leveraging big data software innovations and the most recent hardware technologies, this project will create an integrated system able to deliver unprecedented throughput rates in terms of data processing, analysis and visualization. The long-term

[2] Somani, Gaurav, et al. "DDoS attacks in cloud computing: issues, taxonomy, and future directions." *Computer Communications* (2017).

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

vision of DEOS is tied to open new frontiers and foster research towards efficient algorithms capable of processing large quantities of EO data, a stepping stone before the more ambitious far-end goal of global machine reasoning.

1.1 Objectives

The objective of DEOS is to make **a giant leap forward towards a fast and energy-efficient processing of high-resolution EO data to produce valuable insights for society**. To achieve this goal, we will improve by one order of magnitude the state-of-the-art of EO-based computing, on the four pillars of big data processing, namely space, time, energy and information. Therefore, we can breakdown DEOS goal in the following objectives:

Space: DEOS will be able to process EO datasets with a spatial resolution of up to a 100 meters produced by the recently launched Sentinel-2 satellites, which is over an order of magnitude higher resolution than commonly used in land surface analysis (i.e., 25 km2). While most regional data analysis studies focus on relatively small areas, usually in the order of several hundreds of thousands of square kilometers (e.g., Iberian peninsula 582,000 km2), DEOS is more ambitious and targets to cover the entire Europe (i.e., 10'180,000 km2) at high resolution. Similarly, we will cover ocean data, both Mediterranean and North Atlantic areas, with unprecedented spatial detail, and atmospheric profiles with high vertical resolution. This imposes significant demands on storage capacity and data streaming capabilities. **DEOS will develop novel data reduction techniques based on principal component analysis (PCA) as well as other approximate computing strategies to substantially reduce the size of those datasets and optimize data movements and storage (See Section 3.1.4 WP5).**

Time: The speed at which EO services are delivered is fundamental in some domains (e.g., wildfire prediction). DEOS aims to reduce the processing time of compute-bound EO data processing kernels by one order of magnitude in order to achieve NRT services. To achieve this objective, **we will design novel techniques to make efficient use of new computing technologies, both in storage and processing, such as non-volatile memories (NVM), graphic processing units (GPU) and field-programmable gate arrays (FPGA) (See Section 3.1.4 WP6)**. **Furthermore, we will introduce new aggressive approximate computing and trans-precision techniques that dramatically speed up data processing by allowing as little as 1% of error in the final results.** Furthermore, multiple accuracy levels can be offered in parallel giving the flexibility to the distributed task scheduler of deciding on the precision according to the current platform workload and NRT demands.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Energy: Processing more data faster has been a common goal in parallel processing and data analysis for several decades. Achieving the same objective while consuming much less energy is a relatively recent research focus. Indeed, in the era of mobile devices, huge datacenters and supercomputers that consume over 15 megawatts, reducing power consumption has become increasingly important. This point is even more critical given the alarming rate of global warming and the disastrous consequences of climate change. For this reason, DEOS takes an unprecedented green computing approach, in which we leverage the heating needs of modern society to process huge EO datasets on a daily basis. In this way, DEOS aims to **implement a continent-wide decentralized EO platform with a perfect power usage effectiveness (i.e., PUE=1.0)**. This distributed platform is much more resistant to DDoS attacks than solutions focused on a centralized datacenter. However, this imposes significant challenges, such as the data transfer and storage, as well as the scheduling of EO data processing tasks on such a distributed infrastructure. There are other issues of computing in heaters distributed across the country such as resiliency and robustness to errors. **DEOS will develop a software stack capable of leveraging such a widely distributed infrastructure for data processing (See Section 3.1.4 WP3)**. In addition we will study the impact of memory errors (common at high temperatures) on EO applications.

Information: Big data processing is only as valuable as the insights it produces for society. DEOS aims to **analyse EO datasets across the entire European region in order to detect i) extremely dry soil and temperature conditions posing a risk of wildfire, ii) soil moisture extremes to predict droughts and floods, iii) atmospheric profiles of essential climate variables and trace gases, such as water vapor, NO2 and ozone concentration and iv) sea ice monitoring and diagnostics (See Section 3.1.4 WP4)**. While some of these environmental analysis have been done in the past at low-resolution in relatively small geographical locations, a comprehensive analysis at continental scale and at this resolution has never been achieved. This is a dramatic game-changer enabled by the big data era, in which we will switch from the classic "*show me there*" EO service to a much more global and insightful "*show me where*" service.

1.2 Relation to the work programme

In order to accomplish the objectives of the DEOS project (See Section 1.1 above), it is mandatory to touch over all the topics of the work programme. The DEOS workflow starts by accessing "*extremely large numbers of high volume streams of noisy and possibly incomplete data*" (due to clouds and other factors) produced by the EU initiative, Copernicus. The Sentinel satellites produce over 1 terabyte of EO data on a daily basis. Those datasets will be decomposed in independent chunks and distributed across computing heaters (actually cloud High

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Performance Computing energy-efficient units) throughout different European cities. The objective is to perform a completely decentralized "*distributed data and process mining, predictive analysis and visualization*". These computing heaters constitute a green cloud computing grid where actual computing power is subject to evolve according to external constraints. Therefore, the energy-efficient data processing algorithms deployed must interact at the highest efficiency with the low-power hardware underneath. This will require the development of a new "*software stack designed to help programmers and big data practitioners take advantage of novel architectures*" such as deep memory hierarchies and processing accelerators. Techniques such as approximate computing and robust machine learning algorithms that produce accurate results under the presence of errors will be optimized to run in this distributed green computing platform. The result of those computations will be then offered as a service for the entire European Union, both to public agencies and to interested enterprises as well. Some of the services will be offered in near real time possibly upon a fee if a potential market is found (See Section 2.2.1) while others will be offered as open data for world wide researchers and citizens.

1.3 Concept and methodology

Big data also means big processing and huge storage capacity. Put in terms of energy consumption this translates into a potential power consumption catastrophe as data centers spend huge amounts of energy. Efforts to understand the climate and the environment could end up even hurting the environment if measures are not taken to limit the amount of energy consumed in data centers. A clear example of this is that by 2015, all the datacenters of Europe consumed more energy than Switzerland. Moreover, the growth of such infrastructures seems to follow an exponential trend. **The DEOS consortium firmly believes that an important part of the energy spent into computing and storing big datasets is actually waste that could be dramatically reduced without damaging the final results.** Hence, it is critical to design innovative and disruptive approaches to process all that data in the most green and sustainable fashion and this requires efforts in all levels of the big data software stack.

1.3.1 Concept

DEOS is overall a big data project, and as such it aims to improve all the layers of the software stack, from the high level EO data processing application layer down to the low level optimizations to efficiently leverage new hardware technologies, passing by the machine learning algorithms and the distributed data handling middleware layers. DEOS will focus on the specific aspects of each software layer as follows:

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

EO Application Layer : DEOS will leverage the data produced by the European Sentinel Satellites launched under the Copernicus initiative, as well as other space platforms, in order to provide high resolution services. The feasibility of the DEOS platform will be demonstrated through four use cases covering land, atmosphere and ocean applications:

- Case 1 - Soil moisture retrievals and early warning of flood and drought events: Present operations at BEC provide daily soil moisture maps at 1 km spatial resolution over the Iberian Peninsula from the optimal blend of SMOS and MODIS NDVI/LST data using the algorithm in (Piles et al., 2014; 2016)[3][4]. SMOS data stream is received at BEC facilities in NRT. MODIS data over the Iberian Peninsula is acquired and processed by a ground-segment at the University of Valladolid and provided to BEC in NRT. Both coverage and spatial resolution of BEC fine-scale soil moisture estimates are at present limited by the use of MODIS data, which is only available globally from NASA with a 15-day latency. The focus of this use case is to enhance the spatial coverage and spatio-temporal resolution of soil moisture retrievals with the use of Sentinel 2 NDVI and Sentinel 3 LST data, which is available in NRT over the whole European continent. Additionally, **the presence of clouds masking LST measurements will be mitigated by the combination of Sentinel 3 with MSG SEVIRI data, which provides data every 15 minutes**. DEOS target is to set up an inversion scheme able to provide daily soil moisture maps at 100-m spatial resolution over Europe and to exploit this information by setting up an alert and monitoring system in case of hydrological extremes (drought and floods).

- Case 2 - Wildfire prevention service: This use case aims at the development of maps of wildfire risk based on anomalies and extremes of soil moisture (from Case 1) and temperature and wind speed predictions. Risk assessment will be based on an upgraded version of the modeling framework[5] and will be provided in near real-time. Recent climate trends evidence a rise of temperatures and an increase in the duration and intensity of droughts which is in turn leading to the occurrence of larger wildfires, which threaten the environment as well as human lives and beings. In this context, improved wildfires prediction tools are urgently needed. **DEOS will develop an interactive service interface to explore soil moisture and temperature trends and anomalies and probabilistic wind speed and wildfire predictions in near real time**. The aim is to provide support to present fire risk prevention services with the new capabilities provided by Copernicus EO data. In addition, wind farm operators and energy traders might also benefit from DEOS wind speed forecasts.

[3] **Piles, María**, et al. "A downscaling approach for SMOS land observations: Evaluation of high-resolution soil moisture maps over the Iberian Peninsula." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7.9 (2014): 3845-3857.
[4] **Piles, María**, et al. "Towards improved spatio-temporal resolution soil moisture retrievals from the synergy of SMOS and MSG SEVIRI spaceborne observations." *Remote Sensing of Environment* 180 (2016): 403-417.
[5] Chaparro, David, et al. "Surface moisture and temperature trends anticipate drought conditions linked to wildfire activity in the Iberian Peninsula." *European Journal of Remote Sensing* 49 (2016): 955-971.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

- Case 3 - Predictions of air quality, CALIOPE : CALIOPE is an operational air quality system developed in the Earth Science department and daily operated in the Mare Nostrum supercomputer. The system provides air quality forecast at 24h and 48h for Europe (12x12 km), the Iberian Peninsula (4x4 km) and Canary Islands (2x2 km). The focus of this use case is to estimate atmospheric profiles of ECV and trace gases -temperature, water vapor, nitrogen dioxide (N2O), methane (CH4), carbon monoxide (CO), carbon dioxide (CO2) and ozone (O3) concentration- from hyperspectral infrared sounders with unprecedented speed, accuracy and vertical resolution for enhanced air quality retrievals. **The proposed inversion scheme is based on machine learning regression methods that enable fast retrievals of atmospheric profiles from high volumes of data**[6]. They will be applied during the project life to hyperspectral infrared data from the Infrared Atmospheric Sounding Interferometer (IASI) onboard the MetOp-A satellite, and could also be applied to the Meteosat Third Generation Infrared Sounder (MTG-IRS) instrument on-board future Sentinel 4.

- Case 4 - Sea ice diagnostics: The focus of this use case is to compute some derived variables from the sea ice observations at high resolution from Sentinel 1 Synthetic Aperture Radar (SAR) data. From the sea ice concentration and sea ice thickness -2+1D (time, longitude, latitude)- fields provided by the satellites, the sea ice volume, sea ice area and sea ice extent (only depending of the time) are computed. **These new "observational" variables can be compared to the outputs of the models (EC-Earth for example) and help in the process of the forecast verification.**

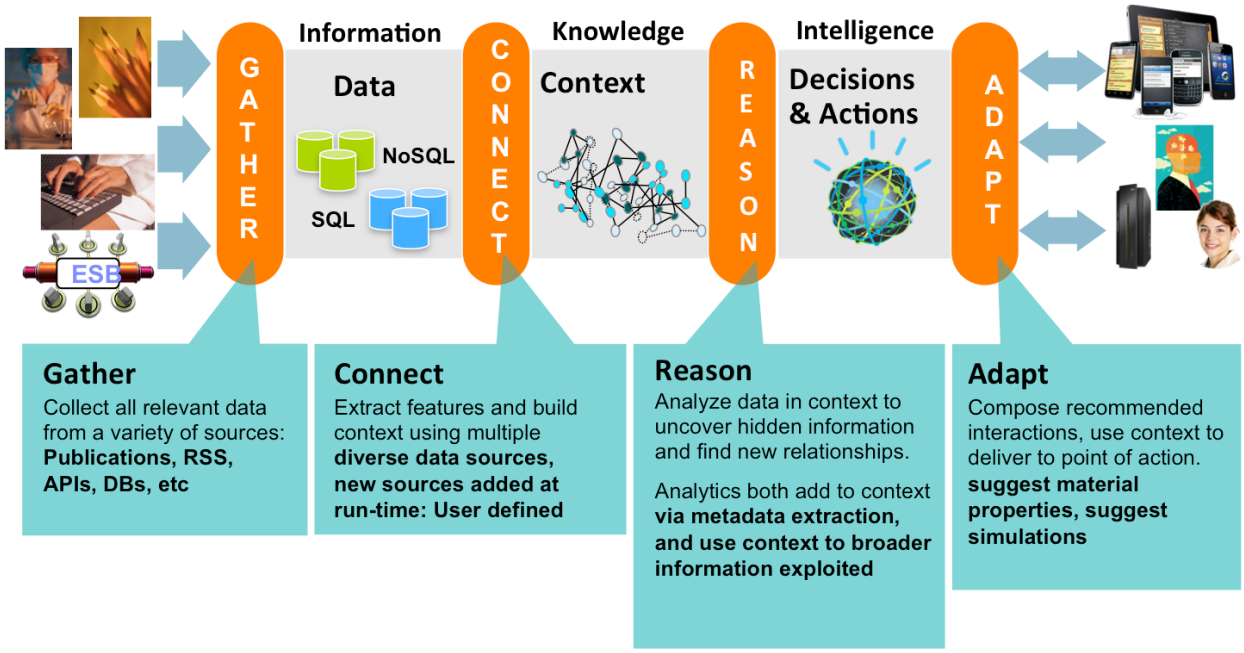
Data Mining and Knowledge Extraction Layer : Processing large volumes of data to extract knowledge and information on low-power platform is challenging. Indeed, Big Data analytics is one of today's hot topics. In the following figure we schematically show a typical pipeline of operations when processing large datasets to extract knowledge and information. This kind of workflows for parallel processing are usually implemented using a specialized workflow framework such as PyCOMPSs[7]. In the *gather* phase, data are collected from different sources. Here the data needs to be pre-processed to run the subsequent analytic in an efficient manner. This is particularly relevant in a low-power platform, where I/O and memory might represent a real bottleneck in the full pipeline. Thus, novel methods to compress and reduce the complexity of the data without loosing information must be evaluated, optimised and deployed. Recent works[8] have shown that techniques such as PCA can reduce satellite data from 6 observation bands to only 3 layers while extracting 99.3% of the relevant information. DEOS proposes the systematic use of PCA as well as the development of more aggressive data reduction techniques

[6] Verrelst, Jochem, et al. "Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and-3." *Remote Sensing of Environment* 118 (2012): 127-139.
[7] Tejedor, Enric, et al. "PyCOMPSs: Parallel computational workflows in Python." *The International Journal of High Performance Computing Applications* 31.1 (2017): 66-82.
[8] Estornell, J., Martí-Gavliá, J. M., Sebastiá, M. T., & Mengual, J. (2013). Principal component analysis applied to remote sensing. *Modelling in Science Education and Learning*, *6*, 83-89.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

to reduce the stress on storage, communication and processing needs. This will be further optimized to run efficiently in a low-power platform.

Example of typical pipeline for knowledge extraction from data



In the next phase, *connect*, links between the new data and the existing one are established. This phase is particularly important when in the third phase, *reason*, techniques based on graphs (networks) or trees are employed. To establish connections, different methods can be used depending on the type of input data. In the case of images, deep learning and particularly convolutional neural networks represent effective tools to automatically detect and understand images content, so that it can be linked with other type of data. Both the *connect* and *reason* phases need to run in near real time on the low-power platform, to cope with the continuous flow of new input data. This poses a serious limitation on the complexity of the algorithms and the level of accuracy used in the main routines. **DEOS will develop Deep Learning and Graph Processing algorithms that combine trans-precision computing and energy efficiency techniques to run on the decentralized low-power platform**. As an example, in the connection phase and assembling of the adjacency matrix to be used by the subsequent graph analytics algorithms, we can employ *sparsification* or *randomization* techniques to reduce the complexity of the graph. The processing of the resulting graph will be based on approximated algorithms that will provide accurate prediction within an acceptable tolerance. As an example, stochastic estimators can be used to evaluate node centralities without computing the eigenvalues of the matrix[9]. **By combining this approach with trans-precision,**

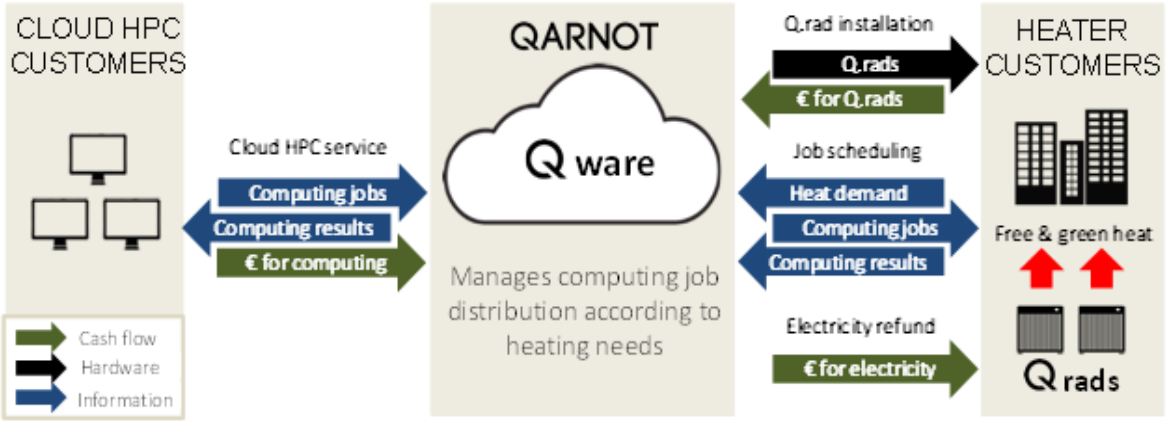
[9] **C. Bekas**, E. Kokiopoulou, and Y. Saad, "An estimator for the diagonal of a matrix," Applied Numerical Mathematics, vol. 57, no. 11–12, pp. 1214–1229, 2007.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

DEOS will push the performance of this algorithm so that it can be efficiently run on a low-power platform. In addition, we will investigate, communication-avoiding approaches to limit the number and volume of communications required for the target analysis. The last phase, *adapt*, closes the pipeline and represents the point of interaction with the users. This phase is typically inexpensive, however depending on the type of output some acceleration can also be delivered, in particular when 3D visualization is needed.

Distributed Computing Layer : DEOS will demonstrate that NRT services can be achieved while using the Qarnot model of *distributed heating clouds* based on *free cooling*[10]. This is a completely disruptive approach to cloud computing that drastically increases energy efficiency. On one side, Qarnot sells and deploys Q.rads computing heaters in residential, public and commercial buildings. On the other side, Qarnot sells HPC services to companies, as shown in the Figure below. A part of this revenue is used to refund electricity consumed by the Q.rads.

Model of a distributed heating cloud



Qarnot's business model can be summarized as follows: The **Q.ware distribution platform** dynamically and securely distributes computations according to heating needs, thus enabling the exploitation of the computing power of Qarnot's deployed Q.rad grid. The **Q.rad digital heater** constitutes a key element in this coming future by bringing Seamless Intelligence for domestic and local usage and distributed computing power to serve the economy. Qarnot infrastructure and platform constitute a totally disruptive alternative to traditional data centres:

- Providing a highly competitive High-Performance Cloud Computing service;
- Reducing the global carbon footprint by avoiding data centres consumption;
- Opening way to a huge computing capacity throughout the European territory.

[10] **Ngoko, Y.** (2016, August). Heating as a Cloud-Service, A Position Paper (Industrial Presentation). In European Conference on Parallel Processing (pp. 389-401). Springer International Publishing.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

To compute on such clouds, we will consider a processing architecture that includes three types of nodes: Server nodes, compute nodes and storage nodes. Here, the servers, compute and storage nodes are completely distributed in cities where they also produce heat for the benefit of buildings. On the server nodes is deployed a cloud resource manager that includes a scheduler and a storage manager. To start a computation, an HTTPS request must be sent to the cloud resource manager. We consider that each request can be associated with a *processing graph* (or workflow) whose nodes correspond to computations performed in a *docker* container[11]. Thus, **the processing of the request will correspond to successive deployments of docker containers, according to the precedence constraints of the processing graph**. During the processing, a container can save data by the mean of a storage system whose API is compatible with Amazon S3[12]. We will in particular use the Ceph system[13] whose scalability and robustness was demonstrated in big data experiments performed by the CERN[14]. We decided to use Docker in order to avoid the extra-virtualization costs of virtual machines[15] while keeping the great flexibility and modularity provided by virtualization.

It is important to observe that the Qarnot platform is highly distributed and built on top of an heterogeneous network that includes Ethernet Gigabytes and 10 Gb links in addition to the optical fiber. So, one could raise some network concerns regarding the suitability of the platform for compute-intensive applications. On this point, it is important to notice that if communication avoiding algorithms are more suitable to the Qarnot platform than highly coupled parallel algorithms, the Qarnot platform is mainly currently exploited by customers interested in 3D rendering, financial and meteorological simulations. Most of these applications are data-intensive and require several data transfer to the compute nodes. These applications are successfully operated on the Qarnot platform because of the intelligent data prefetching and streaming policies, supported by the Q.ware. Finally, let us observe that an extension (already funded) of the Qarnot heating platform is planned for 2019. This will incorporate a new type of highly coupled nodes (more than 32 processors per nodes) that will also serve as a boiler that produces hot water from processors.

Hardware Optimization Layer: With the arrival of new hardware technologies, computing nodes are getting more dense (including for Qarnot nodes as presented above) as they are equipped with many cores and processing accelerators such as GPUs and FPGAs. **DEOS will develop a middleware prototype to easily accelerate the processing of large datasets in those new devices.**

[11] https://www.docker.com/
[12] https://aws.amazon.com/fr/documentation/s3/
[13] http://docs.ceph.com/docs/master/start/intro/
[14] Van Der Ster, D., & Rousseau, H. (2015). *Ceph 30PB Test Report* (No. CERN-IT-Note-2015-002).
[15] Felter, W., Ferreira, A., Rajamony, R., & Rubio, J. (2015, March). An updated performance comparison of virtual machines and linux containers. In *Performance Analysis of Systems and Software (ISPASS), 2015 IEEE International Symposium on* (pp. 171-172). IEEE.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

A particular aspect of computing on radiators is to deal with temperature issues. Indeed, given that the objective of these computing devices is to produce heat, it is mandatory to check for temperature-induced errors. It is well-known in the computing community[16], that high temperatures can have a significant impact on the results produced by the machine. **DEOS will study the robustness of EO applications under the presence of errors and data corruption** and will also introduce lightweight online data monitoring techniques that continuously check for data integrity and result correctness. This could open new opportunities to process EO datasets in unprotected hardware such as a cloud of mobile phones or other low-power devices. This research on error tolerance will be developed in parallel to mixed precision schemes presented previously and both endeavors will exchange and benefit from each other results.

In addition to processing devices, new storage technologies such as NVM and 3D-stacked memories (3DRAM) are starting to arrive into the market offering new opportunities but also bringing with them new challenges. In particular, these young technologies offer a new trade-off between latency, bandwidth and storage capacity[17]. Future big data systems are likely to be populated with a hybrid combination of classic DRAM and new memory technologies. How to efficiently leverage them is still an open question. **DEOS will design a middleware that measures the access frequency of different datasets in order to place those with high frequency access in the low-latency memory bank while the other datasets are placed in the lower-latency (but larger storage capacity) memory**. This data dynamics are likely to be application-dependent and in some cases might even be input-dependent. The results of the PCA-based data reduction will be of great benefit for this research.

Please note that the DEOS consortium has a strong support (See letters of support in Section 6) from multiple industry leaders of the new computing and processing technologies and they will provide high-level guidelines during the project. They are very well aware of the impact that DEOS could make in their industry by demonstrating how to implement NRT EO services using those technologies.

In summary, by implementing NRT EO services, DEOS aims to improve and push beyond the state-of-the-art in all levels of the big data software stack by reducing computation and storage waste and speeding up the processing of large EO datasets at the same time. This involves new innovative software techniques from the machine learning community, as well as the optimal use of novel hardware devices and architectures and the implementation of a disruptive infrastructure that minimizes energy waste. The overall structure of the work that DEOS aims to accomplish is presented in the figure below and is further detailed in the work

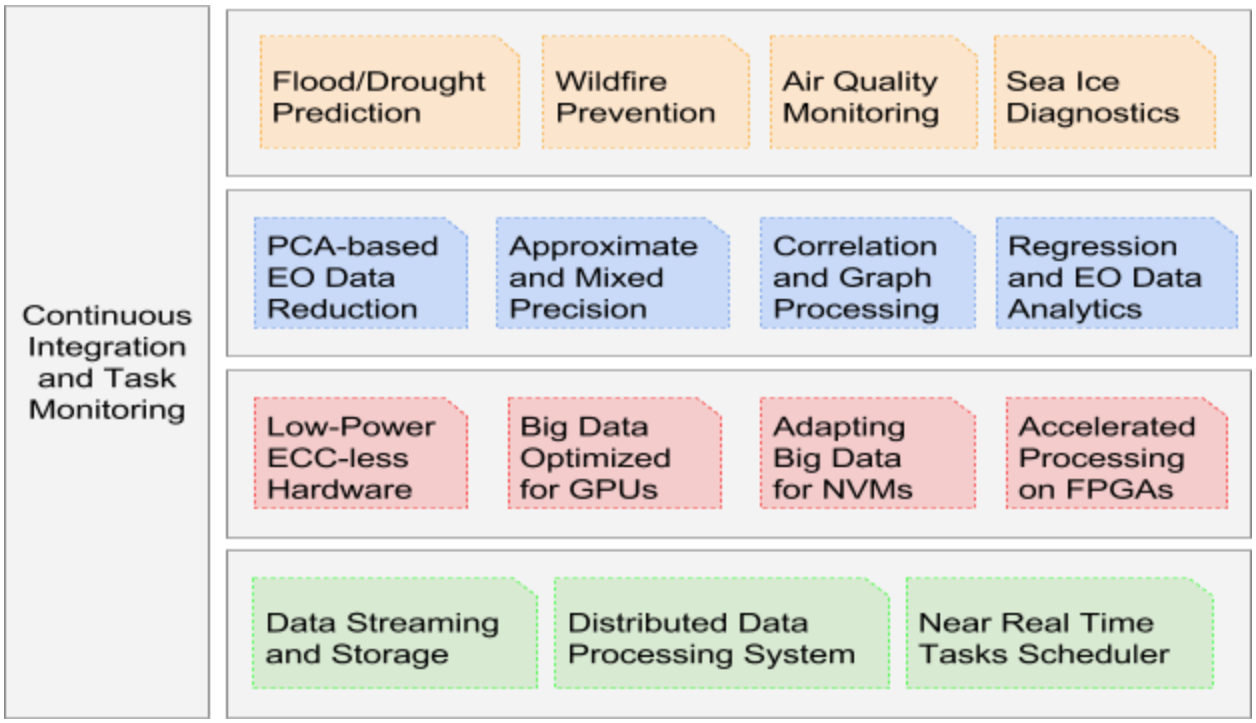
[16] Schroeder, B., Pinheiro, E., & Weber, W. D. (2009, June). DRAM errors in the wild: a large-scale field study. In *ACM SIGMETRICS Performance Evaluation Review* (Vol. 37, No. 1, pp. 193-204). ACM.

[17] Zivanovic, D., Pavlovic, M., Radulovic, M., Shin, H., Son, J., Mckee, S. A., ... & Ayguadé, E. (2017). Main Memory in HPC: Do We Need More or Could We Live with Less?. *ACM Transactions on Architecture and Code Optimization (TACO), 14*(1), 3.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

packages (See Section 3.1.4). **To this day, there is neither a distributed infrastructure aiming to leverage Copernicus datasets at this scale, nor a highly sustainable approach to EO data processing such as the one proposed by DEOS.**

DEOS Software Stack Improvements



1.3.2 Methodology

DEOS aims to demonstrate the viability of producing insightful NRT EO services for the community over a sustainable energy efficient computing platform. To demonstrate the power of the methods develop, we must provide real services running over real datasets and validate the results of the services provided. Thus, it is important to make sure that its services fit the needs of potential future users (See Section 3.2.1). Therefore, the first couple of years of the project will allow us to experiment with different EO services at different spatial and temporal resolutions, as well as different accuracy levels and measure the real need of end-users. For instance, entities interested on satellite ice monitoring might want to cover large surface areas but they might not give much importance to the update frequency (i.e., NRT service) as ice caps do not change dramatically in hours. On the other hand, firewatcher of natural reserves might not be interested on wide continental areas but only focus on a small region and for them time is the number one priority so that they can respond as fast as possible to the crisis. Based on this, we plan to study the European market of EO services following the methodology presented below.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Fast Service Development : DEOS will implement a fast cyclic process in which the EO services provided get **build, evaluated and improve** in a fast iterative process. Starting **from the first month of the DEOS project, we will set up a webpage offering the possibility to potential end-users to register** so that they get notified as soon as the first beta service is launched. Having a pool of early adopters will allow us to measure their needs in terms of spatial resolution, area coverage, processing speed and results accuracy. Not all service might have the same requirements and having this information will allow us recognize areas in which more effort should be spent. **Within the first semester, we expect to have a first minimal viable service** (MVS) that we will offer to beta-testers to get their feedback. In return, they can make specific service requests that will be considered by the DEOS consortium. This MVS is really a minimalistic example that will have a low resolution, area coverage and update frequency; but that will allow end-users to have a first idea about the services DEOS aims to provide. **Within the first year, DEOS will offer a real scale service**. This rapid *build-measure-learn* cycle will allows us to tune efforts where they are needed and improve the quality of the services. **By the end of the second year, DEOS will be providing multiple EO services covering uses cases related to land, sea and atmosphere at the European scale on a daily basis**. Finally, **by the end of the project, DEOS will be able to achieve smart automatic analysis of high-resolution land, atmospheric and ocean data at the European scale in a near real time basis** (For more details see Section 3.2.3). It is important to note that this process differs drastically from the long development phases that sometimes lead to unexpected obstacles by the end of the project. Instead, we aim to learn as much as possible during the first months of the project and then further feed that learning loop.

Technology readiness level : DEOS will rely on technologies already on the market or arriving to the market soon. In other terms, their Technology Readiness Level (TRL) is estimated to be between TRL8 and TRL9. Multiple of the satellites are already in orbit producing data on a daily basis and several other will be launched during the timeline of the DEOS project. On the ground, thousands of building heaters based on computing technology are about to be deployed across Paris and in other European cities. This distributed green computing grid has already demonstrated its processing capabilities with compute-intensive rendering of 3D movies. Next generation heating devices, able to use compute modules using alternative processing units such as GPU for example, are planned to hit the market during the timeline of this project. New storage and processing capabilities are therefore being developed to take advantage of these new computing architecture and thus compute other kind of jobs in optimal conditions such as big data. On the other side, Copernicus open-data servers have been providing datasets to the public for several years already. Thus, it is fair to say that almost all the hardware pieces are already available and have demonstrated their potential. It follows that the main idea of the DEOS project is to take all those pieces and glue them together with the required software components to make them work as a whole.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Relationship with other projects : At least 3 already funded H2020 projects have been or are working in a way or another on distributed computing with Earth Observation data: IGIT (Integrated geo-spatial information technology and its application to resource and environmental management towards the GEOSS)[18], MILLENNIUM (European climate of the last millennium)[19] and MED-SUV (MEDiterranean SUpersite Volcanoes)[20].

In IGIT, the objective was to develop the elements of a prototype system for data collection, analysis and dissemination for decision makers and for the public. A specific work package was dedicated to distributed computing, especially the "*Building multi-scale model based on CEBERS, SPOT and Landsat TM images*", "*New methods for metadata conversion in order to study the metadata of ILWIS operations and OpenGIS WPS specification*" and the "*Development of two grid nodes as WebGIS Map servers*" tasks are relevant.

In MILLENNIUM, the objective was to assess if the magnitude and rate of 20th century climate change exceeded the natural variability of European climate over the last millennium. They used in this project the climateprediction.net distributed computed network, running climate simulations on people's home computers.

In MED-SUV, the objective was to "*develop next generation geo hazard observing systems*" to detect and better understand of volcanic eruptions from Campi Flegrei/Vesuvius and Mt. Etna volcanoes. They used the expertise of one of the consortium's partners, Terradue UK Ltd. whose core business is Grid/Cloud based distributed computing (infrastructure and platform) and distributed data discovery and access, from global environmental analysis using large amounts of EO data, to regional land change detection with complex 3rd party algorithms.

DEOS will analyse the challenges encountered on theses EO distributed computing projects and leverage the outcomes in order to quickly start building on top of the state of the art big data technologies for Earth observation.

Gender Analysis : The DEOS consortium recognizes the importance of the promotion of a fair balance of men and women within the project. Despite the gender gap existing within the IT sector, among the DEOS consortium key personnel there are more than 11% of women. This should be seen as a small first step forward to the equal gender balance objectives of the EC and a step towards our goal of a more fair gender balance during the project. In DEOS, the gender aspects will be considered and promoted by all consortium partners and monitored by the Project Coordinator. Data on gender participation in the project (numbers and roles) will be collected and monitored with a view to benchmarking against employment best practice.

[18] http://cordis.europa.eu/result/rcn/168115_en.html
[19] http://cordis.europa.eu/project/rcn/78551_en.html
[20] http://cordis.europa.eu/result/rcn/191839_en.html

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

1.4 Ambition

DEOS ambition is to be the most efficient and sustainable decentralized platform for NRT EO data processing in the world. Some companies offer services based on the analysis of satellites images. However, they usually focus on very specific localized services that do not require processing large datasets. DEOS aims to switch from this localized perspective to a more wide view where an entire continent is constantly monitored on a daily basis. This is an important step forward, as this continent-wide monitoring will allow us to detect cycles and patterns that cannot be detected otherwise. This is a drastic change in the type of questions one can ask, from the classic "*show me there*" to a more broad "*show me where*" mindset. **DEOS will be a stepping stone towards a NRT planetary monitoring and understanding of the environment and the economy.**

It is understandable that most current EO services focus on specific local questions since not every company can afford a large data-processing center to analyse vast regions. The few companies that do have the infrastructure to process huge datasets in NRT (e.g., Google, Amazon) tend to do it in classic datacenters that are not geographically distributed and that consume an important amount of energy cooling the machine where the processing is taking place. DEOS takes a radical approach by developing the required software infrastructure to perform this challenging processing on a decentralized and highly energy efficient platform. **DEOS will be the only EO platform geographically distributed in tens of thousands of nodes spread out across multiple cities, achieving a perfect power usage efficiency**.

Furthermore, energy efficiency and green computing are not the only goals of DEOS. In order to create new services and expand the market, it is important to be able to deliver NRT data that is still relevant for the end users. This requires fast processing of extremely large datasets, which is of course one of the central challenges of the big data community. DEOS attempts to tackle this issue with a twofold strategy. First, DEOS will develop novel strategies to apply approximate computing on a specific set of EO algorithms to speedup their processing without significant lose of accuracy. Second, DEOS will explore how new hardware technologies can be leveraged to speedup EO data processing. The existing research work in these two topics applied to EO data processing is limited to a few studies, therefore **this can be considered exploratory research with an important potential of delivering groundbreaking results.**

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

2. Impact

The applications for NRT-EO services are so broad and wide that the impact of a project like DEOS is hard to predict accurately. Given that this is an emerging market, the greatest applications of this technology are likely to be those that noone has imagine yet. Nonetheless, there is a great potential of a positive transformation for the environment and the economy, both at medium and long term. DEOS aims to be the enabler of this transformation by putting together an outstanding team with great experience in different research and industry sectors. The extended impact of DEOS and the measures to maximize it are presented as follows.

2.1 Expected impacts

DEOS aims to develop the most sustainable decentralized platform for NRT-EO. To achieve such ambitious objective DEOS has to go beyond the state-of-the-art and provide new technologies. In particular, DEOS will propose a set of "*Powerful (Big) Data processing tools and methods that demonstrate their applicability in real-world settings*", more specifically in the context of Earth Observation and Environment Monitoring.

The Qarnot platform currently has up to 5,000 cores that process hundreds of Gigabytes per day (estimation over the second semester of 2016). Compared to a classical server, the Qarnot heater lowers the carbon footprint of computations by up to 75%, which is an incredibly high efficiency in terms of energy. The heaters are geo-distributed within 7 sites (i.e., each site refers to a building) over the city of Paris. By 2018, Qarnot intends to deploy 60,000 cores in up to 10 European cities. With this computing power, the platform can process several Terabytes of data per day while continuing to deliver NRT services. By promoting and implementing environmental data to be processed in this platform, **DEOS will strengthen competitiveness of this emerging market of green cloud computing** and helping the growth of European companies.

In addition, DEOS will design and implement optimised software schemes to leverage new hardware devices for big data processing. We plan to target several European products and the demonstration of our NRT EO services will **promote the use of such devices within the Big Data community, hence promoting growth of the European IT sector (See Section 6)**.

DEOS impact not only extends to the domain of big data processing. DEOS will provide valuable insights on the hydrological extremes and the atmosphere conditions during multiple years at extremely high resolution and high accuracy. This will allow climate scientists to get a better understanding about several important climate variables such as refined knowledge about the carbon cycle, which is critical to improve current computational climate models. This project output aligns well with the European roadmap to limit the extent of global warming in the next

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

decade. But DEOS not only helps to understand the impact of CO2 emissions on the climate, it does so in a very responsible way by implementing the most sustainable platform for big data computing in the world. In other words, **DEOS impact on the environment is twofold, it helps humanity better understand the climate and it provides a role model of sustainable behavior to the community.**

2.2 Measures to maximise impact

As mentioned earlier, potential applications vary from wildfire prediction, to cargo tracking systems, passing by crop growth analysis and air quality control touching many different industry sectors, from biodiversity preservation to national security. Therefore, one of the main focus of the consortium will be to guarantee that the services developed during the project are adapted to the different sector requirements and land to the market converted into a socio-economic benefit for society.

The DEOS consortium is deeply convinced of the importance of communicating science and explaining the importance of our work. In particular, it is critical to inspire the young generations to take over our efforts in the future. **DEOS will take extensive actions to maximise the reach of our results by implementing creative ways to communicate our research to the general public (See Exordium)**.

2.2.1 Dissemination and exploitation of results

Dissemination and exploitation activities within DEOS aim to build a viable context for the future operation guaranteeing its sustainability and profitability.

Dissemination of project results

Project dissemination will aim to share research findings with scientific and engineering stakeholders, draw the technical breakthroughs to the attention of industry stakeholders, raise awareness among policy makers, application developers and end-users or potential users of DEOS' outcome, and position DEOS in the forefront of this area. Research results will be published in selected journals and presented at targeted international conferences and symposia on each of the research and industrial areas. Partners will also participate in locally organised workshops and seminars to present their results. The project will also seek to collaborate with other related projects in this research area in order to cross-fertilise results and maximise the impact of dissemination activities.

The project website will serve as a central hub to keep the DEOS' community updated on project results, with papers being published here as well as information about participation in

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

conferences and dissemination material. A suite of basic communication tools will also be provided to ensure that all partners can easily disseminate project results (See Section 2.2.2).

Exploitation of project results

The exploitation plan will vary depending on the specific layer and if the subject matter is of Public Interest or Commercial interest. DEOS will evolve into a self-sustainable business unit with a "Fast Service Development" approach following these assumptions:

- DEOS will develop a service offered to public agencies and to interested enterprises.
- Basic services, such as downloading Historic data, will be offered for free upon registration.
- Users willing to upgrade the basic services to more sophisticated ones, like NRT Forecast, could join the community becoming customers upon a fee. They will be able to personalise their request with the accuracy and spatio-temporal resolution.
- This fee will be calculated based on the ROI delivered to customers, the revenue stream generated and the operational cost incurred. The membership fee will be market-viable.

DEOS will offer smart automatic analysis of high-resolution land, atmospheric and ocean data at the European scale with the following main services:

- Early warning of flood/drought events & wildfire prevention: to provide daily soil moisture maps at 100-m spatial resolution over Europe and exploit this information to 1) set up an alert and monitoring system in case of hydrological extremes (drought and floods) and 2) identify extremely dry soil and vegetation conditions posing a risk of fire.
- Predictions of air quality: to estimate atmospheric profiles of ECV and trace gases.
- Sea ice diagnostics: for forecast verification
- Land and Offshore Wind farms traders: create a solid link between climate and energy research by co-designing a climate service for network management that will enable the industry to use reliable predictions (or by extension to Offshore wind farms, Satellite Observations) of temperature and wind speed at sub-seasonal to seasonal time scales. Outcomes are expected to provide better scientific understanding and guidance enabling the players concerned (e.g. energy network operators, regulatory authorities) to frame strategic choices concerning cost-effective network management. Network operators will as a consequence be able to minimise the risk of an unbalanced energy system requiring costly decisions to be taken, and the energy industry will grow and adapt in an efficient and flexible manner as its landscape continues to evolve.

Due to the structure of our proposed Work Implementation Plan, each work package is associated with a software layer and a partner specialised in that domain. The consortium partners have already identified their intentions for using and exploiting the results of the project:

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Partner	Exploitation intentions/plans
BSC	An important point for BSC is technology transfer with industry. Some of the BSC developments in this project will be used inside in other in-development platforms, increasing BSC exposition to industrial relevant problems. In particular, development on the usage of hardware resources for big data processing will be integrated in open source libraries that will be offered for the benefit of the entire HPC and big data community. Services that target a more specific industry sector might be transformed into products upon a fee if a viable market is found during the development of the project.
IBM	IBM will exploit the development of DEOS within the OpenPOWER foundation. This implies that the most effective algorithms developed in the project will enter in future POWER math libraries. This will be immediately available in an open form to all the member of the OpenPOWER foundation with a clear impact for data and computing centers, cloud computing infrastructures, research centers and SME.
UVEG	The Image Processing Lab at UVEG will exploit the activities developed within DEOS to further collaboration with the European Space Agency (ESA), the European Center for Medium-Range Weather Forecast (ECMWF) and the European Organisation for Exploitation of Meteorological Satellites (EUMETSAT). The algorithms developed in the project to fully exploit new Copernicus Sentinel data could be set in operations at a research center (e.g. SMOS BEC, BSC) or through an SME to provide added-value EO services in near real time upon a fee. Also, the data set of Essential Climate Variables retrieved during the project will be exploited to deliver valuable knowledge of present water cycle changes and land-atmosphere-ocean interactions. Results will be published in high-impact scientific journals and presented at international conferences.
QARNOT	As a direct result from the DEOS project, NRT EO commercial services using Qarnot's grid will generate fees. Indirectly, the work done by Qarnot regarding smart distributed storage and intelligent scheduling adapted to big data will be leveraged

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

	to extend Qarnot's markets, to computing tasks with higher requirements. These developments are also important to improve Qarnot heterogeneous grid usage and therefore increase overall level of grid direct commercial use.

Knowledge and IPR Management

For an effective exploitation of the project results it is necessary to establish already at this proposal stage how the Intellectual Property (IP) is going to be managed. It will be done in full respect of the **DEOS Consortium Agreement**, which will be finalised, signed and enforced before the project start. It will regulate IP both during and after the project. It aims to protect the interests of each partner, allow good cooperation, and appropriate access. Intellectual Property Right (IPR) management is based on the following principles:

- Background: Each partner owns the background that it brings to the project.
- Patents: The Innovation Manager (See Section 3.2.1) will check the compliance of the partners with the IP process and support the Partners on questions concerning patents. The process will be that any partner who wants to file a patent has to refer to the Innovation manager, who will determine if it is a joint foreground or not. In the case of a joint foreground, the General Assembly will decide on the allocation of intellectual property and exploitation modes. In case of conflict, there will be a vote by a majority of two-thirds (2/3) where the Coordinator will have the casting vote.
- Foreground: generated by only one partner. Foreground shall be the property of the partner carrying out the work generating such foreground.
- Joint foreground: Where the generated foreground is the foreground of several partners, the partners concerned shall have joint ownership of such foreground, according to the proportion of their intellectual, human, material and financial contributions unless they establish an agreement regarding the allocation of property rights relating to it and the terms of exercising that joint ownership.
- Access Rights: For the sole purpose of implementing the project, the right to use a partner's background shall be granted to the other partners, if it is needed to enable those partners to carry out their own part of the work. Such use rights shall not be assignable or exclusive. They shall not be subject to sub-licensing and shall be granted on a royalty-free basis.
- Data Management: It will be carefully considered how research data is handled, shared and curated paying special attention to confidentiality. We will elaborate a Data Management Plan (DMP) (See Section 3.1.4 WP2) to describe the data management life cycle for all data sets that will be collected, processed or generated by the research project. It is a document outlining how research data will be handled during a research

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

project, and even after the project is completed, describing what data will be collected, processed or generated and following what methodology and standards, whether and how this data will be shared and/or made open, and how it will be curated and preserved.

The Innovation Manager will be in charge of all issues related to pre-existing know-how, management of knowledge, protection and publication of the project results in coordination with the leaders of the Exploitation and Dissemination WP (WP2). He/she will produce a management deliverable explaining the IPR plans in the project, and be in charge of managing IPR activities arising from the Project, such as patents, in full respect of intellectual property rights clauses defined in the European Commission model contract and the DEOS Consortium Agreement. The Innovation Manager will count on external advisors (legal service and market research consultant).

The Innovation Manager will also lead an End-Users group (EUG) which will be an advisory body made up of academic and end-users of the DEOS technology from areas such as climate, meteorology and oceanography. The EUG will review the project plan and suggest possible additions to better align the project with the needs of user communities. It will also assist in directing the work of the project to ensure the compatibility of the technology planned and developed with user requirements. The EUG will also identify and establish contact with other potential end-users. It will interact with end users in relation to future opportunities and commercial leads. It will also target such users when disseminating project results. The benefits of implementing The EUG are twofold:

- Project members will be able to adapt the project outcome
- End-Users Group members will have early access to the technologies of the project

With regards to publications, the consortium is committed to provide at least “green” open access wherever feasible. We have reviewed the provisions of “*The Guidelines on Open Access to Scientific Publications and Research Data*” in Horizon 2020 and defined a strategy for knowledge management and protection. Green open-access is also known as self-archiving and means that authors deposit a preprint or a (potentially revised) author version of their publication at an institutional or subject repository that allows public access. If permitted, this may also be the publisher's version of record.

All candidate publications will be circulated among all partners for approval 30 days before the due publication date. Each partner will have the right to suggest changes or removal of confidential information. Conflicts will be addressed by the conflict resolution process to be defined to that end. In case a partner has exploitable knowledge that they do not want to exploit, they may make it available to the consortium and will notify the EU of such situation.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

2.2.2 Communication activities

Effective communication is essential to ensure that the project reaches its intended beneficiaries and to promote the value of European research to key stakeholders and to the wider public. To achieve this, we must define a clear dissemination and communication strategy at the outset of the project by identifying target audiences and selecting the appropriate means of communicating with each of them, as shown in the table below:

Target audience for communication activities.

Target audience	Key messages	Value to target audience
Applications developers	Know-how and advice to maximize the efficacy and efficiency of applications in novel architectures.	Help them improve their applications and the post-processing of the generated results.
HPC and Big Data research community	Novel schemes and techniques on approximate computing and efficient use of novel hardware devices.	Improve the efficiency of their systems in order to increase global machine's throughput.
Industry stakeholders	Feedback about their technologies and guidelines for further improvement.	Information about market expectations and opportunities for product enhancement.
Policy makers	Status of the environment and efficacy of near real time satellite data processing.	Guidelines for future investment and weakness identification for future consolidation.
General public	News about advancement on satellite data processing and a better IT understanding.	Provides new business opportunities and ideas to leverage at all levels.

Moreover, the DEOS consortium has well-established partner communication channels and the coordinator has an in-house Communication and Technology Transfer unit at BSC with wide range of excellent contacts in industry and research, upon which it will draw in order to

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

communicate project benefits. Communication activities will seek to: (1) Raise awareness about the project and its results, (2) update key stakeholders on project progress and (3) build community and encourage participation.

Overall, DEOS communication strategy will consider the following activities:

- *Branding:* A distinctive brand identity will be created for the project, represented by a logo and corporate image. This will be applied consistently across all project dissemination materials and templates in order to reinforce awareness of the project.
- *Website and social media:* The public project website will be the first scenario where all project information and results will be presented. It will allow project updates to be communicated to a large and diverse audience. All services and results will also be promoted via potential social media channels. A news section will provide project updates and dissemination material will be available for download, as well as open-source software and services.
- *Digital communication:* Leaflets and factsheets are a useful and attractive way of providing the key facts about the project at a glance, and can be easily distributed at events. Press releases addressed to scientific media can be prepared and released as required to disseminate project results. White papers, webinars are useful to disseminate technical results targeting industry. Forming media partnerships, writing articles for relevant blogs and giving interviews are all ways of giving the project greater visibility and updating the core community on project results. Preferably, the dissemination material will be digital material to be disseminated via social media (infographies, videos, factsheets, etc.)
- T*rainings and organization of one final workshop:* Along the project various trainings will be organized. All results will be presented in key conferences such as SC, ICS, ISC, BigData, among others, and we will organize a final workshop in order to invite all stakeholders involved in DEOS and present all project results. Well known and well established workshops and training held at the partner institutions will include sessions on distributed computing, EO data processing, machine learning and hardware optimization for big data. Workshops are frequently the basis for collaborative application development, e.g. the PRACE training courses done by the PRACE training centers (such as the ones organized by BSC) and other seasonal schools where the partners will have presence (HiPEAC, PUMPS summer school) will be taken into account. In particular, three trainings will be organized both for the consortium partners as well as to potential DEOS users.
- *Newsletter to households and companies heated by Qarnot's computing heaters:* As the DEOS data processing is performed in a distributed computing grid hosted by many customers across European cities, we will inform those Q.rad users about the DEOS project and what is being computed on their heaters. These communications will be

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

performed in a biannual basis and will include a summary of the result of the project and the positive environmental impact of what it was achieved by heating their place. This will raise awareness of the project among the general public and will foster a sentiment of contribution to protect the environment. The communication will be non-technical, so that even children can understand it. Similar newsletters will be online and distributed to a subscribe option on the DEOS website, as well as promoted via social media channels of all project partners. It will also include a Coordinator’s welcome with latest updates of the project.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

3. Implementation

3.1 Work plan - Work packages, deliverables and milestones

3.1.1 Overall strategy

The objectives of the project will be achieved through activities organized in six work packages. These include Project management (WP1) and Dissemination and Exploitation of project results (WP2). The technical work is organized in one transversal work package of continuous integration and evaluation (WP3), and three vertical WPs devoted to different software levels (WP4), (WP5) and (WP6).

- **WP1 Project Management and Coordination** deals with the administrative and financial management of the project, including the monitoring and assessment of the quality of the project outcomes.

- **WP2 Project Dissemination and Exploitation** will take care of all dissemination and exploitation activities concerning the results of the DEOS project. This includes dissemination in the research domain as well as to the general public.

- **WP3 Distributed Operations, Integration and Validation** will be in charge of the distributed computing platform as well as the continuous integration of the outcomes of tasks from other WPs into a permanent service that will be delivered to the end-users throughout the project.

- **WP4 Near Real Time Earth Observation Applications** is responsible for the design, implementation and evaluation of the main services that the DEOS project will deliver. This WP will be in close communication with the End Users (see Section 3.2.1) in order to get feedback and improve the services.

- **WP5 Low- power Approximate Computing and Data Analytics** will lead the efforts on data analytics and machine learning techniques to improve the quality of services. In particular, they will get feedback from the WP4 to know which services and specific sub-modules are numerically robust and can benefit from mixed-precision and other approaches to speed-up the data transfer and processing.

- **WP6 Optimizing EO Processing on New Hardware Architectures** is responsible for the optimization aspects on the new hardware technologies. For instance, most of the mixed-precision results obtained in WP5 will be leveraged in WP6 to then be ported to

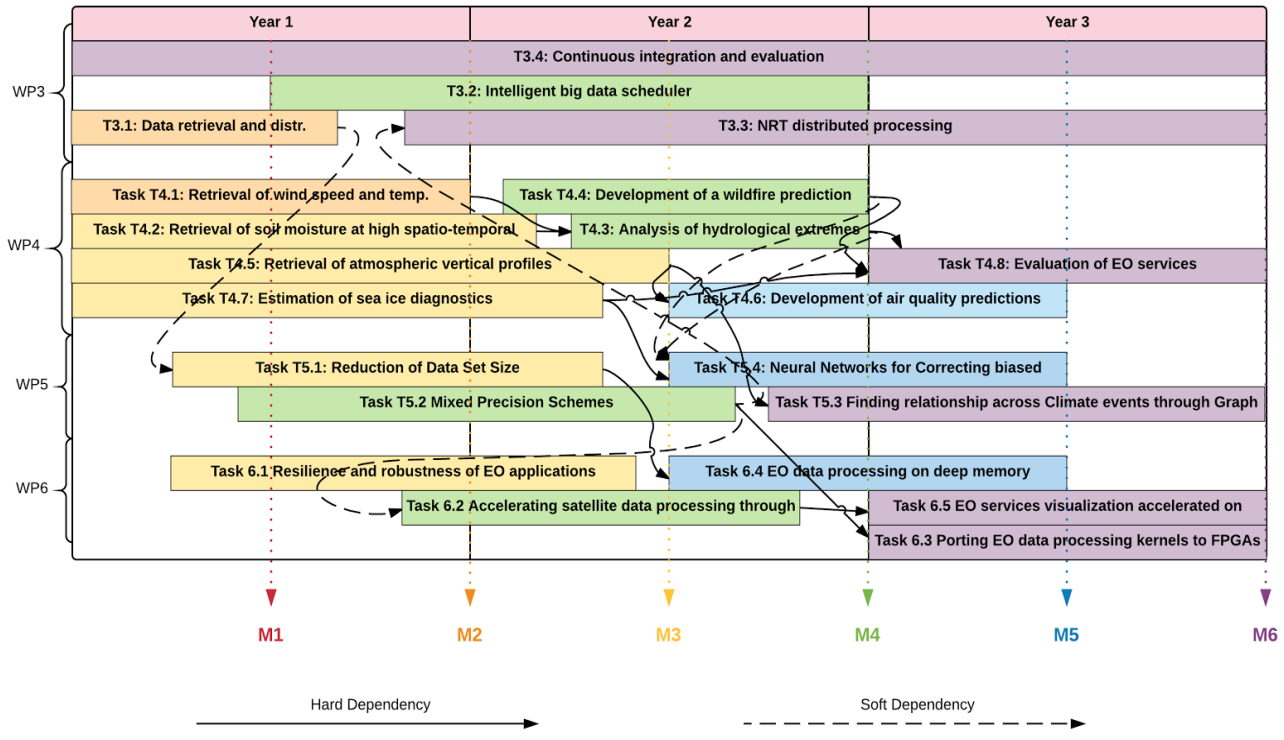
This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

hardware accelerators (such as GPUs) that allow for half-precision. Other optimisations such as leveraging deep-memory hierarchies will be studied and implemented to improve the data processing and deliver near real time services.

3.1.2 PERT Diagram - Interdependencies among Work Packages

The DEOS consortium plans to act as a cohesive team with strong links across the different work packages. This is visible in the task dependencies shown in the PERT diagram below. Tasks are grouped by work package as shown in the left-side of the figure. There are two type of dependencies: *hard dependencies* and *soft dependencies*. Hard dependencies show tasks that cannot start until the results of the preceding task are delivered. Soft dependencies show relationships between two tasks, in which the results of a task can enhance the results of the other task but are not required for the other task to start. Some soft dependencies are shown as going back in time in the figure but this is just to have a clear visualization of the dependencies. In reality, the other task can start, make progress and leverage the results of its dependency when they become available without causing any type of delay.

PERT Diagram showing Tasks Dependencies across WPs



This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

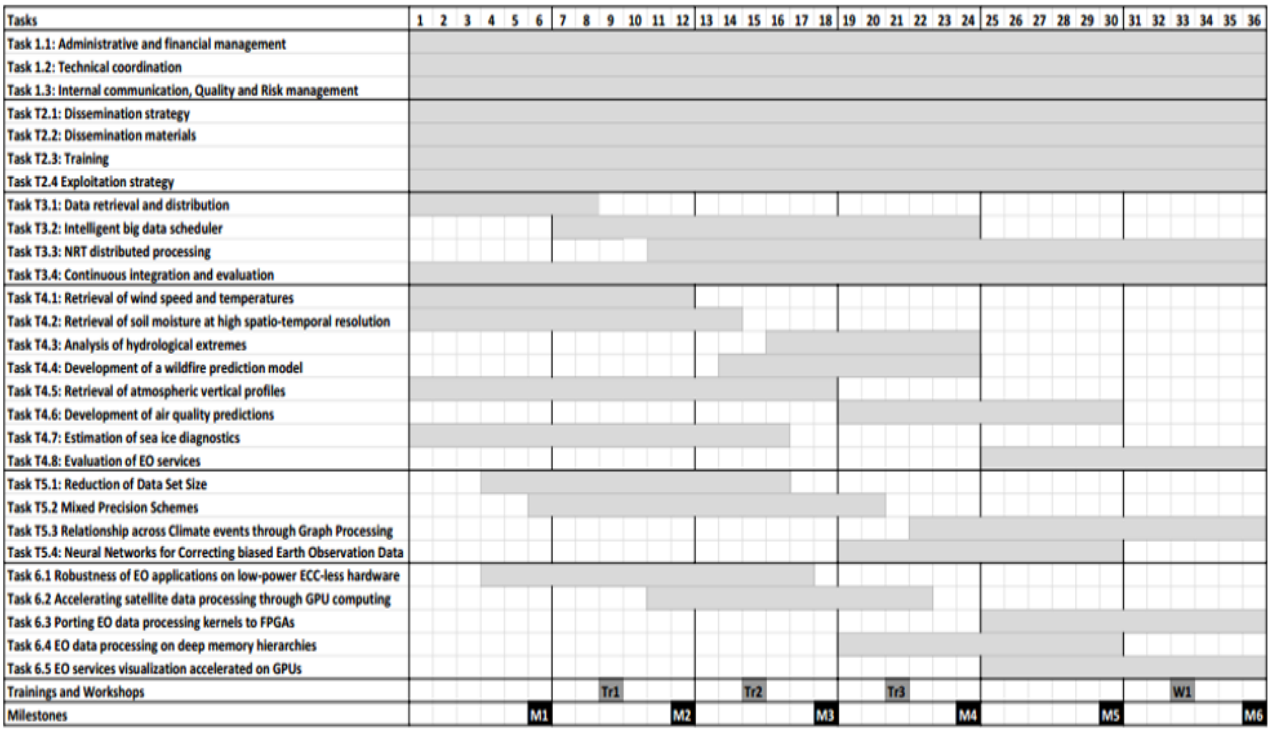
The figure also show the six milestones of the project (See Section 3.2.3). In addition, to facilitate the understanding of the link between tasks and milestones we have used a color code, in which tasks that are required to reach a milestone are colored in the same color as the targeted milestone. Please note that the first milestone (i.e., red) is an internal milestone to have a working canvas and is related to deliverable D4.1 (See Section 3.1.6).

Finally, please note that tasks generally do not start just after the end of a previous task, but there are one or two month spaces in between. This is to adjust for possible short delays that some tasks could potentially experience. When tasks are completed in time, and if all hard dependencies have been delivered, then some tasks might start one or two months in advance to achieve early results and to mitigate potential short delays in future tasks.

3.1.3 Timing of the different work packages and their components

The Gantt diagram below shows the timing of the tasks of the different work packages and the milestones, as well as the planned timing for the trainings and the final workshop. Most tasks are at least 12 months long, with the exception of a couple tasks that can be executed faster.

Gantt Diagram of DEOS Work Packages, Milestones and Dissemination Events



Work packages 5 and 6 will start after 3 months, when the dataset and *toy* application are already deployed in the distributed platform. Task 3.4 is responsible for the continuos integration

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

of the results of the other work packages in order to enhance the EO services that we aim to provide. In addition, most services should be deployed by the end of the second year, leaving the the third year for a complete and thorough cross verification and validation process.

As it is shown in the diagram, the organization of the tasks put more weight on WP3 and WP4 during the first 18 months of the project, and more weight on WP5 and WP6 during the second half of the project. This is because the first part is about delivering a canvas of the EO services that we want to provide, while the second part of the project is more focused on optimizing them to make them run faster using software and hardware techniques.

3.1.4 Detailed work description

Work package number	1	**Lead beneficiary**		BSC
Work package title	Project Management and Coordination			
Participant number	1	2	3	4	
Short name of participant	BSC	QARNOT	UVEG	IBM	
Person months per participant:	18	1	1	2	
Start month	1		**End month**	36	

Objectives

This work package deals with the overall DEOS management; more specifically, it has the following objectives:

- To deliver on the technical objectives of the project within the time and budget constraints of the project.
- To ensure that there is clear and effective communication between partners; to detect management and technical issues as early as possible and bring them to resolution.
- To establish and enforce effective management and quality procedures that will result in high quality project deliverables.
- To provide efficient operational management support including: administrative and financial planning, reporting to the EC, management of project legal aspects including project-related contracts and IPR, and management of day-to-day operational and technical progress.

Description of work

The following is a list of the tasks required to achieve the objectives of this work package. The high level Management Structure as well as the individual roles and responsibilities within this structure are explained in Section 3.2. It also includes a brief overview of the most important management procedures of the project which will be further defined in the early months.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Task 1.1: Administrative and financial management (M01-M36) [Leader: BSC; Contri: ALL]

Lead by the project manager, this task will establish the corresponding procedures, tools and methodologies to enable a correct project management, including administrative and financial management. It will also coordinate the timely production of deliverables, organize the kick-off meeting and reviews, and organize and manage audits requested by the commission. On a 6-month basis, the project coordinator will monitor resources usage, producing internal use of resources reports to ensure the project resources expenditure is in track with the work progress.

Task 1.2: Technical coordination (M01-M36) [Leader: BSC; Contributors: ALL]

Lead by the technical manager, this task will perform the technical co-ordination of the project, by means of monitoring the progress of the work packages, technical co-ordination of the meetings, appointing reviewers to assess the quality of the deliverables before their delivery to the EC, and solving technical conflicts.

Task 1.3: Internal communication, Quality and Risk management (M01-M36) [Leader: BSC; Contributors: ALL]

In this task, we will determine the appropriate strategy to ensure clear communication channels between all partners in order to facilitate the exchange of critical project documentation and news and to encourage participation in the decision-making process. The task will require defining and maintaining internal collaborative tools for sharing documentation and communicating work status. One of the outcomes of this task will be a Project Portal (shared workspace) and a series of Distribution Lists.

In this task, we will also define and implement the appropriate quality assurance processes that ensure accurate documentation, reporting and justification of the work being carried out. A process will be developed to ensure that the deliverables have been reviewed by a broad spectrum of individuals against a well-defined set of criteria. Moreover, we will determine the minimum level of quality required for presentation of the official outcomes of the project to the EC. The high level principles guiding these procedures will be agreed to at the start of the project at the Kick-off Meeting. The administrative project management procedures defined in Task 1.1, quality assurance and risk management processes will be documented in D1.1.

Deliverables

D1.1 Project Management and Quality Guidelines (M03) [Responsible: BSC]

This deliverable will describe the project's internal management procedures, detailing the project's Quality assurance process as well as a detailed Risk evaluation and internal communication tools and mechanisms.

D1.2: First Year Periodic report (M12) [Responsible: BSC]

This deliverable will include a technical report describing the project activities performed during the first reporting period, and a financial report containing the financial statements and explanation

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

of the use of resources per beneficiary for the period.

D1.3: Second Year Periodic report (M24) [Responsible: BSC]

This deliverable will include a technical report describing the project activities performed during the second reporting period, and a financial report containing the financial statements and explanation of the use of resources per beneficiary for the period.

D1.4: Final Periodic report (M36) [Responsible: BSC]

This deliverable will include a technical report describing the project activities performed during the third reporting period, and a financial report containing the financial statements and explanation of the use of resources per beneficiary for the period.

Work package number	2	**Lead beneficiary**			BSC
Work package title	Project Dissemination and Exploitation				
Participant number	1	2	3	4	
Short name of participant	**BSC**	QARNOT	UVEG	IBM	
Person months per participant:	18	3	7	3	
Start month	1		**End month**	36	

Objectives

- Disseminate the project ideas and results via a public website, considering technical, educational, cultural and commercial aspects.
- Participate to and present articles to scientific conferences to publish the project results and possible products.
- Build and support a DEOS community to facilitate acceptance and make European industry aware of the DEOS potential.
- To facilitate cross-fertilisation with other projects working on overlapping areas
- Identify potential end-users and customers and engaging them for the project results
- Market analysis to adapt our services to the needs and identify competitors

Description of work

The dissemination and exploitation working group will liaise closely with other work packages, as well as maintain close contact with the End-Users Group in order to provide up-to-date information on the achievements of the DEOS project, and to disseminate this to the appropriate audiences in each case. With the same objective, this information collected and the feedback obtained from the

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

stakeholders will be used to redefine the exploitation management plan and establish commercial and non-commercial relation with them.

Task T2.1: Dissemination strategy (M1-M36) [Leader: BSC; Contributor: All]

In this task, the team will define the objectives of the communication plan, finding adequate channels and messages to target audience, and define its activities or tools to achieve these objectives. The outcome of this task will be the Dissemination and Communication Plan (D2.1). This document will be regularly updated during the project's lifetime.

The team will carry out the activities detailed in the dissemination plan, such as:

• A public-facing website for external communication. This will provide updates on project progress as well as provide the open data available to the defined audiences, as well as project documentation and scientific publications. It will provide information about events and lectures in which the project will be participating. Furthermore, it will show a set of open-source available data that will be opened on a login-based requirements in order to obtain information about the possible target audiences and end-users.

• Branded templates for project presentations and documents to ensure a consistent image and reinforce awareness of the project.

• Dissemination materials such as factsheets, leaflets or posters, depending on the target audience and event as well as trial demos of the available data.

• Press releases, which allow key messages and project news to be highlighted to the scientific and general press, both national and international. Partners can assist in raising awareness about the project by translating press releases and distributing them to national and local media.

• Audiovisual media: several videos will be produced to provide an engaging and informative means of communicating the project's results to various identified target audiences.

All communication activities and tasks will be carefully monitored and quality metrics will be reported in the Dissemination Reports. Quantitative indicators may include:

- Website statistics
- Social media statistics, such as the number of participants in a LinkedIn group.
- Number of press articles in national and international media as well as scientific press.
- Number of press releases.
- Number of subscribers to the newsletters.
- Number of events in which the project has participated.
- Number of lectures given.
- Number of visits to potential end-users.
- Number of app users.

Acknowledgement of the EC funding sources will be included in all dissemination materials.

Task T2.2: Dissemination materials (M01-M36) [Lead: BSC; Contributor: All]

Dissemination materials such as factsheets, leaflets or posters, depending on the target audience will be developed. This material will be distributed at events, exhibitions or conferences attended

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

by the partners. The focus will be mainly to disseminate the project on digital communication media such as LinkedIn group by creating articles, as well as developing an online newsletter.

Task T2.3: Training (M01-M36) [Lead: BSC; Contributor: All]

This task will include the organization and logistical support of training courses which will be organized three times during the project. The training plan will be included in the Dissemination and Exploitation Plan (D2.1), as well as identifying the training needs of the various researchers during the project, or any other defined targets from particular industrial sectors. Training material should be uploaded onto the intranet.

Task T2.4: Exploitation strategy (M01-M36) [Lead: BSC; Contributor: All]

This task involves all the activities related to approach the market and guarantee the profitability and sustainability of the project following the above described "Fast Service Development" strategy. The Innovation Manager will be in charge of identifying potential users and customers for the services offered and manage the pool of early adopters in order to align continuously the project with the needs of user communities and manage the End-Users group. IPR and Data Management will be key in this task to protect the results while exploiting and sharing them (D2.2 and D2.3).

Deliverables

D2.1) Dissemination and Communication Plan (M03) [Responsible: BSC]

This deliverable should include the dissemination strategy defining the objectives, different targets and channels, communication activities in order to ensure a societal impact and correct exploitation of the project's results.

D2.2) Exploitation Plan (M06) [Responsible: BSC]

This deliverable, in close relation with the Dissemination plan, will provide a solid base for exploitation actions within the project lifespan and after. It will be a first approach to the strategy being updatable with the input from the end users and other market analysis activities as well as the project´s results. This deliverable will include possible conclusions and results from the Advisory group. It will be evolving in time and it will contain: i.Analysis of the context: identify the early adopters and analyse their characteristics and requirements ii.Validation of Hypothesis, iii.End User Meeting feedback, iv.Conclusions and suggestions and v.End-users database.

D2.3) Data Management Plan (M06) [Responsible: BSC]

The DMP will describe the life cycle for all data sets that will be collected, processed or generated by the research project. It is a document outlining how research data will be handled during a research project, and even after the project is completed, describing what data will be collected, processed or generated and following what methodology and standards, whether and how this data will be shared and/or made open, and how it will be curated and preserved. The DMP is not a fixed document; it evolves and gains more precision and substance during the lifespan of the project. It

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

should include the following information: i.Description of Data, ii.Data Collection/Generation, iii.Data management: documentation & Metadata, iv.Intellectual Property Rights and v.Accessibility: Data sharing, archiving and preservation.

D2.4) First Dissemination and Exploitation Report (M12) [Responsible: BSC]

The annual reports will summarize all dissemination and exploitation activities realized during the period 1 of the project. It will also have to include analysis of all activities and its evaluation, if possible. This report will analyse how the first minimal viable service (MVS) is getting into the market and the feedback from the early adopters. The last report will give a broader overview of the project and focus more on its results and exploitation/impact.

D2.5) Second Dissemination and Exploitation Report (M24) [Responsible: BSC]

The annual reports will summarize all dissemination and exploitation activities realized during the period 2 of the project. In particular, this deliverable will analyse how the three MVS are being accepted by the early adopters and how to intensify the feedback loop. It will also have to include analysis of all activities and its evaluation, if possible. The last report will give a broader overview of the project and focus more on its results and exploitation/impact.

D2.6) Third Dissemination and Exploitation Report (M36) [Responsible: BSC]

The annual reports will summarize all dissemination and exploitation activities realized during the period 3 of the project. It will also have to include analysis of all activities and its evaluation, if possible. The last report will give a broader overview of the project and focus more on its results and exploitation/impact.

Work package number	3	**Lead beneficiary**			QARNOT
Work package title	Distributed Operations, Integration and Validation				
Participant number	1	2	3	4	
Short name of participant	BSC	QARNOT	UVEG	IBM	
Person months per participant:	6	60	3	1	
Start month	1		**End month**	36	

Objectives

This work package aims to have a global vision of the distributed computing platform and manage the retrieval of the datasets as well as scheduling of the computing tasks. The objectives are the following ones:

- To implement a data retrieval and efficient storage system for large satellite datasets and

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

prepare them to be processed in a distributed infrastructure.
- To design and develop a scheduling algorithm taking into account the service requirements and the data volumes in order to deliver NRT services.
- To oversight the continuous integration of all the outputs of the project and evaluate them in the production platform.

Description of work

Task T3.1: Data retrieval and distribution (M1-M08) [Lead: QARNOT; Contributor: UVEG]

This task's objective is to manage huge data quantities efficiently for distributed processing. The data will be retrieved from the Copernicus data servers and will next be distributed to the geo-distributed storage of the Qarnot platform. Qarnot is currently adapting the Ceph system to the management of its storage nodes. The main challenge will consist in developing prefetching policies to anticipate situations in which compute nodes could wait for data in order to continue a processing. For this purpose, locality-aware policies specialized to the processing of Earth observations must be proposed. The policies will guarantee that on compute nodes, the latency to get the input data is enough to expect the output in NRT.

Task T3.2: Intelligent big data scheduler (M07-M24) [Lead: QARNOT; Contributor: IBM]

Processing satellite data and delivering near real time services on a distributed grid requires a good management of priorities in order to guarantee some quality of service. We will investigate and implement scheduling algorithms that guarantee that services are delivered in a near real time fashion, even in the context of an earth observation service. For this purpose, Qarnot intends to make its scheduler data-aware. This means for instance to deploy on each compute node the first tasks that could already be done, depending on the data available from the first storage, close to the compute node.

Task T3.3: NRT distributed processing (M11-M36) [Lead: QARNOT; Contributor: BSC]

In order to ensure NRT service delivery, the project will develop approximate computing algorithms for Eo processing. These parameterized algorithms adapt the precision of the output depending on a walltime, allowed for the processing. The goal of this task is to develop an intelligent adapter that will constantly calibrate the approximate computing algorithms in order to choose the *best* compromise between the precision and the Near real time delivery of outputs. These adapters will be deployed on the compute nodes in order to constantly adapt to walltime objectives, decided in the processing.

Task T3.4: Continuous integration and evaluation (M01-M36) [Lead: QARNOT; Contrib: All]

Qarnot will be responsible for the integration of all the other developments in the project and their evaluation in the production platform. Given that DEOS will undertake a *Fast Service Development* strategy (See Section 1.3.2) and it aims to produce a *minimum viable service* within the first six

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

months, the integration work will start from the first month of the project. This contrast with other project where the standard procedure is to start the integration after two years in which the partners develop their tasks on their side in an almost independent fashion.

Deliverables

D3.1) Large dataset management tool for distributed architecture (M12) [Resp.: QARNOT]
This deliverable will be a software tool able to manage large dataset within distributed computing perspective, retrieve dataset from Copernicus, send it for processing, and gather results for backup in the appropriate location. A report that explain the software will also be produced.

D3.2) Big data adapted smart scheduler (M24) [Responsible: QARNOT]
This deliverable is a documented software which purpose is to distribute processing job smartly and adaptively to the Qarnot grid infrastructure.

D3.3) Resource adaptative NRT processing (M36) [Responsible: QARNOT]
This deliverable is a documented software algorithm able to adapt computing process in near real time.

D3.4) Grid evaluation tool (M30) [Responsible: QARNOT]
This deliverable is a software tool to monitor grid performance for continuous integration and evaluation purposes.

Work package number	4	**Lead beneficiary**		UVEG
Work package title	Near Real Time Earth Observation Applications			
Participant number	1	2	3	4	
Short name of participant	BSC	QARNOT	UVEG	IBM	
Person months per participant:	18	0	70	0	
Start month	1		**End month**	36	

Objectives

The main goal of this work package is to develop the application services that will ultimately deliver valuable knowledge to the society and industry users. We will focus on a group of services

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

(See Section 1.3.1) and make sure that the output of those applications are within the accuracy parameters expected by the end users and that the results are relevant. Here is the breakdown of those objectives:

- Detection of hydrologic extremes -floods and droughts- over Europe based on satellite-based soil moisture retrievals with unprecedented accuracy and spatio-temporal resolution.
- Provide a European-wide risk of wildfires based on anomalies and extremes of soil moisture and temperature and wind speed predictions.
- Provide more accurate predictions of air quality over Europe based on satellite-derived atmospheric profiles of trace gases.
- Provide sea ice diagnostics (extent, volume and area) based on high resolution spaceborne observations.

Description of work

Task T4.1: Retrieval of wind speed and temperatures (M01-M12) [Lead: BSC; Cont.: UVEG]

This task uses the ten-metre wind speed forecasts from the ECMWF(European Centre for Medium-Range Weather Forecasts) System 4[21] operational seasonal prediction system. It consists of running ECMWF 4 system off-line for the period 2010-present to extract wind speed and temperatures meteorological forecasts and perform a spatial kriging to 1x1 km on Qarnot.

Task T4.2: Retrieval of soil moisture at high spatio-temporal resolution (M01-M14) [Lead: UVEG; Contributor: BSC]

Present operations at BEC provide daily soil moisture maps at 1 km spatial resolution over the Iberian Peninsula (IP) from the optimal blend of 25-km SMOS and 1-km MODIS NDVI/LST data using the algorithm in (Piles et al., 2014)[22]. Both coverage and spatial resolution of BEC fine-scale soil moisture estimates are limited by the use of MODIS data, which is only available globally from NASA with a 15-day latency. This task aims at running on Qarnot a tailored version of BEC downscaling algorithm to provide daily soil moisture maps at 100-m spatial resolution over Europe with the use of SMOS, Sentinel 2 NDVI and Sentinel 3 LST data, which are available in NRT over the whole European continent. Additionally, the presence of clouds masking LST measurements will be mitigated by the combination of Sentinel 3 with 15-min repeat geostationary data from MSG SEVIRI data. Maps will be generated from June 2015 (launch of Sentinel-2) to present. Techniques for data dimensionality reduction as well as for fast processing of big data sets will be

[21] Molteni, F., Stockdale, T., Balmaseda, M., Balsamo, G., Buizza, R., Ferranti, L., ... & Vitart, F. (2011). *The new ECMWF seasonal forecast system (System 4)* (p. 49). European Centre for Medium-Range Weather Forecasts.

[22] **Piles, M**., Sánchez, N., Vall-llossera, M., Camps, A., Martínez-Fernández, J., Martínez, J., & González-Gambau, V. (2014). A downscaling approach for SMOS land observations: Evaluation of high-resolution soil moisture maps over the Iberian Peninsula. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(9), 3845-3857.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

implemented to ensure real time capabilities.

Task T4.3: Analysis of hydrological extremes (M16-M24) [Lead: UVEG; Contributor: BSC]

This task will implement the automatic detection of anomalies on high resolution (100-m) soil moisture maps using machine learning approaches (e.g. convolutional neural network). Anomalies will be computed using as a reference a SMOS-based soil moisture climatology for years 2010-present at 25 km spatial resolution.

Task T4.4: Development of a wildfire prediction model (M14-M24) [Lead: UVEG; Contributor: BSC]

This task aims at providing maps of wildfire risk based on anomalies and extremes of soil moisture and temperature and wind speed predictions. Causal inference techniques will be applied to explore the relationship between pre-fire conditions of soil moisture, temperature, wind, and wildfire properties such as extension and intensity. Results will be used to upgrade the forest fire modelling framework first presented in (Chaparro et al., 2016)[23] and provide a fire risk assessment in near real time.

Task T4.5: Retrieval of atmospheric vertical profiles (M01-M18) [Lead: UVEG; Contri: BSC]

The objective of this task is to provide enhanced air quality retrievals using satellite-based retrievals of atmospheric profiles of ECV and trace gases, such as temperature, water vapor, nitrogen dioxide (N_2O), methane (CH_4), carbon monoxide (CO), carbon dioxide (CO_2) and ozone (O_3) concentration. Retrievals will be provided with unprecedented speed, accuracy and vertical resolution using a tailored version of the algorithm first presented in (Camps-Valls, 2012).

Task T4.6: Development of air quality predictions (M19-M30) [Lead: BSC; Contrib.: UVEG]

CALIOPE is an operational air quality system developed in the Earth Science department and daily operated in the Mare Nostrum supercomputer. The system provides air quality forecast at 24h and 48h for Europe (12x12 km), the Iberian Peninsula (4x4 km) and Canary Islands (2x2 km)**.** This task aims at improving the present accuracy of CALIOPE air quality predictions by the assimilation of satellite-based profiles of trace gases and EIONET stations using an adaptive kalman filter routine running on Qarnot.

Task T4.7: Estimation of sea ice diagnostics (M01-M16) [Lead: BSC; Contributor: UVEG]

The focus of this task is to compute sea ice diagnostics from high resolution Sentinel 1 data. From the sea ice concentration and sea ice thickness fields provided by the satellites, the sea ice volume, sea ice area and sea ice extent will be computed. These new "observational" variables can be compared to the outputs of the model EC-Earth and help in the process of the forecast verification.

[23] Chaparro, D., **Piles, M.**, Vall-llossera, M., & Camps, A. (2016). Surface moisture and temperature trends anticipate drought conditions linked to wildfire activity in the Iberian Peninsula. *European Journal of Remote Sensing*, *49*, 955-971.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Task T4.8: Evaluation of EO services (M25-M36) [Lead: UVEG; Contributor: BSC]

This task will focus on the evaluation of DEOS services and applications through direct engagement with end-users. Incremental improvements in the service will be implemented according to the received feedback.

Deliverables

D4.1) Requirements for data retrieval and distribution (M6) [Responsible: UVEG]

This deliverable will address the generic handling of Copernicus data, including tools for I/O of the products and formats defined within the project, as well handling data from external sources (BEC and BSC).

D4.2) Soil moisture from SMOS-S2/S3/SEVIRI Algorithm Theoretical Baseline Document (M18) [Responsible: UVEG]

The purpose of the *Algorithms Theoretical Baseline Document* (ATBD) is to "describe the algorithms which will produce higher level products. The document should focus on the scientific justification for the algorithms selected to derive the product, an outline of the proposed approach and a listing of the assumptions and limitations of the algorithm". This deliverable consists on an ATBD of the high resolution soil moisture retrievals, including evaluation results from comparison with in-situ data from the International Soil Moisture Network[24].

D4.3) Atmospheric profiles of trace gases from IASI Algorithm Theoretical Baseline Document (M18) [Responsible: UVEG]

This deliverable consists on an ATBD of the atmospheric trace gases retrievals, including an evaluation of the use of different machine learning techniques such as Kernel Ridge Regression or Deep Neural Networks.

D4.4) Sea ice diagnostics (M18) [Responsible: BSC]

This deliverable is a software feature that allows visualizing the status of sea ice (extent, volume, and area) in near real time as well as its evolution in the last decades[25]. Summary sheets of forecast verification will also be provided.

D4.5) Analysis of extreme hydrologic events (M24) [Responsible: UVEG]

This deliverable is a software feature that allows visualizing the status of water in soils in near real time as well as looking at trends and anomalies for the period 2010-present.

[24] https://ismn.geo.tuwien.ac.at/

[25] Mathiot, P., König Beatty, C., Fichefet, T., Goosse, H., Massonnet, F., & Vancoppenolle, M. (2012). Better constraints on the sea-ice state using global sea-ice data assimilation. *Geoscientific Model Development*, *5*(6), 1501-1515.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

D4.6) Prediction of Air quality (M30) [Responsible: BSC]

This deliverable consists on the provision of air quality forecasts by means of concentration maps (Hourly and maximum concentration for 24 and 48 h forecast), and air quality index maps (with tables of maximum values of air quality, planned for 24 and 48h). Summary sheets of the annual forecast (statistical and graphics) as well as forecast evaluation at measuring stations (EIONET) will be provided.

D4.7) Maps of Wildfire risk (M30) [Responsible: UVEG]

This deliverable consists on the provision of maps of wildfire risk in near real time. The accuracy of the model predictions will be assessed for the period 2010-present using the forest fire dataset provided by the European Forest Fire Information System—EFFIS (http://effis.jrc.ec.europa.eu).

D.4.8) Best practices and roadmap for distributed EO services (M36) [Responsible: UVEG]

This deliverable will contain all best practices, lessons learnt, and specifications from DEOS regarding the use of open geospatial data in distributed computer environments.

Work package number	5	**Lead beneficiary**		IBM
Work package title	Low-power Approximate Computing and Data Analytics			
Participant number	1	2	3	4	
Short name of participant	BSC	QARNOT	UVEG	IBM	
Person months per participant:	24	0	9	33	
Start month	4		**End month**	36	

Objectives

This work package aims to explore and design novel techniques of approximate computing for satellite data processing to speedup the computations without hurting the accuracy of the results. This effort can be breakdown in the following objectives:

- To identify opportunities for reducing the size of the datasets that need to be transferred to processing modules like heaters.
- To accelerate the processing modules by using mixed precision mechanisms without hurting the final accuracy of the computations.
- To carry out exploratory studies on more aggressive mechanisms to find relationships across different climate events.
- To use machine learning techniques to better identify related event as well as anomalies.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Description of work

This work package is structured into 4 main tasks, each one of them targeting one of the above presented objectives:

Task T5.1: Reduction of Data Set Size (M04-M16) [Lead: IBM; Contributor: BSC, UVEG]

This task will make use of well known data reduction techniques like Principal Components Analysis (PCA), clustering algorithms like k-means or Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to reduce the dimensionality of the data sets that need to be transferred from the sensors to the processing modules. While techniques like PCA will identify redundancies in data sets dimensionality, DBSCAN or k-means will be used to identify repetitive data points and to select the meaningful ones. These techniques are expected to have a very strong impact in reducing and identifying redundancies on the data sets obtained from satellites.

Task T5.2 Mixed Precision Schemes (M06-M20) [Lead: BSC; Contributor, IBM]

This task will complement the efforts done in Task 5.1 in two ways: first, it will apply mixed precision schemes to data reduction algorithms like PCA or k-means to accelerate the compression process. Second, it will speed up the satellite-data processing by identifying those kernels that are less sensitive to reduced precision schemes. The methodology is going to be two fold: first, the task will explore the possibility of reducing the number of bits of floating point representations; second, it will apply fixed point arithmetic, which is much cheaper and faster than the floating point one, to improve performance while reducing energy.

Task T5.3 Finding relationship across Climate events through Graph Processing (M22-M36) [Lead IBM, Contributor BSC, UVEG]

This task aims at developing automatic techniques to analyze climate data base. It is typically the case that such databases contain events that are related but it is extremely hard to discover such relationships by means of human inspection due to the large size and complexity of such data bases. This task will explore automatic techniques like graph partitioning (if the data based has a graph structure) or other approaches like convolutional neural networks to automatically classify and correlate climate events.

Task T5.4: Neural Networks for Correcting biased Earth Observation Data (M19-M30) [Lead: BSC; Contributor IBM, UVEG]

This task targets the bias correction problem that climate models typically exhibit. Such biases arise from inaccuracies due to limited spatial resolution or truncating errors of numerical schemes. While methods like multiple linear regressions partially mitigate bias correction problems, this task will explore more aggressive methods based on anomaly detection or convolutional neural networks to detect the biased data and correct it.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Deliverables

D5.1) Report on reduction of data set size (M18) [Responsible: IBM]
This deliverable will summarize the results of Task T5.1.

D5.2) Report on mixed-precision schemes (M24) [Responsible: BSC]
This deliverable will summarize the results of Task T5.2.

D5.3) Report on techniques for detecting relationship across climate events (M36) [Responsible: IBM]
This deliverable will summarize the results of Task T5.3.

D5.4) Report on machine learning techniques applied on Earth Observation (M30) [Responsible: BSC]
This deliverable will summarize the results of Task T5.4.

Work package number	6	**Lead beneficiary**			BSC
Work package title	Optimizing EO Processing on New Hardware Architectures				
Participant number	1	2	3	4	
Short name of participant	BSC	QARNOT	UVEG	IBM	
Person months per participant:	48	8	12	3	
Start month	4		**End month**	36	

Objectives

The main focus of this work package is to study, design and implement software optimizations in order to make efficient use of novel hardware storage and processing technologies. The objectives of the work package can be broken down as follows.

- To study how EO data processing is affected while computing in hardware devices without error correction codes and under the safety voltage threshold.
- To evaluate big data processing on ARM-based architectures, identify bottlenecks and propose solutions to optimize those workloads.
- To study how big data applications perform on architectures with multiple levels of memory, including in-package memory, classic DRAM and persistent memory. Optimize the processing of EO data for such hybrid systems with new latency/bandwidth tradeoffs.
- To port compute-intensive EO processing kernels to hardware accelerators in order to speedup those tasks and deliver near real time services.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Description of work

This work will only focus on optimizing EO data processing on novel hardware technologies that are just arriving into the market or that offer new opportunities that are not fully understood.

Task 6.1 Resilience and robustness of EO applications on low-power ECC-less hardware (M04:M17) [Leader: BSC; Contributors: IBM]

Mainstream personal computers are equipped with error correcting codes (ECC) that continuously monitor the data written in memory devices. This is because the data can be corrupted by multiple reasons, such as high temperature, cosmic rays, among others. However, the cost of ECC is not low, it consumes over 12% of the storage capacity which also represents an important consumption of energy. In the modern era of mobile computing with limited power and graphic devices used for non-critical tasks, unprotected hardware has been gaining popularity, as a measure to reduce costs and to be more energy efficient. Such approaches could potentially be applied to big data and to EO data processing under some restrictions. Moreover, systems equipped with ECC can also suffer from data corruption in some cases, such as when operating at high temperatures. Given that we are operating in a platform made out of heaters it is important to study how those errors could affect the computation. If a system can guarantee that all errors larger than a particular threshold will always be detected, then one could remove the hardware protection and gain over 12% on storage capacity. This also increases the confidence of running EO data processing in hardware that is systematically working at very high temperatures. This could be translated into a substantial gain in energy efficiency. We will evaluate the viability of computing EO services under the presence of errors at different levels and with different levels of accuracy. For this we will make use of the EO applications defined previously. This task is closely related to the WP5 as errors can be viewed as some kind of approximate result.

Task 6.2 Accelerating satellite data processing through GPU computing (M11:M22) [Leader: BSC; Contributors: QARNOT]

The recent advances in GPU computing have eased the efficient porting of big data applications into such devices. For instance, developments in hardware, such as the NVIDIA Pascal architecture [26], enable the use of half-precision floating point numbers and virtual memory addresses beyond the accelerator's physical memory. NVIDIA's Unified Virtual Memory (UVM) or more efficient OpenACC/OpenMP 4.5 compilers greatly enhance accelerator coding productivity. Upcoming accelerator architectures such as NVIDIA Volta are expected to improve UVM performance further by leveraging finer-grained automatic GPU-host data migrations. We will identify compute-intensive kernels that represent significant bottlenecks for the targeted applications and port them into GPUs. In order to develop efficient kernels, we will consider state-of-the-art techniques such as UVM and mixed-precision methods. In particular, we will leverage the results of Task 5.2 to accelerate kernels that can work on mixed precision.

Task 6.3 Porting EO data processing kernels to FPGAs (M25:M36) [Leader: BSC]

[26] http://www.nvidia.com/object/gpu-architecture.html

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

In this task, we will leverage FPGA accelerators to map appropriate compute-intensive tasks into the FPGA fabric. Compared to GPUs, FPGAs are a better fit to do real-time processing on streaming data, especially if the processing can be expressed using fixed-point data. On the other hand, GPUs are a better fit if the kernel has a lot of floating point processing requirements. For this reason, we will analyze and select a suitable subset of DEOS application kernels to run on the FPGAs. To address the programmability issues related with FPGAs, we will leverage a high level synthesis language such as OpenCL or Vivado[27].

Task 6.4 EO data processing on deep memory hierarchies (M19:M30) [Leader: BSC]

Future computing architectures will be composed by multiple storage levels, from in-package high-bandwidth with low capacity memory to high capacity with low bandwidth persistent memories. From SRAM to classic SSD there is a wide range of different type of memories, passing by DRAM and NVM. The new Intel Optane[28] SSD is a good example of the technologies arriving in the market. How to map all the datasets of a big data application into such a hybrid architecture is still an open question. We will develop a strategy to make efficient use such deep-memory hierarchies and we will design a software that can automatically detect and provide guidelines to map the dataset in such a way to optimize the storage use and reduce the number of lost computing cycles. In particular, we will leverage the results of Task 5.1 where the information of each data variable is measured using PCA to get a first idea of the access frequency of the different variables. We will evaluate the performance of the targeted EO services on a deep-memory system.

Task 6.5 EO services visualization accelerated on GPUs (M25:M36) [Leader: BSC; Contributors: UVEG]

An important part of the big data analytics workflow is the visualization of the results obtained after the analysis. This is particularly important for satellite imaging and the understanding of climate events in large geographical areas. Therefore this task focuses on the creation of sophisticated visualization tools to manipulate large amounts of data and render the information in a easy to understand fashion. The visualization products will be accelerated using GPUs and updated into our website.

Deliverables

D6.1) Final report on the impact of memory errors on NRT EO applications (M18) [Responsible: BSC]

Final report on the impact of memory errors on the targeted EO applications. This includes a detailed analysis for each of studied applications showing the frequency and extent of erroneous results while computing on unprotected hardware.

[27] Schmid, M., Schmitt, C., Hannig, F., Malazgirt, G. A., Sonmez, N., Yurdakul, A., & Cristal, A. (2016). Big Data and HPC Acceleration with Vivado HLS. In *FPGAs for Software Programmers* (pp. 115-136). Springer International Publishing.

[28] http://www.intel.com/content/www/us/en/architecture-and-technology/intel-optane-technology.html

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

D6.2) Final report on GPU acceleration for NRT EO applications (M24) [Responsible: BSC]
Final report on the results of porting and optimizing a subset of EO kernels into GPUs. The deliverable will provide detailed speedup numbers as well as identified issues and bottlenecks.

D6.3) Final report on FPGA porting of NRT EO applications (M36) [Responsible: BSC]
Final report on the acceleration of fixed-point EO data processing kernels optimized on FPGAs and the impact on the final application speedup.

D6.4) Report on adapting EO applications to deep memory hierarchies (M30) [Respo: BSC]
Final report on the optimization of multi-spectral EO data processing on deep memory hierarchies. The report includes techniques designed to detect access frequency and data positioning strategies.

D6.5) Final report on GPU-accelerated visualization tools (M36) [Responsible: BSC]
Final report on the visualization of the EO services and the challenges faced to render an interactive product for the End Users. The report will give detailed information about the GPU implementations, optimizations and performance results.

3.1.5 List of Work Packages

Work pack No	Work Package Title	Lead Partic. No	Lead Participant Short Name	Person-Months	Start Month	End month
1	Project Management and Coordination	1	BSC	22	1	36
2	Project Dissemination and Exploitation	1	BSC	31	1	36
3	Distributed Operations, Integration and Validation	2	QARNOT	70	1	36
4	Near Real Time Earth Observation Applications	3	UVEG	88	1	36
5	Low- power Approximate Computing and Data Analytics	4	IBM	66	4	36
6	Optimizing EO Processing on New Hardware Architectures	1	BSC	71	4	36
				348		

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

3.1.6 List of deliverables

Deliv. No	Deliverable name	WP No	Short name of lead participant	Type	Diss. level	Deliv. date
D1.1	Project Management and Quality Guidelines	WP1	BSC	R	PU	M03
D1.2	First Year Periodic report	WP1	BSC	R	CO	M12
D1.3	Second Year Periodic report	WP1	BSC	R	CO	M24
D1.4	Final Periodic report	WP1	BSC	R	CO	M36
D2.1	Dissemination and Communication Plan	WP2	BSC	R	PU	M03
D2.2	Exploitation Plan	WP2	BSC	R	CO	M06
D2.3	Data Management Plan	WP2	BSC	R	CO	M06
D2.4	First Dissemination and Exploitation Report	WP2	BSC	R	PU	M12
D2.5	Second Dissemination and Exploitation Report	WP2	BSC	R	PU	M24
D2.6	Third Dissemination and Exploitation Report	WP2	BSC	R	PU	M36
D3.1	Large dataset management tool for distributed architecture	WP3	QARNOT	DEM	PU	M12
D3.2	Big data adapted smart scheduler	WP3	QARNOT	DEM	PU	M24
D3.3	Resource adaptative NRT processing	WP3	QARNOT	DEM	PU	M36
D3.4	Grid evaluation tool	WP3	QARNOT	DEM	PU	M30
D4.1	Requirements for data retrieval and distribution	WP4	UVEG	R	PU	M06
D4.2	Soil moisture from SMOS-S2/S3/SEVIRI Algorithm Theoretical Baseline Document	WP4	UVEG	R	CO	M18
D4.3	Atmospheric profiles of trace gases from IASI Algorithm Theoretical Baseline Document	WP4	UVEG	R	CO	M18
D4.4	Sea ice diagnostics	WP4	BSC	DEM	PU	M18
D4.5	Analysis of extreme hydrologic events	WP4	UVEG	DEM	PU	M24
D4.6	Prediction of Air quality	WP4	BSC	DEM	PU	M30

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

D4.7	Maps of Wildfire risk covering the European continent	WP4	UVEG	DEM	PU	M30
D4.8	Best practices and roadmap for distributed EO services	WP4	UVEG	R	PU	M36
D5.1	Report on reduction of dataset size	WP5	IBM	R	PU	M18
D5.2	Report on mixed-precision schemes	WP5	BSC	R	PU	M24
D5.3	Report on techniques for detecting relationship across climate events	WP5	IBM	R	PU	M36
D5.4	Report on machine learning techniques applied on Earth Observation	WP5	BSC	R	PU	M30
D6.1	Final report on the impact of memory errors on NRT EO applications	WP6	BSC	R	PU	M18
D6.2	Final report on GPU acceleration for NRT EO applications	WP6	BSC	R	PU	M24
D6.3	Final report on FPGA porting of NRT EO applications	WP6	BSC	R	PU	M36
D6.4	Final report on adapting NRT EO applications to deep memory hierarchies	WP6	BSC	R	PU	M30
D6.5	Final report on GPU-accelerated visualization tools	WP6	BSC	R	PU	M36

3.2 Management structure and procedures

3.2.1 Management structure

The project organizational structure includes the following key components:

- Coordinator:
  - Technical Manager (TM),
  - Project Manager (PM).
- Innovation Manager (IM).
- Work Packages Leaders (WPL).
- End Users Group Coordinator (EUC).
- End Users Group (EUG).

The interactions between the different management components of the project are described in the text below.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Coordinator

The Barcelona Supercomputing Center will serve as Coordinator of the DEOS project. This role is a responsibility shared between the Technical Manager (TM), Leonardo Bautista-Gomez, and the Project Manager (PM), Sergi Madonar, or the individuals assigned to these roles during any interim absences from the project.

Technical Manager (TM)

The Technical Manager (TM) ensures that the scientific and technical objectives of the project are met. The TM defines the high level technical strategy and drives the project team to implement according to that strategy. In implementing this strategy, the TM also ensures that the project maintains its relevance to the H2020 ICT Work Programme 2016-2017 and its strategic objectives. Moreover, the TM organizes technical presentations of project progress to external parties and ensures the appropriate involvement and visibility of the members of the project. The Technical Manager is supported by the Project Manager (PM), who is responsible for the day-to-day execution of the project. The TM collaborates closely with the PM to provide clear and accurate Periodic Reports.

Project Manager (PM)

The Project Manager (PM) is responsible for the day-to-day execution of the project. The PM will ensure the timely delivery of project objectives and deliverables by continuously monitoring the project progress against the plan of record. The Project Manager identifies and tracks issues as well as proposes suitable corrective actions (i.e. resource reallocation, etc) that might require a formal decision by the General Assembly. The PM is also responsible for calling the General Assembly meetings and reviews as well as compiling and distributing Minutes and Actions. The PM defines the procedures for change control (proposed changes to the plan of record), risk management, quality assurance and IPR management. The administrative and financial management of the project is also responsibility of the PM, including internal use of resources monitoring on a 6-month basis, the provisioning of Periodic Reports and Financial Statements, and ensuring an efficient distribution of EU funding. The Project Manager will also act as the official point of contact between the Commission and the Beneficiaries.

Innovation Manager (IM)

The Innovation Manager (IM) has the task to understand and assess innovations and innovators in a project, as well as commercialization opportunities and related strategies. For a given innovation the IM should identify the project partner best placed to take it to market and

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

provide advice on fulfilling the innovation potential. The IM is an expert with a clear affinity for identifying market opportunities and overcoming commercialization hurdles.

The key task of the IM is to collect relevant information on potential innovation and innovators by reading project materials and engaging in discussions with partners at the review meeting. By doing so, and depending on the stage to project (just started, progressed or nearly finished), the IM assesses how ready the consortium/innovator is for entering the market and how they intend to anticipate changing market conditions. At the same time the interaction between the IM and innovators in the consortium is meant to raise their awareness of the issues at hand and to help them develop a more compelling exploitation attitude.

Work Package Leaders (WPLs)

Work Package Leaders are responsible for the scientific and technical work of their respective Work Packages. This includes the planning and control of all activities within the Work Package, the preparation of deliverables and the collection of the contributions from other partners participating in the respective Work Packages for internal and external reports. They meet regularly via teleconference or face-to-face as a part of the GA and arrange for additional technical meetings when necessary. They are expected to raise critical issues to the GA and to support the Technical Manager in coordinating cross-work package relationships within the appropriate activity area. They must actively participate in the regular project-related meetings as well as prepare technical and status presentations as required. Each WPL is appointed by the organisation responsible for the respective WP. Partners appointed as WPL leaders are indicated in Section 3.1.5. The WPLs may nominate separate task leaders when necessary.

End Users Group Coordinator (EUC)

The End Users Group Coordinator consists of one representative of the consortium who works with the Technical Board and the End-Users Group to establish the priorities for the user communities needs and requirements. It will also have close contact with the WP2 Leader to plan any needed industry-related events or workshops, as well as defining the industrial related targets and activities relevant to the project.

End Users Group (EUG)

The End Users Group provides a cost effective mechanism for quickly obtaining real-world feedback on project interim results. Moreover, it facilitates industry’s direct participation in identifying and pursuing exploitation opportunities. The EUG will be an advisory body made up of industrial end-users of Big Data technology. The members will be proposed by the EUC and approved by the General Assembly. The EUG will be responsible for providing feedback on their experience, suggesting possible additions to better align the project with the

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

needs of user communities, as well as reviewing key deliverables and orienting the work of the project to ensure the compatibility of the technology developed with user requirements.

3.2.2 Management procedures and tools

Internal communication tools

In order to support the cooperation among all partners and encourage participation in the decision-making process, a set of internal collaboration tools will be set up. BSC, as the leader of WP1 will be responsible for providing the project with the necessary internal collaborative tools, including: a Project Portal (shared workspace) to facilitate the exchange of critical project documentation and news, an issue tracking system and a set of distribution mailing lists for working sub-groups as appropriate (i.e. by WP). Also, an internal repository will be set for software development and report writing purposes. In addition, the consortium will use modern team communications tools, such as Slack[29].

Progress monitoring

Progress monitoring will be performed through the set of milestones as part as the work plan structure, and summarized in the List of Milestones in Section 3.2.3. The DEOS project work plan structure has been defined in order to ease charting the progress of the project. The Coordinator will ensure that monitoring the work progress and use of resources is done in order to ensure the detection of errors and deviations as early as possible in the project's life cycle. This will enable the consortium to apply systematically corrective actions or contingency plans, if necessary. WPLs will report to the Coordinator the effort spent by their work packages, status of achievement of milestones, production of deliverables and completion of tasks.

Quality Control & Assurance

A quality assurance process will be defined in order to ensure accurate documentation, reporting and justification of the work being carried out. An internal peer-review process will be set up, including a well-defined set of criteria, to assure the project deliverables meet the minimum quality standards before being sent to the European Commission as official outcomes of the project. The high level principles guiding these procedures will be agreed to at the start of the project. The quality assurance process will documented in the Quality Plan (D1.1).

Emergency and Conflict Resolution

Any event that may jeopardize the overall completion date of the Project should be reported immediately to the Project Manager. The PM will call an emergency GA meeting or teleconference as required. Each party involved in the issue must present a short document describing their respective understanding of the conflict that includes at least one proposed solution. The GA reviews the conflict documents and following the procedures of the GA, each member votes on one of the proposed solutions. The solution receiving the simple majority is

[29] https://slack.com/

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

implemented with the chairperson casting the tie-breaking vote as necessary. The formal procedure to deal with such conflicts will be detailed in the Quality Plan (D1.1). Additionally, the CA will describe the process for settlement of disputes.

3.2.3 List of milestones

Milestones are control points where decisions are needed with regard to the next stage of the project. The DEOS consortium aims to be a compact team in where all partners work together for the same goal. Therefore, we have defined global milestones that gather and integrate the progress of the different work packages. For the monitoring and measurement of the DEOS work performance we have defined the following milestones:

Table 3.2a: List of milestones

Mile. No	Milestone name	Related work package(s)	Due date (in month)	Means of verification
1	**Red** : - Toy use case running on the platform	WP3 WP4	6	Prototype
2	**Orange** : - Regular data retrieval and distribution - Wind speed predictions	WP4	12	Service running
3	**Yellow** : - Sea Ice Diagnostics - IASI-derived atmospheric profiles - High resolution soil moisture data - PCA-based data reduction - Data Corruption Robustness Analysis	WP4 WP5 WP6	18	Service running
4	**Green** : - Smart distributed task processing - Wildfire risk maps - Flood/Drought detection - Mixed precision processing - Acceleration on GPUs	WP3 WP4 WP5 WP6	24	Service running
5	**Blue**: - Air quality predictions available - EO Data Regression and Analytics - Runs on Deep Memory Architectures	WP4 WP5 WP6	30	Service running
6	**Purple** : - Dynamic NRT scheduler - Validation of EO services - Correlation and Graph Processing - Acceleration on FPGAs - Visualizations GPU-accelerated	WP3 WP4 WP5 WP6	36	Validated Services

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

3.2.4 Innovation management

In order to ensure that the results of project DEOS will not remain confined in academia or research labs but will find their route toward the market, a proper innovation management is of paramount importance. The innovation manager will work closely with project coordinator and the consortium exploitation team to ensure a proper exploitation path as explained in section 2.2.1 and as defined in the WP2. Innovation management processes include both day-to-day management of knowledge and IPR issues and the iterative creation of exploitation plan and technology roadmaps. More concrete, it will include: (1) Creation of an IPR repository; (2) monitor IPR compliance with H2020 and consortium agreement rules; (3) facilitate any related conflict; (4) facilitate the creation of commercial agreements between partners leading to joint exploitation after the end of the project; (5) Monitor the project to guarantee consistency between technical and marketing choices; (6) monitor the market during the whole duration of the project, particularly concerning evolution of the technology, potential customers, and existing and emerging competitors; (7) plan initiatives that combine technical and exploitation objectives to create business models for defining an exploitation path of most relevant innovations.

3.2.5 Risk assessment and contingency strategies

The DEOS project will elaborate in the first months of the project a Risk assessment and management strategy including planned contingency measures. It will be based on a continuous risk analysis methodology involving all consortium members through the GA. Such Risk management strategy will be described in detail in the Quality Guidelines (D1.1).

The following table provides a list of potential risks identified per work package at the proposal stage. All risks that have been identified to date are classified with low probability but with the potential for high impact. Addressing potential risks will be part of the normal operation of the project, being addressed in the General Assembly monthly meetings. This regular review of potential concerns will ensure the early warning of potential risks an ample time to employ the necessary corrective actions.

Risks considered to be of importance, in particular risks associated with partners not performing or conflicts between partners will be closely monitored by the Coordinator. In general, risk management will be the responsibility of the Coordinator, and the status of any risk situations will be informed to the EC via the Periodic Reports, except when there is a clear need for earlier EC intervention upon decision of the General Assembly.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Description of risk			Work Packages involved	Proposed Risk-mitigation measures
Potential Risk	**Impact**	**Probability**		
Possible delays in appointment of personnel	The project start will be slower than planned	Low	ALL	Partners already have personnel with the required expertise However, partners will start early (before actual project kick-off) to search for qualified personnel
Key milestones or deliverables are delayed	The project results will be delayed	Low	ALL	The PM will foresee possible problems and take early corrective actions to improve the performance of concerned partners.
Expertise risks	Partners are not capable to perform the planned activities	Low	ALL	Partners have been chosen carefully. Partners will react quickly if replacements are required. The Technical Manager will contribute by identifying alternatives.
The planned platform is not large enough to process the satellite data	The project results will not reach the targeted processing speed	Low	WP3	Qarnot grid is continuously growing, the available IT power is expected to double (at least) every year for the next 3 years. Also, it is possible to overflow on public clouds if necessary, and thus maintain SLAs and QoS. These hybrid cloud capabilities can also mitigate unexpected

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

				downtime and increase grid resilience. In addition, we have the support of another distributed grid and they are happy to test our EO services on their platform (See Section 6.4)
Bandwidth and file transfer rates are limiting	Execution time might be higher than expected and compromise NRT objectives	Low	WP3	Dataset splitting and adapted scheduling process will help to mitigate this risk by smartly adapting data transfer to actual grid processing capacity.
Selected algorithms to to merge observational land surface temperatures do not provide full coverage due to the presence of persistent clouds	Predictions of geohazards may not provide full coverage over Europe (specially in northern latitudes)	Low - Medium	WP4, WP5	In case coverage does not reach expectations Europe-wide, predicted land surface temperatures provided by BSC will be used in the predictions.
Selected algorithms based on approximation will not deliver the expected acceleration	The project results will not reach the targeted processing speed	Low - Medium	WP3, WP4, WP5	In case of results below expectations during early testing, the effect of mixed precision will be emulated to predict impact on final applications. Different approaches will be considered early enough to reach the final target.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

The hardware devices for optimization are not delivered in time	The research on optimizing NRT EO for new hardware cannot be done in time	Low - Medium	WP6	BSC already own several clusters with some early prototypes of the targeted hardware architectures. In case the new technologies get delayed we will perform the experiments in our current prototypes and leverage simulators for future architectures.

3.3 Consortium as a whole

DEOS is composed by a consortium of partners with very complementary skills and experience. QARNOT has a large infrastructure that has been working in production providing computing hours to real world customers for several years. Manipulating large datasets and processing tasks in a distributed platform with state-of-the-art big data tools is part of QARNOT's daily routine. This know-how will be critical for the integration, evaluation and validation of the DEOS infrastructure. Their platform is expected to double by next year and reach 5MW in computing power during the timespan of the DEOS project. Additionally, BSC has a long background in developing workflow tools for large-scale data analytics.

From the high-level application perspective we have an excellent team from UVEG with great expertise in Earth Observation and its applications. Their work and outcome of their applications has a great impact on society as they target critical disaster scenarios that could be prevented or mitigated. Additionally, the Earth Science department of BSC contributes with state-of-the-art techniques in a broad spectrum of applications that are applied in current production systems.

On the machine learning and approximate computing domain, IBM and its cognitive computing team are some of the leaders in the field with cutting-edge technology that they have applied to multiple industry areas with great results. The collaboration between IBM and BSC is strong and they already work in joint projects on approximate computing and machine learning.

Finally, BSC has a long history of strong HPC research and working with cutting-edge hardware technology and developing tools and frameworks to efficiently exploit such devices. BSC has received the GPU Center of Excellence Award by NVIDIA (See Section 6.1) and is one of the leaders in large scale accelerated computing. Overall, the structure of the consortium aims to touch all the levels of the big data ecosystem thanks to the great expertise of its partners, with BSC acting as a glue between all the layers, strongly consolidating the entire consortium.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

3.4 Resources to be committed

The total budget estimated for the full duration of DEOS project is **2'247.847 €** subdivided in the following cost categories. The budget distribution between European participants is detailed in the A3 Forms.

Personnel costs

The major part of project costs is dedicated to human resources within the partner organizations. Person effort invested by DEOS European partners totals **348** person months for the duration of the project, and correspond to **1.641.900 €** of direct Personnel Costs. Around 15% of the effort is dedicated to Management, Communication, and Dissemination and Exploitation activities (WP1 and WP2), and around 85% is dedicated to R&D activities. The following table summarizes the effort distribution among the different WPs by beneficiary (WP leaders marked in bold). The table shows the effort allocated among the partners:

Table 3.4a: Summary of staff effort

	WP1	**WP2**	**WP3**	**WP4**	**WP5**	**WP6**	**Total Person-Months per Participant**
1/BSC	**18**	**18**	6	18	24	**48**	132
2/QARNOT	1	3	**60**	0	0	8	72
3/UVEG	1	7	3	**70**	9	12	102
4/IBM	2	3	1	0	**33**	3	42
Total Person / Month	22	31	70	88	66	71	348

Other Direct Costs

Budget for 'Other direct costs' have been carefully considered to ensure that they deemed necessary for ensuring the effective consecution of the work planned and, as a consequence, the successful achievement of the project goals. Other Direct Costs at the EU project level amounts for a total of **148.378 €.** Academic partners are allocated a total amount of **105.378 €** under this category, and industrial partners a total amount of **43.000€** for covering costs for travels to project meetings and attendance to scientific conferences for dissemination purposes as well as costs of publishing in open access.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Table 3.4b: 'Other direct cost' items (travel, equipment, other goods and services)

1 / BSC	Cost (€)	Justification
Travel	17.000	Expenses to cover the face to face meetings, review meetings and conference or dissemination events.
Equipment	15.000	Estimated cost for a few hardware devices to realize the foreseen optimization work (See Section 3.1.4 - WP6)
Other goods and services	52.177,5	Audit costs (2.500 €). Trainings, workshops, promotional material and IP management (49.677,5 €).
Total	84.178	

2 / QARNOT	Cost (€)	Justification
Travel	17.000	Expenses to cover the face to face meetings, review meetings and conference or dissemination events.
Equipment	0	
Other goods and services	5.000	Audit costs.
Total	22.000	

3 / UVEG	Cost (€)	Justification
Travel	17.000	Expenses to cover the face to face meetings, review meetings and conference or dissemination events.
Equipment	0	
Other goods and services	4.000	Audit costs.
Total	21.000	

4 / IBM	Cost (€)	Justification
Travel	17.000	Expenses to cover the face to face meetings, review meetings and conference or dissemination events.
Equipment	0	
Other goods and services	4.200	Audit costs.
Total	21.200	

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

4. Members of the consortium

4.1. Participants

4.1.1 Barcelona Supercomputing Center

Partner Full Name	Barcelona Supercomputing Center	**Participant Number**	1
Partner Short Name	BSC	**Country**	Spain
Type of Organization	Research Center	**Website**	www.bsc.es

Brief description of the entity

The Barcelona Supercomputing Center (BSC) was established in 2005 and is the Spanish national supercomputing facility and a hosting member of the PRACE distributed supercomputing infrastructure. The Center houses MareNostrum, one of the most powerful supercomputers in Europe. The mission of BSC is to research, develop and manage information technologies in order to facilitate scientific progress.

BSC was a pioneer in combining HPC service provision, and R&D into both computer and computational science (life, earth and engineering sciences) under one roof. The centre fosters multidisciplinary scientific collaboration and innovation and currently has over 400 staff from 41 countries. In 2011, BSC was one of only eight Spanish research centres recognized by the national government as a “Severo Ochoa Centre of Excellence”.

BSC has collaborated with industry since its creation, and has participated in projects with companies such as ARM, Bull and Airbus as well as numerous SMEs. BSC also participates in various bilateral joint research centers with companies such as IBM, Microsoft, Intel, NVIDIA and Spanish oil company Repsol. The centre has been extremely active in the EC Framework Programmes and has participated in over one hundred projects funded by it. BSC is a founding member of HiPEAC, the ETP4HPC and participates in the most relevant international roadmapping and discussion forums and has strong links to Latin America.

Education and Training is a priority for the centre and many of BSCs researchers are also university lecturers. BSC offers courses as a PRACE Advanced Training Centre, and through the Spanish national supercomputing network among others.

The Computer Sciences Department focuses on building upon currently available hardware and software technologies and adapting these technologies to make efficient use of supercomputing infrastructures. The department proposes novel architectures for processors and memory hierarchy and develops programming models and innovative implementation approaches for these models as well as tools for performance analysis and prediction.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Main task in the project
BSC will coordinate the project leading the management work package and will lead also the dissemination and exploitation activities in the project. BSC leads also the WP6 about optimizing EO data processing for new hardware technologies. In addition, BSC plays an important role on all of the work packages acting as a cohesive component that guarantees that the different software layers do not get disconnected.
Short profile of key staff members
Dr. Leonardo Bautista-Gomez (male) is a senior researcher at the Barcelona Supercomputing Center where he leads the European Marie Curie project on Deep-memory Ubiquity, Resilience and Optimization. He was awarded the 2016 IEEE TCSC Award for Excellence in Scalable Computing (Early Career Researcher). Before moving to BSC he was a Postdoctoral researcher for 3 years at the Argonne National Laboratory, where he investigated data corruption detection techniques and error propagation. Prior to that, he did his PhD. in resilience for supercomputers at the Tokyo Institute of Technology. He developed a scalable multilevel checkpointing library called Fault Tolerance Interface (FTI) to guarantee application resilience at extreme scale. For this work, he was awarded the 2011 ACM/IEEE George Michael Memorial High Performance Computing Ph.D. Fellow at Supercomputing Conference 2011 (SC11), Honorable Mention. Before moving to Tokyo Tech, he graduated in Master for Distributed Systems from the Paris 6 University. **Dr. Marc Casas** (male) is a senior researcher at the Barcelona Supercomputing Center (BSC). He received a 5-years degree in mathematics in 2004 from the Technical University of Catalonia (UPC) and a PhD degree in Computer Science in 2010 from the Computer Architecture Department of UPC. He was a postdoctoral research scholar at the Lawrence Livermore National Laboratory (LLNL) from 2010 to 2013 working on algorithmic-based fault tolerance and active measurement methods based on software interference. He has received several awards, like a Marie Curie Fellowship, and some of his papers have been awarded in conferences like Euro-Par or Supercomputing (SC). His current research interests are high performance computing architectures, runtime systems and parallel algorithms. He is currently involved with the RoMoL and the Montblanc3 projects as well as the IBM-BSC Deep Learning Center. **Prof. Francisco Doblas-Reyes** (male), is the Director of Earth Science Department at BSC. He is involved in the development of the EC-Earth ESM, and has been since its inception. Prof. Doblas-Reyes is a worldwide expert in the development of seasonal-to-decadal climate prediction systems and has more than 20 years of experience in weather and climate modelling, climate prediction, as well as the development of climate services. For his work in seasonal forecasting, he was awarded the Norbert Gerbier-MUMM International Award from the UN World Meteorological Organization (WMO) in 2006. He serves in several panels of the World Climate Research Programme (WCRP) and the World Weather Research Programme (WWRP) under the UN WMO

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

(among them the steering group of the Polar Prediction Project), is a member of the European Network for Earth System modelling HPC Task Force and has participated in numerous national and European FP4, FP7 and H2020 projects. Currently, Prof. Doblas-Reyes is the principal investigator (PI) or co-investigator in 6 FP7 and H2020 European projects, is coordinator of the FP7 collaborative SPECS project, he is also leading a COPERNICUS action and supervises numerous postdoctoral scientists and software engineers. He has won 50 Million hours of computing time for the High Resolution Ensemble Climate Modeling project through the PRACE network. He is a lead author of the IPCC and member of the steering group of the Polar Prediction Project. Overall, Prof. Doblas-Reyes has authored and co-authored more than 100 peer-reviewed papers on climate modelling and prediction, as well as climate services, and currently has a total of 6103 citations with a h-index of 39.

Pierre-Antoine Bretonnière (male), holds a Masters Degree in "Mathematical and Mechanical Modelling" from the Matmeca engineer school in Bordeaux (France). Graduated in 2010, he has worked in several climate research institutes (CERFACS - Toulouse - France, Catalan Institute of Climate Sciences - Barcelona - Spain and the Earth Sciences Department of the Barcelona Supercomputing Center). His work focuses on climate models outputs and diagnostics, data management and model coupling. He was the person in charge of the data management plan and data conventions definitions in the SPECS FP7 project and has participated in several other European projects. He is also involved in the Research Data Alliance (RDA) framework as chairman of the "Weather, climate and air quality" interest group.

Dr. Alicia Sánchez Lorente (female), holds a PhD in “High-Precision Gamma and X-Ray Spectroscopy” from the Johannes Gutenberg University, Mainz, (Germany). After working at the Helmholtz Institute Mainz (Germany) in the Nuclear Physics Division as a research manager, she joined the Barcelona Supercomputing Center Earth-Sciences co-leading the computational Earth Sciences group. She has a long experience in developing computational techniques as well as data analysis algorithms to deal with huge amount of data (Big Data) and improve Signal-to-Noise Ratio. She has participated on several international Scientific Collaborations worldwide and has helped with her interdisciplinary knowledge to consolidate important european research projects within the framework of High Precision Energy technologies. Currently, her interest focuses on developing Big Data and Machine Learning techniques in the field of earth sciences.

Kim Serradell Maronda (male), holds a Bachelor (2005) in Computer Sciences for the Facultat d'Informàtica de Barcelona (FIB-UPC) and for the Grande école publique d'ingénieurs en informatique, mathématiques appliquées et télécommunications de Grenoble (ENSIMAG). Since 2014 is also Master on High Performance Computing from the Facultat d'Informàtica de Barcelona (FIB-UPC). Currently, he is the manager of the Computational Earth Science (CES) group at the Earth Sciences department in the Barcelona Supercomputing Center (BSC). At present, he is in charge of the computational resources of the department and he was also responsible of the operational runs of the

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

NMMB/BSC-Dust model for BDFC and CALIOPE Air Quality System. He has been involved in European projects like IS-ENES(1&2), ESiWACE, SDS-WAS, BDFC or CONSOLIDER.

Dr. Antonio J. Peña (male) is a Sr. Researcher in the Computer Sciences Department since 2015 and a Juan de la Cierva Fellow since 2017. Within the Programming Models group, he leads research in accelerators and communications for high performance computing. Antonio is the Manager of the BSC/UPC NVIDIA GPU Center of Excellence. He currently organizes PUMPS, the most popular accelerator computing summer school in Europe along with several PRACE Advanced Training Centre (PATC) courses in accelerator programming and communications. He was previously at Argonne National Laboratory (USA) as a Postdoctoral Appointee (2012-2015), driving the heterogeneous memory and accelerator computing areas of research within the Programming Models and Runtime Systems group, where he was the technical lead of the DMEM and VOCL projects while being part of the core MPICH R&D team. Antonio pursued his doctorate in Advanced Computer Systems in a joint collaboration between the Universitat Jaume I and the Universitat Politècnica de València (Spain). His PhD dissertation, titled "GPU Virtualization for High Performance Clusters", was awarded in 2015 with the Extraordinary Doctoral Award from the Universitat Jaume I. This work started the rCUDA (Remote CUDA) project, from which he is the original developer and architect, acting later as the Development Supervisor of the project. Antonio has published over 20 articles in top-ranked computer science journals and conferences and is currently involved in several European projects including ExaNode, EuroEXA, DEEP-ER, IntertWINE, and HBP.

Sergi Madonar (male) joined the BSC in 2016 as Project Manager providing support for the departments of Computer Applications in Science & Engineering (CASE), Operations and Computer Science. He has experience in preparation of proposals and project management in FP7 and Horizon 2020 frameworks. Before he worked five years at Barcelona City Council managing international initiatives in the Smart Cities sector. He holds a Telecommunications Engineering degree by UPC, a Master in Telecommunications Management by UPC and a Master in Business Administration in Knowledge Society by UOC.

Relevant publications

- Bautista-Gomez, L., Zyulkyarov, F., Unsal, O., & McIntosh-Smith, S. (2016, November). Unprotected computing: a large-scale study of DRAM raw error rate on a supercomputer. In *SC'16 Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. ACM.

- Bautista-Gomez, L., Tsuboi, S., Komatitsch, D., Cappello, F., Maruyama, N., & Matsuoka, S. (2011, November). FTI: high performance fault tolerance interface for hybrid systems. In *Proceedings of 2011 international conference for high performance computing, networking, storage and analysis* (p. 32). ACM. **Award for Perfect Review Score.**

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

- M. Casas, B. R. de Supinski, G. Bronevetsky, M. Schulz, Fault Resilience of the Algebraic Multi-Grid Solver, In *Proceedings of the 26nd International Conference on Supercomputing (ICS)*, pages 91-100, 2012.

- Pay MT, Piot M, Jorba O, Basart S, Gassó S, Jiménez-Guerrero P, Gonçalves M, Dabdub D, Baldasano JM, 2010. A full year evaluation of the CALIOPE-EU air quality system in Europe for 2004: a model study. Atmos Environ, 44, 3322-3342.

- Soret, A., Guevara, M., Baldasano, J.M., 2014. The potential impacts of electric vehicles on air quality in the urban areas of Barcelona and Madrid (Spain). Atmospheric Environment, 99, 51–63

Relevant projects

- MONT-BLANC, MONT-BLANC 2 & MONT-BLANC 3
  The MONT-BLANC project, an Exascale project funded by the EU 7th framework programme and H2020, aims to develop a European Exascale approach leveraging on commodity power-efficient embedded technologies. The project has developed a HPC system software stack on ARM, and is deploying the first integrated ARM-based HPC prototype. MONT-BLANC project is also working on a set of 11 scientific applications to be ported and tuned to the prototype system. The rapid progress of MONT-BLANC towards defining a scalable power efficient Exascale platform has revealed a number of challenges and opportunities to broaden the scope of investigations and developments. Particularly, the growing interest of the HPC community in accessing the Mont-Blanc platform calls for increased efforts to setup a production-ready environment. On the third extension of Mont-Blanc (MONT-BLANC3) the main target is the creation of a new high-end HPC platform (SoC and node) that is able to deliver a new level of performance / energy ratio whilst executing real application.

- DEEP, DEEP-ER (DEEP-Extended Reach) & DEEP-EST (DEEP - Extreme Scale Technologies)
  DEEP was an Exascale project funded by the EU 7th framework programme to develop a novel supercomputer architecture based on two different sets of heterogeneous nodes. BSC role in the project was to augment OmpSs programming model to facilitate the porting of large and complex MPI applications on the proposed heterogeneous cluster architecture.
  DEEP-ER (DEEP-Extended Reach) is an extension to DEEP project, also funded by the EU 7th framework programme to address two significant Exascale challenges: the growing gap between I/O bandwidth and compute speed, and the need to significantly improve system resiliency. BSC role in the project is to extend OmpSs with novel task--based checkpoint/restart capabilities to provide a local, asynchronous and lightweight resiliency mechanism that complements traditional resiliency libraries based on synchronized application checkpoint/restart.
  DEEP-EST is an extension to DEEP-ER project, funded by the Horizon 2020 framework programme. project will create a first incarnation of the Modular

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Supercomputer Architecture (MSA) and demonstrate its benefits. In the spirit of the DEEP and DEEP-ER projects, the MSA integrates compute modules with different performance characteristics into a single heterogeneous system. Each module is a parallel, clustered system of potentially large size.

- APPLICATE Advanced Prediction in Polar regions and beyond: Modelling, observing system design and LInkages associated with a Changing Arctic climaTE investigates ways to improve weather and climate prediction in the face of a rapidly changing Arctic. The BSC Earth Sciences department is involved in 5 of the 9 WPs and leads one of them.

- EUPORIAS aims to increase the resilience of the European society to future climate variability by demonstrating how climate forecast information can become directly usable by decision makers in different sectors. IC3 leads research related to wind power planning and operations, and is the coordinator of the work package "Climate Information for Decision Making Processes". EUPORIAS provides a cross-sectoral assessment of the role and perception of probabilistic climate forecasting over seasonal-to-decadal timescales in society and industry, and an experimental basis to test initial strategies for the development of climate services.

Relevant products connected to the subject of this proposal

- CALIOPE (Air Quality Forecast System)

- WMO Dust Centers:Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS), R+D project operated by AEMET and BSC under the umbrella of WMO to study R+SDS modelling and Barcelona Dust Forecast Center (BDFC), EUDAT (EUropean DATa infrastructure), BSC presents as service provider: and as use case pilot.

- RDA (Research Data Alliance): BSC is leader of the Interest Group on Weather, Climate and Air Quality.

Relevant infrastructure

- MareNostrum 4 will have a performance capacity of 13, 7 Petaflop/s. This innovative supercomputer will be made from IBM, which will integrate in one unique machine its own technologies alongside those of Lenovo, Intel and Fujitsu. The general purpose element, provided by Lenovo, will have 48 racks with 3,456 nodes with next generation Intel Xeon processors and a central memory of 390 Terabytes. Its peak power will be over 11,1 Petaflop/s, which is to say that it will be able to perform more than 11,000 trillion operations per second. It will have an Elastic Storage of 15 PBytes.

- MinoTauro is a heterogeneous cluster where the main computational power is provided by NVIDIA GPUS. This cluster is the second most powerful at BSC and it is available as part of the RES resources and as Tier-1 system at the

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

DECI-PRACE calls. The system provides more than 300 TFlops in total. NVIDIA GPU is a heterogeneous cluster with 2 configurations, the first one with 61 Bull B505 blades and the second with 39 bullx R421-E4 servers.

- The BSC-CNS is a key element of and coordinates the Spanish Supercomputing Network, which is the main framework for granting competitive HPC time to Spanish research institutions. Furthermore, BSC-CNS is one of six hosting nodes in France, Germany, Italy and Spain that form the core of the Partnership for Advanced Computing in Europe (PRACE) network. PRACE provides competitive computing time on world-class supercomputers to researchers in the 25 European member countries.

4.1.2 Qarnot Computing

Partner Full Name	Qarnot computing	**Participant Number**	2
Partner Short Name	QARNOT	**Country**	France
Type of Organization	SME	**Website**	www.qarnot.com
Brief description of the entity			

Qarnot computing, founded in 2010, developed and deployed a totally innovative device: **the Q.rad, a connected heater using processors as a heat source**. Totally silent, it remotely gets its computing instructions through **Qarnot's dedicated Q.ware dispatching software platform** over the Internet. In 2013, Qarnot deployed the first version of the Q.rad to provide free and green heat to a building in Paris. It highlighted new and very complementary business opportunities. In 2015 the power consumption of all server and data centres in Europe, with 65 TWh, exceeds the total power consumption of Switzerland. Q.rad heaters are installed in buildings and are used exactly as any standard heater. Driven by the Q.ware, they compute remotely for companies and heats people for free. Instead of using concentrated data centres, Qarnot distributes computing power in buildings.

Qarnot computing is a young SME selling HPC services on one side and heaters on the other side that are actually HPC servers. Regarding hybrid cloud computing Qarnot proposes access to its infrastructure through dedicated pipeline or directly through its API according to commercial contracts. Clients are mainly in the banking / insurance industry and 3D animation, the goal is to pursue the development of these segments and start fostering new opportunities such as this big data approach.

Qarnot computing is glad to participate to this project in order to exploit optimally its heterogeneous computing grid according to heat constraints toward edge computing capabilities. Indeed, DEOS is an opportunity for Qarnot to offer more flexibility to our clients and to increase the actual level of computing usage throughout the year. Our involvement in the project could have a significant impact on Qarnot's grid profitability.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Main task in the project

Qarnot will contribute to the software stack strategy definition to process optimally large data quantity on a widely distributed architecture. This work will imply specific research on computing and storage distribution. Therefore, Qarnot will be responsible for WP3: Software Stack for Distributed Computing and Storage. Its role in this WP3 will be first to work on large datasets retrieval and intelligent transfer to remote distributed grid. Then the scheduling within distributed grids is the second main objective for Qarnot in this WP3. Finally, as NRT execution is seeked in this project, Qarnot will provide continuous integration and grid evaluation to demonstrate its efficiency in terms of big data processing and energy efficiency.

Short profile of key staff members

Mr. Paul Benoit, (male), Chief Executive Officer and founder of Qarnot computing. A graduate of Ecole Polytechnique and Telecom Paristech (X-Telecom), he is a 7 years banking expert, in particular in the field of High-Performance Computing (risk analysis, structured products). Paul previously created a web start-up in 2000. Paul is strongly involved on long-term strategy, product design, along with the company's national and international commercial development.

Mr. Nicolas Saintherant, (male), is Innovation Manager at Qarnot computing. Nicolas holds an engineering degree from Paris-Sud University and a master degree in Technology and Innovation Management obtained at Paris Dauphine University in 2007. He started as scientific project manager for the French embassy in Dublin. He then worked during 9 years as a consultant specialized in innovation management and R&D funding, setting European and French projects for startups, SMEs and large companies. He joined Qarnot in 2016 to strengthen innovation management, from patents and partnerships to R&D funding.

Mr. Yanik Ngoko, (male), research engineer at Qarnot computing, received his M.Sc. in Parallel and Numerical Computing from the University of Yaoundé I and his doctorate in Computer Science from the Institut National Polytechnique de Grenoble, France (2011). Since October 2014, he works as a research engineer at Qarnot computing while being an associate researcher of the University of Paris 13 (Laboratoire d'Informatique de Paris Nord). His research interests include the modelling and resolution of scheduling problems in distributed middlewares, the usage of machine learning techniques for boosting algorithms, the optimization of energy efficiency in clouds, the automation of services composition.

Relevant publications

- Leila Abidi, Souha Bejaoui, Christophe Cérin, Jonathan Lejeune, Yanik Ngoko and Walid Saad; "Data Management for the RedisDG Scientific Workflow Engine", The 6th IEEE International Symposium on Cloud and Service Computing, Pages 1-9, Nada, Fiji, 2016.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

- Yanik Ngoko. “Heating as a Cloud-Service, A Position Paper (Industrial Presentation).” Springer International Conference on Parallel and Distributed Computing (EuroPar), 389-401, Grenoble, France, 2016.

- Yanik Ngoko, Denis Trystram, Valentin Reis, Christophe Cérin. “An Automatic Tuning System for Solving NP-Hard Problems in Clouds.” IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 1443-1452, Chicago, USA, 2016.

- Christian Toinard, Timothee Ravier, Christophe Cérin, Yanik Ngoko. “The Promethee Method for Cloud Brokering with Trust and Assurance Criteria.” IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 1109-1118, Hyderabad, India, 2015.

- T. Baker, B. Al-Dawsari, H. Tawfik, D. Reid, Y. Ngoko; “GreeDi: An energy efficient routing algorithm for big data on cloud.” International Journal on Ad How Networks (Special Issue on Big Data Inspired Data Sensing, Processing and Networking Technologies), Volume 35, pages. 83-96, 2015.

Relevant Products and services

- Hybrid cloud HPC services, IaaS and PaaS

- Digital heater and digital boiler with interfaces and sensors for smart buildings

Relevant projects

- SMARTHEAT: E!8239 Eurostars project 2013. Combination of a High Performance Computing (HPC) server with an electric heater including home automation capabilities.

- EeHPC: H2020 SME instrument phase 1. ICT-37-2014-1 Open Disruptive Innovation Scheme. Energy efficient HPC.

- COMP’HEAT: H2020 SME instrument phase 2 (2016) Seal of Excellence. COMPuting HEATing.

- HPC BOILER: E!11017 Eurostars project 2016. Building a boiler based on HPC rack watercooling.

- GRECO (Gestionnaire de ressources pour clouds d’objets): ANR PRACE 2016 (french collaborative project), Resource manager for cloud of Things.

Relevant infrastructure

Heaters propose 500W of computing power, the grid is currently composed of about 500

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

heaters at the end of 2016. This represents about 250kW of IT power distributed in 7 different locations. As the commercial activity is rising quickly, the installed IT power is expected at least to double in 2017 according to contracts already signed. By 2018, we expect to have more than 10,000 heaters (i.e. more than 5 MW) of IT power available. In 2017, Qarnot computing will build a living lab that will be used as a showroom and as a real environment test bed for R&D activities and research projects.

4.1.3 University of Valencia

Partner Full Name	University of Valencia	**Participant Number**	3
Partner Short Name	UVEG	**Country**	Spain
Type of Organization	Education and Research	**Website**	http://isp.uv.es

Brief description of the entity

The University of Valencia is a public centre for education and research more than 500 years old. The Image Processing Laboratory (IPL) at UVEG is formed by about 50 researchers and lecturers from Electrical Engineering, Physics, and Optics. The lab is globally committed to the development of new imaging systems and vision algorithms, and combines statistical learning theory with the understanding of biological and physics processes to advance knowledge on remote sensing techniques and geosciences.

The Image and Signal Processing (ISP) group is currently formed by 18 members, and is funded both through national and international projects. We receive steady funds from the Spanish Ministry of Science and Innovation, and through joint ventures with the National Institutes of Statistics, Geography, and Meteorology. We have started two spin-off companies, and belong to several NoEs for the advance and transfer of applied computational intelligence (ATICA); COST action on user modelling, interaction and uncertainty; and multimodal interaction in pattern recognition and computer vision (MIPRCV). We have been awarded excellence group in vision research, and become key node in the Campus of Excellence on Multimodal Interaction. We have joint projects with international agencies and companies, such as the European Space Agency (ESA), the European Organisation for Exploitation of Meteorological Satellites (EUMETSAT), Analog Devices Inc, and Google Inc. We have recently been awarded with an ERC Consolidator Grant (2015-2020) to Prof. Gustau Camps-Valls and with a “Ramón y Cajal” tenure-track position (2017-2021) to Dr. María Piles.

Main task in the project

UVEG will contribute to the development of the EO applications and retrieval algorithms that will be run in the DEOS platform. This work will imply specific research to tailor state-of-the-art inversion methods and prediction models to the enhanced spatio-temporal resolution of the new constellation of Sentinel satellites. UVEG will be responsible for

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

WP4: Near Real Time Earth Observation Applications. Its role in this WP4 will be to coordinate the efforts of the different partners in the defined use cases, validate the soil moisture and atmospheric profile retrievals, and provide an evaluation of the wildfire and extreme event alarms, in contact with end-users. UVEG will also participate in the implementation and validation of data analytics and machine learning techniques in WP5 as well as providing guidelines on the acceleration efforts of WP6.

Short profile of key staff members

Dr. Maria Piles Guillem (Female). Msc and PhD in telecommunication engineering (2005, 2010), mastering in remote sensing. Research Fellow at Melbourne University (2010), Research Scientist at Universitat Politècnica de Catalunya (2011-2015), Research Engineer at ICM-CSIC (2016) and Research affiliate at Massachusetts Institute of Technology (2011-2016). Since 2017, she is a "Ramón y Cajal" senior researcher at the IPL, Universitat de València. Her research activity is centered in the retrieval of soil moisture and vegetation geophysical parameters from space observations (microwave radiometers, radars and hyperspectral imagers), and the development of multi-sensor downscaling techniques. She has participated in 10 projects (6 national, 4 international), has advised 9 Ms.C. Theses and 2 Ph.D. Thesis (1 underway). She received the Med-Storm Prize for Young Researchers in the Plinius conference (2011) and the UPC special doctoral award in Information Technology and Communication (2012). She has been awarded with a MIT-Spain/La Cambra de Barcelona Seed Fund (2010-2012), a MIT-MISTI Global Seed Fund (2014-2015) and a BBVA fund to young researchers (2015). Since 2015, she is an external reviewer for the European Commission in the 7th framework program. She is currently serving as president of the IEEE Geoscience and Remote Sensing Society (GRSS), Spanish Chapter. She has published 34 papers (307 cites) in international peer-reviewed journals, 3 book chapters and more than 65 international conference presentations (8 invited).

Prof. Dr. Gustau Camps Valls (Male). Bsc and Msc in Electrical Engineering (1998, 2000), and PhD in Physics (2002). Full professor (habilitation) in the Electrical Eng. Dep. at UVEG, Visiting Researcher at the Univ. Trento (Italy), the MPI Bio. Cyb. (Tübingen, Germany) and Invited Professor at the EPFL (Switzerland). He conducts, supervises and evaluates national and international research projects. He has authored 140 international peer-reviewed journal papers, more than 120 international conference papers, 20 international book chapters, and edited several books on neural and kernel machines. He holds a Hirsch's h index h=48, entered the ISI list of Highly Cited Researchers in 2011, and he is a co-author of the 3 most highly cited papers in relevant journals in geoscience and remote sensing. Thomson Reuters ScienceWatch identified one of his papers as a Fast Moving Front research. Recently he obtained an ERC Consolidator Grant for Statistical Learning for Earth Observation Data Analysis. He is an AE of the IEEE Trans Sig. Proc. and IEEE Geosc. Rem. Sens. Lett. Research interests tied to machine learning for remote sensing and geoscience data.

Dr. Luis Gómez Chova (Male) is Associate Professor at the Electronic Engineering Department of the University of Valencia. He is also a researcher at the Image Processing

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Laboratory (IPL). His work is mainly related to pattern recognition and machine learning applied to remote sensing multispectral images and cloud screening. He is the author (or coauthor) of more than 30 international journal papers, 90 international conference papers, and has authored book chapters in international books. Hirsch's h index = 28.

Dr. Jordi Muñoz Marí (Male). B.Sc. degree in Physics (1993), B.Sc. degree in Electronics Engineering (1996), and Ph.D. in Electronics Engineering (2003). He is currently an Associate Professor in the Electronics Engineering Department at the Universitat de València, where he teaches Electronic Circuits and Digital Signal Processing. He also teaches the subjects of Machine Learning, Active Learning and Big Data in a Data Science Master. He is a research member of the Image Processing Laboratory (IPL). His research activity is tied to the study and development of machine learning algorithms for signal, image processing and remote sensing.

Relevant publications

- BEC L4 high resolution soil moisture product (available at cp34-bec.cmima.csic.es), based on the downscaling technique proposed in “M. Piles; N. Sánchez; M. Vall-llossera; A. Camps; J. Martínez-Fernández; J. Martínez; V. González-Gambau. A downscaling approach for SMOS land observations: evaluation of high resolution soil moisture maps over the Iberian Peninsula. IEEE J. Sel. Topics Applied Earth Obs. Rem. Sens., vol.7, no.9, pp.3845-3857, 2014”

- An emulator toolbox to approximate radiative transfer models with statistical learning. Rivera, J.P. and Verrelst, J. and Gómez-Dans, J. and Muñoz-Marí, J. and Moreno, J. and Camps-Valls, G. Remote Sensing 7 (7) :9347-9370, 2015.

- Piles, M., G. Petropoulos, N. Sánchez, A. González-Zamora & G. Ireland, Towards improved spatio-temporal resolution soil moisture retrievals from the synergy of SMOS and MSG SEVIRI spaceborne observations, Remote Sensing of the Environment,vol 180, pp 403-417, 2016.

- Chaparro, D, M. Vall-llossera, M. Piles, A. Camps, C. Rüdiger and R. Riera-Tatché, Predicting the Extent of Wildfires Using Remotely Sensed Soil Moisture and Temperature Trends, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 9, no. 6, pp. 2818-2829, 2016.

- Camps-Valls, G., J. Munoz-Mari, L. Gomez-Chova, L. Guanter and X. Calbet, "Nonlinear Statistical Retrieval of Atmospheric Profiles From MetOp-IASI and MTG-IRS Infrared Sounding Data," in IEEE Transactions on Geoscience and Remote Sensing, vol. 50, no. 5, pp. 1759-1769, May 2012.

Relevant projects

- Generation of added-value SMOS-derived land products at the Barcelona Expert Center (years 2011-2016) : BEC was created in 2007 to coordinate Spanish contributions to the Soil Moisture and Ocean Salinity (SMOS) satellite mission.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

SMOS brightness temperatures are received at BEC premises in NRT, and the center is responsible for archiving, analysing and distributing SMOS-derived products. The centre is world-leader in the soil moisture and ocean salinity retrieval and nowadays more than 30 users access CP34-BEC services every day. Dr. María Piles led the operations of SMOS-derived land products during years 2011-2016, including the generation of global L3 Soil moisture maps, and of high resolution L4 soil moisture maps over the Iberian Peninsula. A near real-time service of fine-scale soil moisture maps was established in 2012, and these maps are since then being used by the Barcelona's fire prevention services to detect extremely dry soil and vegetation conditions posing a risk of fire. In 2015, an experimental wildfire risk index based on soil moisture and temperature trends was also set in operations.

- Soil Moisture on Planet Earth, MIT International Science and Technology Initiative General Seed Funds (2014-2015 awards) : This project, directed by Prof. Dara Entekhabi (NASA SMAP science lead) and Dr. Maria Piles (science team member of SMOS BEC), aims at developing algorithms that produce high-resolution science data products based on the satellite measurements of the two first missions devoted to measuring the Earth's surface soil moisture: the ESA SMOS and the NASA SMAP. These missions use microwave radiometry at L-band to make soil moisture measurements at daily time scales and with levels of accuracy previously not attained. However, due to technological limitations, the spatial resolution of the observations is limited to 40-50 km; this resolution is adequate for many global applications, but is a limiting factor to its application in regional scale studies, where a resolution of (at least) 1 km is needed. The development of downscaling algorithms to improve the spatial resolution of observations is central to these missions, since it could greatly extend the applicability of the data.

- ERC Consolidator Grant 2015-2020. SEDAL-647423 : Statistical Learning for Earth Observation Data Analysis (SEDAL) is a research project funded by the European Research Council (ERC) Consolidator Grant 2015-2020, and directed by Prof. Gustau Camps-Valls at the Universitat de València, Spain. SEDAL is an interdisciplinary project that aims to develop novel statistical learning methods to analyze Earth Observation (EO) satellite data. In the coming few years, this problem will largely increase: several satellite missions, such as the operational EU Copernicus Sentinels, will be launched, and we will face the urgent need to process and understand huge amounts of complex, heterogeneous, multi-source, and structured data to monitor the rapid changes already occurring in our Planet.

- Sentinels Synergy Framework (SenSyF). EU FP7-SPACE-2012-313117 : The SenSyF project provides a specialised Sandbox Service with tools and development/validation platforms where developers are able to implement and test their applications, and then tap into a distributed pool of cloud resources when ready for the exploitation phase. This project will allow for the development and testing of new processing chains and methods for Sentinel and Copernicus/GMES

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

contributing mission data on a continuous basis, and the delivery of higher-level products and services complementing the information provided by the operational services.

- Cloud detection in the Cloud. Google Research Award 2015 : A method for multitemporal cloud masking and estimation of a cloud-free image proposed in Gomez-Chova et al. (submitted) is being developed using the google earth engine (GEE) platform. It has been developed in the framework of the GEE Award project titled Cloud detection in the cloud (L. Gomez-Chova).

Relevant infrastructure

The ISP has a cluster of 25 interconnected servers with a total of 188 CPUs + 10 GPUs, along with a storage capacity of 200 TB and a small cluster dedicated to the web sites. We also count with computer servers and have access to external computer grids, such as Tirant (at UVEG) and MareNostrum (SCB, Barcelona) facilities.

Thanks to our long standing relations with ESA and EUMETSAT, as well as an agreement with DigitalGlobe, we have a huge library of both real and synthetic time series of archived satellite products including (1) reflectance data (Quickbird, WorldView2, HyMap, SPOT, RapidEye, LandSat, VTG, MODIS and SEVIRI, IASI), and (2) Biophysical products from MODIS, COPERNICUS Global Land, MERIS and LSA SAF. We have access to thematic maps and databases on land-cover and land-use (CORINE, SIOSE) and reference databases such as LUCAS and BIOSOIL. We also work on SMOS-related activities in close collaboration with the BEC team and have direct access to SMOS data through their databases.

4.1.4 IBM Research GmbH

Partner Full Name	IBM Research GmbH	**Participant Number**	4
Partner Short Name	IBM	**Country**	Switzerland
Type of Organization	Business and Research	**Website**	www.zurich.ibm.com

Brief description of the entity

IBM Research GmbH, (IBM Research - Zurich), with approximately 300 employees, is a wholly owned subsidiary of the IBM Research division with headquarters at the T.J. Watson Research Center in Yorktown Heights, NY, USA. IBM Research - Zurich, which was established in 1956, represents the European branch of IBM Research. At the lab, scientific and industrial research is conducted in three scientific and technical departments: Cloud & Computing Infrastructure, Cognitive Computing & Industry Solutions, Science and Technology. Main research topics are nanotechnology, supercomputing, advanced server technology, RAID systems & tape storage, security and privacy. IBM Research - Zurich employs a steady stream of postdoctoral fellows, PhD candidates, and summer

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

students who pass through the laboratory. More than 30 nationalities, primarily from European countries, are represented among the research staff members, including such specialists as computer scientists, mathematicians, electrical engineers, physicists, and chemists. IBM Research - Zurich is involved in more than 80 joint projects with universities throughout Europe, in research programs established by the European Union and the Swiss government, and in co-operation agreements with research institutes of industrial partners.

Main task in the project

IBM will contribute in the project development of new algorithms for data analysis. The algorithm will be designed to speedup the process of large volume of data and therefore will be based on stochastic approach and trans-precision computing. IBM will contribute to most of the WPs, with main focus on WP5.

Short profile of key staff members

Dr. A. Cristiano I. Malossi, (male, PhD 2012), Research Staff Member Foundations of Cognitive Solutions. Dr. Malossi received his B.Sc. in Aerospace Engineering and his M.Sc. in Aeronautical Engineering from the Politecnico di Milano (Italy) in 2004 and 2007, respectively, and his Ph.D. in Applied Mathematics from the Swiss Federal Institute of Technology in Lausanne (EPFL) in 2012. His research focus over the past 5 years include: High Performance Computing, Energy-Aware Algorithms and Architectures, Approximate Computing, Graph Analytics, Numerical Analysis, Computational Fluid Dynamics, Aircraft Design, Computational Geology, and Cardiovascular Simulations. Dr. Malossi was a recipient of the 2013 IBM Research Prize for Computational Sciences, the 2015 ACM Gordon Bell prize, and the 2016 IPDPS Best Paper Award. He will be the primarily contact and responsible person for IBM in the project.

Dr. Costas Bekas, (male, PhD 2003) Manager Foundations of Cognitive Solutions. Dr. Bekas received B. Eng., Msc and PhD diplomas, all from the Computer Engineering & Informatics Department, University of Patras, Greece, in 1998, 2001 and 2003 respectively. In 2003-2005, he worked as a postdoctoral associate with Professor Yousef Saad at the Computer Science & Engineering Department, University of Minnesota, USA. Dr. Bekas's main focus is in HPC systems and their impact in everyday life, science and business. His research agenda spans large scale analytics with an emphasis in graph algorithms/DBs, numerical and combinatorial algorithms, energy aware and fault tolerant systems/methods and computational science. Dr. Bekas brings more than 10 years of experience in high performance computing. During the past several years he has been very active in the field of energy aware HPC and new computing paradigms and architectures as well as the effects of inexact arithmetic and demonstrating their impact in HPC and large scale industrial problems with an emphasis in analytics applications. Dr. Bekas was a recipient of the 2012 Prace Award and the 2013 and 2015 ACM Gordon Bell prizes, and the 2016 IPDPS Best Paper Award.

Relevant publications

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

- D. Rossinelli, B. Hejazialhosseini, P. Hadjidoukas, C. Bekas, A. Curioni, A. Bertsch, S. Futral, S. Schmidt, N. Adams, P. Koumoutsakos. 11 PFLOPS Simulations of Cloud Cavitation Collapse. SC13, ACM Gordon Bell prize, Nov. 2013.

- J. Rudi, A. C. I. Malossi, T. Isaac, G. Stadler, M. Gurnis, P. W. J. Staar, Y. Ineichen, C. Bekas, A. Curioni, O. Ghattas. An extreme-scale implicit solver for complex PDEs: highly heterogeneous flow in earth's mantle. SC15, ACM Gordon Bell prize, Nov. 2015.

- P. Staar, P. Barkoutsos, R. Istrate, C. Malossi, I. Tavernelli, N. Moll, H. Giefers, C. Hagleitner C. Bekas, A. Curioni. Stochastic Matrix-Function Estimators Scalable Big-Data Kernels with High Performance. IPDPS16, Best Paper Session, May 2016.

- C. Bekas, A. Curioni, Very large scale wavefunction orthogonalization in Density Functional Theory electronic structure calculations. Computer Physics Communications, 181(6), 1057-1068, 2010.

- A. C. I. Malossi, Y. Ineichen, C. Bekas, A. Curioni. Fast Exponential Computation on SIMD Architectures, HiPEAC 2015 - 1st Workshop On Approximate Computing (WAPCO), 2015.

Relevant projects

- OPRECOMP, H2020: This project aims to replace the conservative “precise” computing abstraction and with a flexible transprecision computing that exploits approximation in both hardware and software (the former due to the presence of variability or unavoidable physical fluctuations, the latter due to computation with limited precision) to boost energy efficiency.

- EXA2GREEN, FP7: This project focuses on energy aware HPC. We developed linear algebra kernels and algorithms as well as energy aware performance metrics. The platform used include IBM BG/Q, Power 7, Power 8 as well as Intel based blades.

- TEXT, FP7: The focus of this project was to deploy key large scale scientific codes such as CPMD on the task based parallelization framework OMPSS. The work was conducted on massively parallel platforms such as the IBM BG/Q supercomputer and on high performance multicore CPUs such as IBM Power 7.

- HPCDJ, CHIST-ERA FP7: This project deals with introducing multicore HPC techniques and programming models directly into the JAVA language with minimal change to the semantics of the language. We are developing massive graph analytics methods directly in Java and deploy them on multicore shared memory solutions based on IBM Power 8 servers.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Relevant infrastructure
• 2 racks BG/Q system with 2040 nodes and 32768 cores for a total of 420 Tflop/s, 32 Tbytes of RAM and 100 Tbytes of GPFS distributed file system. Possible access to largest IBM BG/Q installations worldwide. • 26x POWER8 Minsky nodes with P100 Pascal NVIDIA GPUs. Each of them consist of 2x POWER8 sockets, 10x cores each, at 3.5 GHz, with 160 threads in total and 1 TB of RAM, plus 4x Pascal NVIDIA GPUs each. The OS is Ubuntu/RHEL. • Access to many IBM worldwide computing infrastructure. • Access to the full Watson Portfolio of APIs to build cognitive systems based on machine learning[1]. • Office space and infrastructure as required. Infrastructure for advanced videoconferences.

[1] http://www.ibm.com/smarterplanet/us/e_n/ibmwatson/

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

4.2. Third parties involved in the project

4.2.1 Barcelona Supercomputing Center

Does the participant plan to subcontract certain tasks (please note that core tasks of the project should not be sub-contracted)	N
Does the participant envisage that part of its work is performed by linked third parties	N
Does the participant envisage the use of contributions in kind provided by third parties (Articles 11 and 12 of the General Model Grant Agreement)	Y

The third party, **Institució Catalana de Recerca i Estudis Avançats (ICREA)** will not carry out any part of the work and just lends resources to the beneficiary. These resources are directly used by the beneficiary, the work is performed in its premises and there is no reimbursement by the beneficiary to the third party (free of charge, Article 12). The third party makes available some of its resources to the beneficiary, which does not reimburse the cost to the third party, but which charges the costs of the third party as an eligible cost of the project. Its costs will be declared by the beneficiary in its Form C but must be recorded in the accounts of the third party. In that context, ICREA resources corresponding to dedicated time of Prof. Francisco J. Doblas-Reyes (ICREA personnel) will be available for the whole duration of the project, mainly for RTD activities.

Prof. Francisco J. Doblas-Reyes is the Director of the Earth Science Department which brings together around 50 people working on the prediction of global weather, climate and air quality, as well as in the analysis of the computational efficiency of Earth science codes.

4.2.2 Qarnot Computing

No third parties involved.

4.2.3 University of Valencia

No third parties involved.

4.2.4 IBM Research GmbH

No third parties involved.

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

5. Ethics and Security

5.1 Ethics

The research carried out as part of DEOS project does not enter any ethics issues in the ethical issue table in the administrative proposal forms.

5.2 Security

Please indicate if your project will involve:

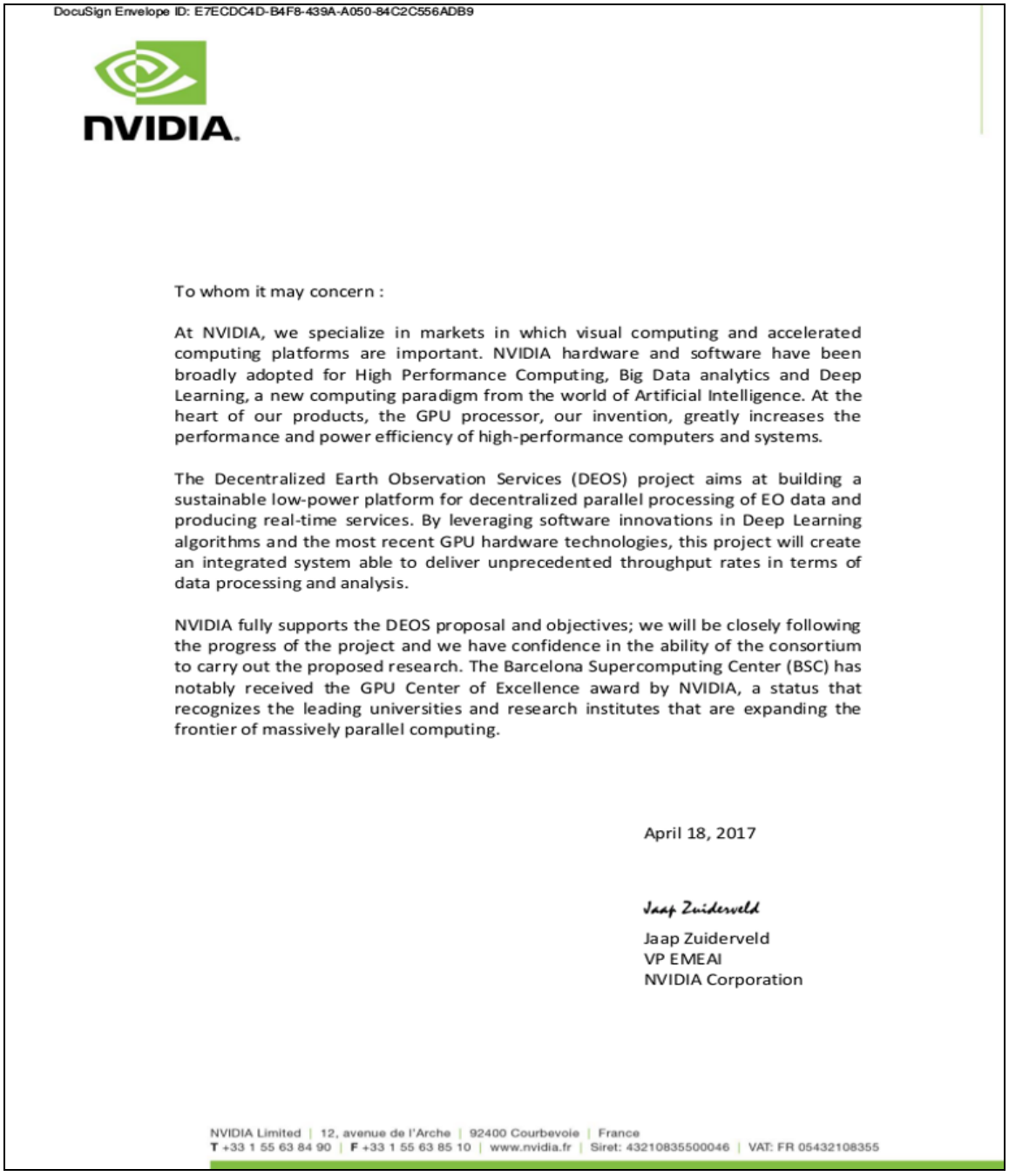
- Activities or results raising security issues: NO
- 'EU-classified information' as background or results: NO

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

6. Annexes

Annex 1: NVIDIA supports DEOS

An important goal of the DEOS project is to speedup the processing of large amounts of data in novel hardware architectures such as GPUs. Thus, having the support of NVIDIA is a great advantage for the DEOS consortium. The letter copied below shows that support.



DocuSign Envelope ID: E7ECDC4D-B4F8-439A-A050-84C2C556ADB9

NVIDIA

To whom it may concern :

At NVIDIA, we specialize in markets in which visual computing and accelerated computing platforms are important. NVIDIA hardware and software have been broadly adopted for High Performance Computing, Big Data analytics and Deep Learning, a new computing paradigm from the world of Artificial Intelligence. At the heart of our products, the GPU processor, our invention, greatly increases the performance and power efficiency of high-performance computers and systems.

The Decentralized Earth Observation Services (DEOS) project aims at building a sustainable low-power platform for decentralized parallel processing of EO data and producing real-time services. By leveraging software innovations in Deep Learning algorithms and the most recent GPU hardware technologies, this project will create an integrated system able to deliver unprecedented throughput rates in terms of data processing and analysis.

NVIDIA fully supports the DEOS proposal and objectives; we will be closely following the progress of the project and we have confidence in the ability of the consortium to carry out the proposed research. The Barcelona Supercomputing Center (BSC) has notably received the GPU Center of Excellence award by NVIDIA, a status that recognizes the leading universities and research institutes that are expanding the frontier of massively parallel computing.

April 18, 2017

Jaap Zuiderveld

Jaap Zuiderveld
VP EMEAI
NVIDIA Corporation

NVIDIA Limited | 12, avenue de l'Arche | 92400 Courbevoie | France
T +33 1 55 63 84 90 | F +33 1 55 63 85 10 | www.nvidia.fr | Siret: 43210835500046 | VAT: FR 05432108355

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Annex 2: IMEC supports DEOS

DEOS aims to study how Big Data processing can be optimized in deep memory hierarchies including NVM. Therefore, it is an important advantage for our consortium to have the support of an industrial partner developing NVM technology solutions such as IMEC. The letter copied below shows that support.

imec

To Whom it may concern,

Subject: Recommendation for the DEOS project

I am a distinguished member of technical staff at IMEC leading the group on various research fronts in the area of design enablement for advanced technology. I am in charge of IMEC's activity on system technology co-optimization to drive technology development from a system and design perspectives. One of the key projects I am leading is technology development for machine learning/neuromorphic computation. I have published over 300 conference and journal papers and I hold more than 40 patents.

I would like to offer my strong support for the DEOS project on Decentralized Earth Observation Systems. I have known PI´s from the coordinating team for a long time now, they have a long history of doing high-quality research and development; and I think their leadership in DEOS will ensure the success of attaining the project goals.

One of the main project goals is on using novel hardware (such as NVM) for doing real-time big data processing. This fits very well with IMEC goals of developing cutting-edge processing and memory technologies.

We would like to continue following the project participants even after the end of the project. I think IMEC could benefit from the results of DEOS: it could be a use case for NVM, and the application that we consider (forecasting forest fires based on streaming satellite data) could be interesting for IMEC.

Summarizing, I would like to offer my strong support for the DEOS project. If you would like to contact me about this recommendation, please email me at: Praveen.Raghavan@imec.be

Yours truly,

R. Praveen

Dr. Praveen Raghavan
Distinguished Member of Technical Staff
Semiconductor Technology and Systems

Date: 23rd April 2017

IMEC vzw
Kapeldreef 75

3001 Leuven

Belgium

Register of Legal Entities Leuven
VAT BE 0425.260.668

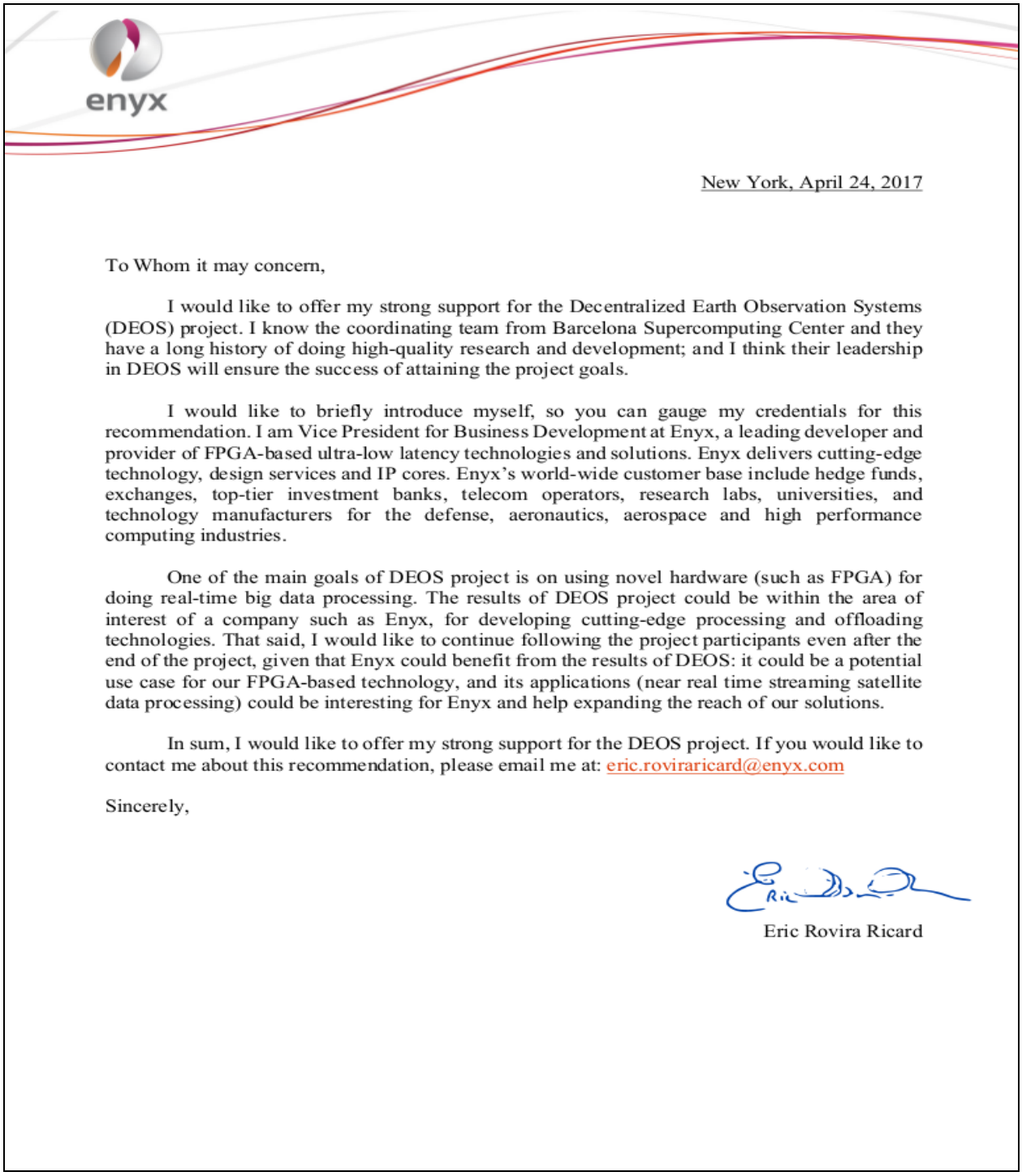
Phone: +32 16 28 12 11

Fax: +32 16 28 94 00
www.imec.be

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Annex 3: Enyx supports DEOS

DEOS aims to accelerate big data processing on hardware devices such as FPGAs. It is of strategic advantage to count with the support of Enyx, one of the global leaders in solutions based on FPGA technology. The letter copied below shows that support.



New York, April 24, 2017

To Whom it may concern,

I would like to offer my strong support for the Decentralized Earth Observation Systems (DEOS) project. I know the coordinating team from Barcelona Supercomputing Center and they have a long history of doing high-quality research and development; and I think their leadership in DEOS will ensure the success of attaining the project goals.

I would like to briefly introduce myself, so you can gauge my credentials for this recommendation. I am Vice President for Business Development at Enyx, a leading developer and provider of FPGA-based ultra-low latency technologies and solutions. Enyx delivers cutting-edge technology, design services and IP cores. Enyx's world-wide customer base include hedge funds, exchanges, top-tier investment banks, telecom operators, research labs, universities, and technology manufacturers for the defense, aeronautics, aerospace and high performance computing industries.

One of the main goals of DEOS project is on using novel hardware (such as FPGA) for doing real-time big data processing. The results of DEOS project could be within the area of interest of a company such as Enyx, for developing cutting-edge processing and offloading technologies. That said, I would like to continue following the project participants even after the end of the project, given that Enyx could benefit from the results of DEOS: it could be a potential use case for our FPGA-based technology, and its applications (near real time streaming satellite data processing) could be interesting for Enyx and help expanding the reach of our solutions.

In sum, I would like to offer my strong support for the DEOS project. If you would like to contact me about this recommendation, please email me at: eric.roviraricard@enyx.com

Sincerely,

Eric Rovira Ricard

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.

Annex 4: GOLEM supports DEOS

Decentralized processing of large datasets is a critical objective to benefit from the data explosion occurring in the recent years. Several startups such as Golem Factory are interested in the same common target. Thus, it is important for the DEOS consortium to have their support. The letter copied below shows that support.

golem

Mr. Leonardo Bautista-Gomez
Centro Nacional de Supercomputación
Carrer de Jordi Girona 29-31
08034 Barcelona, Spain

TO WHOM IT MAY CONCERN

As a CEO of the Golem Factory, I would like to offer my support for the DEOS project on developing a sustainable and highly efficient distributed platform for Earth Observation data processing.

The project goals on a distributed platform for big data computing and our background on decentralized computing are a good match, and I would like to offer our expertise to provide feedback and advice to the DEOS consortium. Indeed, Golem objective is to develop technology for scalable and simple creation of computing infrastructure in decentralized architecture.

The aims of the project fit very well with our strategic agenda as the DEOS use cases focus on processing satellite data in a distributed platform; therefore we would like to maintain our relationship with the project even after the end of the project and we hope that progress achieved within DEOS will make it possible – thanks to the lessons learned and open-source tools and products - to process satellite data also within Golem network. We are also excited about yet another project promoting decentralized approach to computing.

Summarizing, I would like to offer my support for the DEOS project. Should you have questions, do not hesitate to contact me at: julian@golem.network.

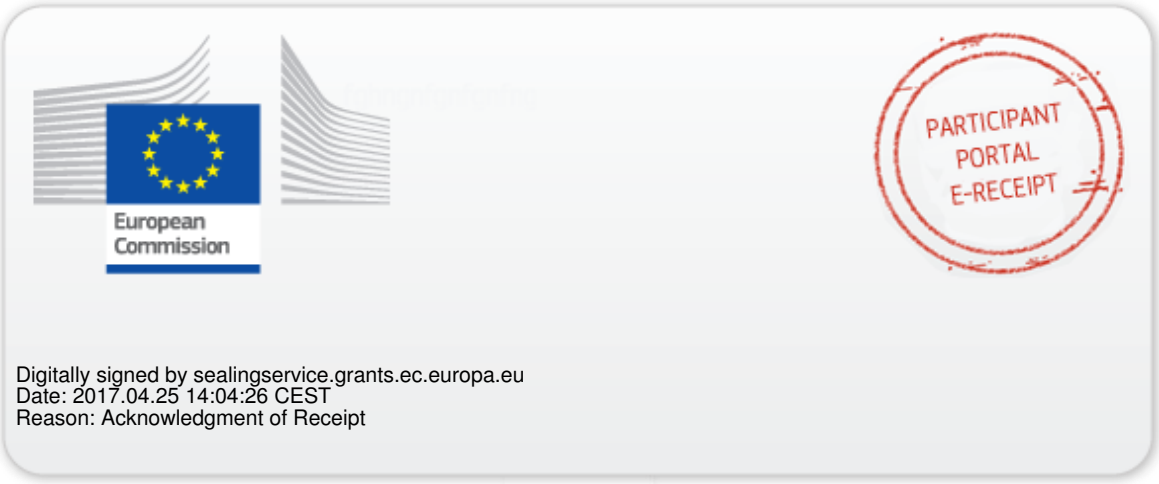
Julian Zawistowski

Golem Factory GmbH

Golem Factory GmbH

c/o Sielva Management SA, Gubelstrasse 11, 6300 Zug, Switzerland

This proposal version was submitted by **Sergi MADONAR** on **25/04/2017 14:04:13** Brussels Local Time. Issued by the Participant Portal Submission Service.



This electronic receipt is a digitally signed version of the document submitted by your organisation. Both the content of the document and a set of metadata have been digitally sealed.

This digital signature mechanism, using a public-private key pair mechanism, uniquely binds this eReceipt to the modules of the Participant Portal of the European Commission, to the transaction for which it was generated and ensures its full integrity. Therefore a complete digitally signed trail of the transaction is available both for your organisation and for the issuer of the eReceipt.

Any attempt to modify the content will lead to a break of the integrity of the electronic signature, which can be verified at any time by clicking on the eReceipt validation symbol.

More info about eReceipts can be found in the FAQ page of the Participant Portal. (http://ec.europa.eu/research/participants/portal/page/faq)