<table>
<thead>
<tr>
<th><strong>Work package:</strong></th>
<th>WP5 – Stakeholder Engagements and Dissemination</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prepared By/Enquiries To:</strong></td>
<td>Dudley Dolan (<a href="mailto:dudley.dolan@q-validus.com">dudley.dolan@q-validus.com</a>) – Q-Validus</td>
</tr>
<tr>
<td><strong>Reviewers:</strong></td>
<td>João Sarraipa ( <a href="mailto:uninova@aquaSmartdata.eu">uninova@aquaSmartdata.eu</a> ) – Uninova</td>
</tr>
<tr>
<td><strong>Status:</strong></td>
<td>Final</td>
</tr>
<tr>
<td><strong>Date:</strong></td>
<td>23/02/2017</td>
</tr>
<tr>
<td><strong>Version:</strong></td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Classification:</strong></td>
<td>Public</td>
</tr>
</tbody>
</table>

**Authorised by:**
Steven Davy  
WIT-TSSG

**Reviewed by:**
João Sarraipa  
Uninova

**Reviewed by:**
Tom Flynn  
Q-Validus

Authorised date: _23/02/2017_
Disclaimer:

This document reflects only authors’ views. Every effort is made to ensure that all statements and information contained herein are accurate. However, the Partners accept no liability for any error or omission in the same. EC is not liable for any use that may be done of the information contained therein.

© Copyright in the document remains vested in the Project Partners.
## AquaSmart Project Profile

**Contract No.:** H2020-ICT-644715

<table>
<thead>
<tr>
<th>Acronym:</th>
<th>AquaSmart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title:</td>
<td>Aquaculture Smart and Open Data Analytics as a Service.</td>
</tr>
<tr>
<td>URL:</td>
<td><a href="http://www.Aquasmartdata.com">www.Aquasmartdata.com</a>, .org, .eu</td>
</tr>
<tr>
<td>Twitter</td>
<td>@AquaSmartData</td>
</tr>
<tr>
<td>LinkedIn Group</td>
<td>AquaSmartData</td>
</tr>
<tr>
<td>Facebook Page</td>
<td><a href="http://www.facebook.com/Aquasmartdata">www.facebook.com/Aquasmartdata</a></td>
</tr>
<tr>
<td>Start Date:</td>
<td>02/02/2015</td>
</tr>
<tr>
<td>Duration:</td>
<td>24 months</td>
</tr>
</tbody>
</table>

### Partners

<table>
<thead>
<tr>
<th>Organisation</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>WATERFORD INSTITUTE OF TECHNOLOGY (TSSG) COORDINATOR</td>
<td>IRELAND</td>
</tr>
<tr>
<td>INTERGRATED INFORMATION SYSTEMS (I2S)</td>
<td>GREECE</td>
</tr>
<tr>
<td>UNINOVA - INSTITUTO DE DESENVOLVIMENTO DE NOVAS TECNOLOGIAS (UNINOVA)</td>
<td>PORTUGAL</td>
</tr>
<tr>
<td>GRAMMOS S.A. (GRAMMOS)</td>
<td>GREECE</td>
</tr>
<tr>
<td>ARDAG COOPERATIVE AGRICULTURAL SOCIETY LTD (ARDAG)</td>
<td>ISRAEL</td>
</tr>
<tr>
<td>NIORDSEAS SL (ANDROMEDA)</td>
<td>SPAIN</td>
</tr>
<tr>
<td>Q-VALIDUS LIMITED (Q-VALIDUS)</td>
<td>IRELAND</td>
</tr>
<tr>
<td>INSTITUT JOZEF STEFAN (JSI)</td>
<td>SLOVENIA</td>
</tr>
</tbody>
</table>

---

**ICT-15-2014: Big data and Open Data Innovation and take-up**  
**H2020-ICT-2014-1**

---

**Public Deliverable**
# PROJECT PARTNER CONTACT INFORMATION

<table>
<thead>
<tr>
<th>WATERFORD INSTITUTE OF TECHNOLOGY.</th>
<th>INTEGRATED INFORMATION SYSTEMS SA.</th>
<th>INSTITUTO DE DESENVOLVIMENTO DE NOVAS TECNOLOGIAS.</th>
<th>GRAMMOS.</th>
</tr>
</thead>
<tbody>
<tr>
<td>T: +353 51 302920 E: <a href="mailto:info@tssg.org">info@tssg.org</a></td>
<td>T: +30 210 8063287 E: <a href="mailto:l2s@l2s.gr">l2s@l2s.gr</a></td>
<td>T: +351 212948527 E: <a href="mailto:rg@uninova.pt">rg@uninova.pt</a> / <a href="mailto:jfss@uninova.pt">jfss@uninova.pt</a></td>
<td>T: +30 26650 29231 E: <a href="mailto:info@grammos-sa.gr">info@grammos-sa.gr</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ARDAG.</th>
<th>ANDROMEDA IBERICA ANDROMEDA GROUP.</th>
<th>Q-VALIDUS LIMITED.</th>
<th>JOŽEF STEFAN INSTITUTE.</th>
</tr>
</thead>
<tbody>
<tr>
<td>T: +972-8-6303200 E: <a href="mailto:nir@ardag.co.il">nir@ardag.co.il</a> <a href="mailto:ido@ardag.co.il">ido@ardag.co.il</a></td>
<td>T: +34 964 587 068 E: AquaSmart @andromedagroup.es</td>
<td>T: +353 1 716 5428 E: <a href="mailto:info@q-validus.com">info@q-validus.com</a></td>
<td>T: +386 1 477 33 77 E: <a href="mailto:dunja.mladenic@ijs.si">dunja.mladenic@ijs.si</a></td>
</tr>
</tbody>
</table>

AquaSmart is co-funded by the European Commission - Agreement Number 644715 (H2020 Programme)
Document Control

This deliverable is the responsibility of the Work Package Leader. It is subject to internal review and formal authorisation procedures in line with ISO 9001 international quality standard procedures.

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Author(s)</th>
<th>Change Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>11/01/17</td>
<td>Dudley Dolan</td>
<td>Table of contents.</td>
</tr>
<tr>
<td>0.2</td>
<td>23/01/2017</td>
<td>Dudley Dolan</td>
<td>Version Ready for CEN review.</td>
</tr>
<tr>
<td>0.3</td>
<td>15/01/2017</td>
<td>Dudley Dolan</td>
<td>Draft version completed.</td>
</tr>
<tr>
<td>0.4</td>
<td>17/02/2017</td>
<td>Dudley Dolan</td>
<td>Version ready for review.</td>
</tr>
<tr>
<td>0.5</td>
<td>21/02/2017</td>
<td>Gary McManus</td>
<td>Document review and edits.</td>
</tr>
<tr>
<td>0.6</td>
<td>22/02/2017</td>
<td>Tom Flynn</td>
<td>Second QA review and issued for approval.</td>
</tr>
<tr>
<td>1.0</td>
<td>23/02/2017</td>
<td>Steven Davy</td>
<td>Approved version release.</td>
</tr>
</tbody>
</table>
Executive Summary

Objectives:
This deliverable sets out the final project version of the CEN Workshop Agreement (CWA), which was planned to be delivered at the end of this project and it was discussed at the kick-off meeting of the CEN Workshop on Big Data. The document establishes the datasets to be used as well as the standardisation approach to be used, and in doing so lays the groundwork for the Big Data standards for the Aquaculture sector.

The formal decision to start work on this CEN Workshop Agreement “Big data Standards for Aquaculture” was taken at the kick off meeting for the aquaSmart project in Luxembourg, February 2015 and the subsequent development of this CEN Workshop Agreement was actioned and executed in line with the CEN CWA workshop protocol. The CWA Project Plan in the premises of CEN/CENELEC in Brussels, on 24th January 2017 to reviewed and discussed at the meeting. CEN duly approved the Plan at the meeting. No objections were registered. The Plan and the CWA is based the work we produced in aquaSmart.

aquaSmart, the Big Data for Aquaculture project, was commissioned by the CEN Workshop on Big Data to identify the requirements for standards in the related area for use by the aquaculture industry, certifying organisations, regulatory authorities and individuals. The aims of the project were to assist in having an effective understanding of the structure of data available, to make proposals for developing analytics and to outline the associated tools that could benefit the aquaculture users. Driven by the business needs of the European aquaculture companies and supporting the EU’s Blue Growth Strategy for marine and maritime sustainable growth Strategy, aquaSmart primarily aims to radically enhance the innovation capacity within the aquaculture sector by helping companies to transform the large volumes of heterogeneous captured data into knowledge, through identification and analysis of this production data, and subsequently using this harvested knowledge to improve performance.

The CWA was developed in accordance with CEN-CENELEC Guide 29 “CEN/CENELEC Workshop Agreements – The way to rapid agreement” and with the relevant provisions of CEN/CENELEC Internal Regulations - Part 2. It was discussed and agreed on 24th January 2017 in a workshop in the CEN/CENELEC premises in Brussels and by representatives of interested parties for approval following an approval period.

Furthermore, this document contains in the appendices a draft Project Plan as was required by CEN for the setting up of a CEN Workshop.
Results:

It was made clear by CEN that the proposed CEN Workshop must not duplicate work being carried out by other CEN activities as well as not duplicating ISO/IEC JTC 1 activities. In order to ensure that this was the case wide ranging consultation took place and it has been concluded that a sectoral approach to Big Data standards will meet the needs of this project and also meet the requirements and restrictions of CEN. The CEN Workshop on Big Data held a successful kick-off meeting on 24th January 2017 where the above was ratified.
# Table of Contents

1 INTRODUCTION ........................................................................................................... 9
2 ABBREVIATIONS AND ACRONYMS ........................................................................... 10
3 Draft CEN Workshop Agreement .................................................................................. 11

3.1 Policy Relevance ........................................................................................................ 11
3.2 Project objectives and aims ....................................................................................... 12
3.3 Data used in Aquaculture ......................................................................................... 13
    3.3.1 Input Variables .................................................................................................. 16
    3.3.2 Data flows and architecture .......................................................................... 18
3.4 Definition of Aquaculture Knowledge Framework ..................................................... 20
    3.4.1 Aquaculture Glossaries or Thesaurus ............................................................. 21
    3.4.2 Aquaculture Domain Ontology ....................................................................... 21
    3.4.3 Aquaculture Training Ontology ...................................................................... 21
    3.4.4 Aquaculture IT Infrastructure Ontology ......................................................... 22
    3.4.5 Aquaculture Knowledge Mechanisms ............................................................. 22
3.5 Target groups for Big Data Standards ...................................................................... 23
    3.5.1 General ........................................................................................................... 23
3.6 Sector specific approach to defining standards for big data ....................................... 23
3.7 Project research outcomes ....................................................................................... 23
    3.7.1 Introduction .................................................................................................... 23
3.8 Recommendations ..................................................................................................... 24
3.9 The work programme and future planning .................................................................. 24
3.10 Methodology ............................................................................................................ 25
3.11 Communication and dissemination ......................................................................... 25

4 CONCLUSIONS ............................................................................................................. 27
5 REFERENCES ................................................................................................................. 28
6 APPENDIX A – CEN Workshop ‘Big data’ Kick-off Meeting ...................................... 29
7 APPENDIX B – CEN Workshop ‘Big data’ Project Plan .............................................. 30
5. Workshop structure ..................................................................................................... 35
TABLE OF FIGURES

Figure 1: Data Flow Diagram................................................................. 18
Figure 2: The aquaSmart Architecture.................................................. 19
Figure 3: The aquaSmart Architecture (expanded)................................. 19
Figure 4: Aquaculture Knowledge Framework........................................ 20
Figure 5: Aquaculture Training Ontology ............................................. 22
TABLES

Table 1: Input Parameters and Second Step................................................................. 17
1 INTRODUCTION

The creation of this CEN Workshop on Big Data was conceived following the identified need for standardization in the domain of Big Data. It is motivated by a number of published European Policy documents including the EU 2016 Rolling Plan for ICT Standardizations and the Digital Single Market Strategy for Europe.

Big Data concerns data sets so large or complex that traditional data processing applications are inadequate. Challenges include analysis, capture, data curation, search, sharing, storage, transfer, visualization and information privacy. The term often refers simply to the use of predictive analytics or other certain advanced methods to extract value from data, and seldom to a particular size of data set. Accuracy in big data may lead to more confident decision making. And better decisions can mean greater operational efficiency, cost reductions and reduced risk.

Analysis of data sets can find new correlations, to "spot business trends, prevent diseases, and combat crime and so on." Scientists, business executives, practitioners of media and advertising and governments alike regularly meet difficulties with large data sets in areas including Internet search, finance and business informatics. Data sets grow in size in part because they are increasingly being gathered by cheap and numerous information-sensing mobile devices, aerial (remote sensing), software logs, cameras, microphones, radio-frequency identification (RFID) readers, and wireless sensor networks. The world's technological per-capita capacity to store information has roughly doubled every 40 months since the 1980s Relational database management systems and desktop statistics and visualization packages often have difficulty handling big data. The work instead requires "massively parallel software running on tens, hundreds, or even thousands of servers". What is considered "big data" varies depending on the capabilities of the users and their tools, and expanding capabilities make Big Data a moving target. Thus, what is considered to be "Big" in one year will become ordinary in later years. "For some organizations, facing hundreds of gigabytes of data for the first time may trigger a need to reconsider data management options. For others, it may take tens or hundreds of terabytes before data size becomes a significant consideration. However, in spite of the relevance of the Big Data today, there is a clear lack and need for regulation concerning its reference Architectures, Technologies, Methods and Applications."
# 2 ABBREVIATIONS AND ACRONYMS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BA</td>
<td>Business Analytics.</td>
</tr>
<tr>
<td>BFCR</td>
<td>Feed Conversion Rate, Biological.</td>
</tr>
<tr>
<td>BI</td>
<td>Business Intelligence.</td>
</tr>
<tr>
<td>BPM</td>
<td>Business Performance Management.</td>
</tr>
<tr>
<td>BPMN</td>
<td>Business Process Model and Notation.</td>
</tr>
<tr>
<td>CEN</td>
<td>Comité Européen de Normalisation.</td>
</tr>
<tr>
<td>COGS</td>
<td>Cost Of Goods Sold.</td>
</tr>
<tr>
<td>EFCR</td>
<td>Feed Conversion Rate, Economic.</td>
</tr>
<tr>
<td>ETL</td>
<td>Extract, Transform, and Load.</td>
</tr>
<tr>
<td>GPD</td>
<td>Growth per Day.</td>
</tr>
<tr>
<td>KPI</td>
<td>Key Performance Indicator.</td>
</tr>
<tr>
<td>KSI</td>
<td>Key Success Indicator.</td>
</tr>
<tr>
<td>LTD</td>
<td>Live To Date.</td>
</tr>
<tr>
<td>MA</td>
<td>Maintainability Answer.</td>
</tr>
<tr>
<td>MQ</td>
<td>Maintainability Question.</td>
</tr>
<tr>
<td>OLAP</td>
<td>Online Analytical Processing.</td>
</tr>
<tr>
<td>OMG</td>
<td>Object Management Group.</td>
</tr>
<tr>
<td>PEPPOL</td>
<td>Pan-European Public eProcurement OnLine.</td>
</tr>
<tr>
<td>PI</td>
<td>Performance Indicator.</td>
</tr>
<tr>
<td>PMC</td>
<td>Portability Metric of a Component.</td>
</tr>
<tr>
<td>PMML</td>
<td>Predictive Model Markup Language.</td>
</tr>
<tr>
<td>SGR</td>
<td>Specific Growth Rate.</td>
</tr>
<tr>
<td>SQuaRE</td>
<td>Systems and software Quality Requirements and Evaluation.</td>
</tr>
<tr>
<td>TGC</td>
<td>Thermal Growth Coefficient.</td>
</tr>
<tr>
<td>TI</td>
<td>Technological Innovations.</td>
</tr>
</tbody>
</table>
3 Draft CEN Workshop Agreement

3.1 Policy Relevance

The project is relevant to EU legislation, policies and actions relating to ICT standardisation, as set out in the 2016 EU ICT Rolling Plan for ICT Standardisation, including the following:

EU 2016 Rolling Plan for ICT Standardisation
(Published December 2015)

“With the continuously growing amount of data (often referred to under the notion Big Data) and the increasing amount of Open Data, interoperability ever more becomes a key issue for leveraging the value of this data.

Standardisation at different levels (such as metadata schemata, data representation formats and licensing conditions of Open Data) is essential to enable broad data integration, data exchange and interoperability with the overall goal to foster innovation on the basis of data. This refers to all types of (multilingual) data, including both structured and unstructured data, as well as data from different domains as diverse as geospatial data, statistical data, weather data, Public Sector Information (PSI) and research data (see also the Rolling Plan contribution on ‘e- Infrastructures for Data and Computing-Intensive Science’), to name just a few.

ACTION 1: invitation to the CEN to support and assist DCAT-AP standardisation process. DCAT-AP is based on the Data Catalogue vocabulary (DCAT). It contains the specifications for metadata records to meet the specific application needs of data portals in Europe while providing semantic interoperability with other applications on the basis of reuse of established controlled vocabularies (e.g. EuroVoc29) and mappings to existing metadata vocabularies (e.g. SDMX, INSPIRE metadata, Dublin Core, etc.). DCAT-AP has been developed by a multi-sectoral expert group. Experts from international standardisation organisations as well as open data portal owners participated in the group to ensure the interoperability of the resulting specification and to assist in its standardisation process.

ACTION 2: promote standardisation in/via the Open Data infrastructure, especially the Pan-European Open Data Portal deployed in the period 2015-2020 as one of the Digital Service Infrastructures under the Connecting Europe Facility programme,

ACTION 3: support of standardisation activities at different levels: H2020 R&D&I activities (see examples in section C above); support internationalisation of standardisation, in
particular for the DCAT-AP specifications developed under the ISA programme (see also action 2 under e-Government, Section D).

**ACTION 4: involvement of stakeholders** in a dialogue about standards for Open Data and Big Data.”

Extracts from the Digital Single Market Strategy for Europe (published in May /June 2015) include the following relevant to the creation of this workshop:

- “**Maximising the growth potential of our European Digital Economy** – this requires investment in ICT infrastructures and technologies such as Cloud computing and Big Data, and research and innovation to boost industrial competitiveness as well as better public services, inclusiveness and skills.”
- “This requires a strong, competitive and dynamic telecoms sector to carry out the necessary investments, to exploit innovations such as Cloud computing, Big Data tools or the Internet of Things.”
- “Only 1.7% of EU enterprises make full use of advanced digital technologies, including mobile internet, cloud computing, social networks or big data while 41% do not use them at all.”
- “Big data, cloud services and the Internet of Things are central to the EU’s competitiveness”.
- “The Big Data sector is growing by 40% per year, seven times faster than the IT market.”

### 3.2 Project objectives and aims

**Project objectives:**

The goal of aquaSmart will be achieved through six key objectives:

1. To facilitate technology transfer in multi-lingual data collection and analytical solutions and services;
2. To implement a multi-lingual Open Data framework that enables companies to seamlessly access global data in order to make knowledgeable decisions;
3. To promote best practices for aquaculture production management in core activities;
4. To develop innovation and deliver state of the art services in the aquaculture sector by tackling the new opportunities to access global data integrated from heterogeneous sources;
5. To develop a training programme and training activities;
6. To deliver a draft CEN standard on ‘Reference Model for Open Data in Aquaculture’.

Project aims:

AquaSmart aims to provide the necessary tools to turn large volumes of heterogeneous aquaculture data into valuable knowledge.

Through the use of the aquaSmart system, you will be able to evaluate feed, feed suppliers, hatcheries, feeding policies, people and management practices and through this identify patterns and trends in your production, identify issues and take the appropriate corrective measures, which lead to improved production and therefore increased profits.

AquaSmart aims to allow analysis and understanding local data and allow benchmarking this against global data.

Using the aquaSmart platform, you will be able to continuously evaluate the performance of your production (see Fig 1) and in turn compare this against similar companies, while adhering to strict privacy rules to protect your specific data from being visible to external parties.

AquaSmart aims to provide the necessary training programme to improve skills and competencies in the aquaculture industry.

Our training programme offers, through adaptable multi-lingual training material, the opportunity to learn about the results of the project as well as specific aquaculture industry knowledge. This is all delivered through the medium of traditional, webinars or mobile.

AquaSmart aims to initiate new working groups within the CEN standardisation body in order to draft a new CEN standard for aquaculture data.

Within AquaSmart we are interested in standardising the use of Open Data in the aquaculture field and standardising the types of analytics that can be sought from big data platforms.

3.3 Data used in Aquaculture

The initial activity of the project is aimed at gathering information on data collected by fish farmers involved in the project, with a primary focus on Europe.

Within the AquaSmart project key datasets have been identified through a series of interviews with the end users (i.e. fish farmers) following a methodology, which is outlined as follows:

- **Identification** – what are the primary questions the users need answered?
- **Analysis** – how is the data used at present to answer those questions?
- **Investigation** – how can technology improve this process?
- **Definition** – define the required dataset attributes.
- **Iteration** – verify, validate, and refine the datasets in cooperation with the users looking for irregularities, missing data etc.
Following an iterative process, datasets were established, which encapsulate the project end user requirements. A refinement process is planned and will continue throughout the lifetime of the project as new sources for datasets become available from fish farms that are external to the project or to support new analyses methodologies.

The End Users (i.e. Andromeda, Argad, Grammos) in the project will use their current and historical production data, through these datasets, to explain their production processes and from this starting point will build the necessary models to support predictive analysis in the aquaculture industry. More specifically, these datasets and models will be used to:

- Evaluate feed effects.
- Evaluate hatcheries and suppliers, in general.
- Evaluate production practices.
- Better estimate the fish number and average weight of the populations that exist in the cages.
- Evaluate the models and adjust them reflect real KPIs.
- Identify bad or spurious data.

In order to accomplish the intention, the datasets that exist in aquaculture production have been identified, and have been classified into the following three main categories:

1. Life to Date (LTD) Data of running cages.
2. LTD Data of closed cages.
3. Periodic data.

And the important dataset according to the end users currently engaged in the project is the periodic dataset between samplings, as this provides real data that can be trusted on a continuous basis.

The dataset, as introduced above for the end users have both input and output variables with the input variables being categorised as Continuous (numeric), Ordinal or Categorical. This is explained as follows:

The **Continuous (numeric)** variables consist of anything that can be measured on a quantitative scale and as such could be any number. The **Ordinal** variables are data that have a fixed, small subset of positive values, which are ordered (i.e. as being poor, fair, good). Finally, the **Categorical** variable consists of data where there are multiple categories involved and they are not ordered (e.g. fish species A Bass, Bream, etc.).
These variables can in turn be further classified:

- Parameters that do not change over time as they may be related to population attributes (e.g., hatchery, stocking year).
- Parameters that may change but due to short sampling periods we consider that their value does not change from sampling to sampling.
- Parameters that change daily, and for these parameters we take into account the averages in the sampling period.

In order to support the integration of open data and global models that are external to the project, we have specified some of the input parameters as being mandatory requirements, with the remainder being optional. As well as these pre-defined parameters, users will also be able to add additional bespoke parameters, which may be relevant to their own production, and thus make these also available for analysis.
### 3.3.1 Input Variables

The input parameters are listed below as follows:

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>TYPE</th>
<th>MANDATORY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographical Region</td>
<td>Categorical</td>
<td>Yes</td>
</tr>
<tr>
<td>Species</td>
<td>Categorical</td>
<td>Yes</td>
</tr>
<tr>
<td>Hatchery</td>
<td>Categorical</td>
<td>No</td>
</tr>
<tr>
<td>Broodstock origin</td>
<td>Categorical</td>
<td>No</td>
</tr>
<tr>
<td>Hatchery Quality (text)</td>
<td>Ordinal</td>
<td>No</td>
</tr>
<tr>
<td>Stocking Year</td>
<td>Categorical</td>
<td>Yes</td>
</tr>
<tr>
<td>Stocking Month</td>
<td>Categorical</td>
<td>Yes</td>
</tr>
<tr>
<td>Days post hatching</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Hatchery CV</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Feed A Feed Category A Feed Supplier *</td>
<td>Categorical</td>
<td>No</td>
</tr>
<tr>
<td>Feed Protein %</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Feed Fat %</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Feed Origin of fish oil</td>
<td>Categorical</td>
<td>No</td>
</tr>
<tr>
<td>Percentage of fish oil</td>
<td>Numeric</td>
<td>No</td>
</tr>
<tr>
<td>Grading Code *</td>
<td>Ordinal</td>
<td>No</td>
</tr>
<tr>
<td>Cage Code</td>
<td>Categorical</td>
<td>No</td>
</tr>
<tr>
<td>Average Weight (start, end)</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Mortality No</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Average Temperature</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Average SFR</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Average Fish Density</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Actual/Model Feed Qty * 100</td>
<td>Numeric</td>
<td>Yes</td>
</tr>
<tr>
<td>Second Step</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Feeder *</td>
<td>Categorical</td>
<td></td>
</tr>
<tr>
<td>Feeding Method</td>
<td>Categorical</td>
<td></td>
</tr>
<tr>
<td>Feeding Parameters **</td>
<td>Categorical</td>
<td></td>
</tr>
<tr>
<td>Inspection Parameters **</td>
<td>Categorical</td>
<td></td>
</tr>
<tr>
<td>Net Code *</td>
<td>Categorical</td>
<td></td>
</tr>
<tr>
<td>Number of fastings</td>
<td>Numeric</td>
<td></td>
</tr>
<tr>
<td>Feeding Rate **</td>
<td>Numeric</td>
<td></td>
</tr>
<tr>
<td>Temperature Variations</td>
<td>Numeric</td>
<td></td>
</tr>
<tr>
<td>Feed Variations</td>
<td>Numeric</td>
<td></td>
</tr>
<tr>
<td>Oxygen</td>
<td>Numeric</td>
<td></td>
</tr>
<tr>
<td>Number of Handlings</td>
<td>Numeric</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1: Input Parameters and Second Step**
3.3.2 Data flows and architecture

A summary of the data gathered highlighting the data flows and processes undertaken within the project is described diagrammatically in the dataflow diagram shown in Figure 1.

![Data Flow Diagram](image)

**Figure 1: Data Flow Diagram**

aquaSmart provides the data mining platform as a cloud service that will be accessible by all the fish farming community. But the scope of the project went beyond that target: by collecting and managing the data mining results from many companies with full respect to confidential data, it generated a knowledge base that will be of maximum usefulness for the aquaculture sector. The companies will be able to transform data to knowledge and use this knowledge to improve efficiency, increase profitability and do business in a sustainable, environment designed for the aquaculture sector. In order to allow even small companies to explore their data and improve in terms of use of feed, environmental impact, growth of the fish, cost, etc.

aquaSmart delivers a cloud based Aquaculture framework (i.e. product, service, training,) supported by an intelligent business model for the analytics of aquaculture data to enable much benefit to be derived in the Aquaculture sector. The introduction of an innovative multilingual knowledge base capacity suitable for the Aquaculture sector, which enables large volumes of data to be accessible as semantic interoperable data and knowledge will improve significantly the sector and ultimately the EU’s competitiveness.
Anonymous data from the companies can be seamlessly imported onto the framework that will incorporate an integrated cloud based data mining services to provide unique data mining insight. This enables the improvement of the knowledge of the system and makes it universal, i.e., the more companies using the framework, the more intelligent the framework becomes.

![Figure 2: The aquaSmart Architecture](image1)

![Figure 3: The aquaSmart Architecture (expanded)](image2)
3.4 Definition of Aquaculture Knowledge Framework

Regarding the theme of knowledge management, the following chapter proposes an Aquaculture Knowledge Framework that formalizes all the data flow shown in the following picture. This picture represents a common reference ontology building process from various enterprises knowledge sources. Such reference ontology integrates the data handled in the domain that once combined represents structured knowledge. Additionally, such overall process also incorporates the assets to support the generation of adaptive multi-lingual e-training service that represents the effective knowledge transfer of the project as illustrated in the right part of the picture bellow.

Knowledge is considered the key asset of modern organizations and industry. The aquaculture domain has a proper nomenclature and the knowledge associated with that economic activity that needs a proper type of knowledge structuring and management. That kind of knowledge organization can be achieved by the development of a specific ontology-based framework aiming to support Aquaculture knowledge, research and operational activities. aquaSmart proposes a framework to be the foundation for the aquaculture knowledge organisation and representation. It specifies the aquaculture knowledge model in four main parts: the aquaculture glossary or thesaurus; the aquaculture domain ontology; the aquaculture training ontology; and the IT infrastructures ontology. The framework also establishes the principles for the knowledge use and management services establishment. It encloses three main parts: searching and reasoning mechanisms; semantic enrichment mechanisms; knowledge and lexicon management mechanisms.

![Aquaculture Knowledge Framework Diagram](image)

**Figure 4: Aquaculture Knowledge Framework**

When an information system intends to represent domain, knowledge needs to be aligned to the community that it represents. Consequently, it is required to have a solution where community
members could present their view on the domain and discuss it with their peers. Additionally, such knowledge must be available and maintained by all the involved actors.

Fundamentally, ontologies are used to improve communication between people and/or computers. By describing the intended meaning of “things” in a formal and unambiguous way, ontologies enhance the ability of both humans and computers to interoperate seamlessly and consequently facilitate the development of knowledge-based (and more intelligent) software applications.

3.4.1 Aquaculture Glossaries or Thesaurus
The main objective of a glossary or thesauri is to be a lexicon reference for a particular community. Thus, an aquaculture glossary or thesauri is such reference but for the aquaculture domain. This domain lexicon integrates terms and concepts with shared definitions (semantics) defined by domain experts. Due to such characteristics, these lexicon elements facilitate the semantic alignment between actors (systems or people) enabling interoperable communications. Additionally, a multi-language glossary that has mappings between the various languages concepts and synonyms outreaches a bigger community.

3.4.2 Aquaculture Domain Ontology
Ontologies allow key concepts and terms relevant to a given domain to be identified and defined in a structure able to express the knowledge of an organisation (Sarraipa et Al., 2010). A good ontology model of any particular domain knowledge facilitates its understanding (Camarinha-Matos and Afsarmanesh 2007). Additionally, its recognised capacity to formally represent knowledge, to facilitate use and maintenance through semantic searching and reasoning, if integrated in a system could be handled for problem solving (Karayel et al. 2004) contributing to such system computational intelligence increasing. Aquaculture domain ontology represents the knowledge in the domain in such way that if defined by domain experts with the support of knowledge engineers, will provide the necessary insights towards the improvement of the efficiency of the aquaculture production processes. Thus, it can enclose knowledge for representing fish diseases, aquaculture production equipment, water quality, etc.

3.4.3 Aquaculture Training Ontology
The aquaculture training ontology will be used to represent the training knowledge base facilitating the categorization of its elements and subsequently reasoning over it. It comprises the model to represent any training curriculum and it is composed by generic training elements as courses, modules, competences, skills, etc. Its main objective is to specify a training curriculum which, addressed by appropriate reasoning mechanisms, will be able to generate customizable training programmes. It should contribute to the skills and competencies development of the trainees as
required for specific understanding and exploitation. Figure 5 presents the relations between training concepts and elements.

![Figure 5: Aquaculture Training Ontology](image)

### 3.4.4 Aquaculture IT Infrastructure Ontology

In the context of any project a set of use cases are normally identified to describe required functionalities that can be provided through particular services. Thus, these services’ can accomplish or support particular business processes and applications. In order to allow future reuse or sharing of these services, an ontology to formalise such IT services or infrastructures in a kind of services UDDI are necessary. This framework will be supported through Semantic Web technologies by providing tools: (i) to define an information model (as an ontology), (ii) to semantically enrich and relate the modelled data and (iii) to query this information. This framework will essentially provide:

1. An information model that allows users to instantiate and catalogue information that describes the functionality and interface of modularized services;
2. A query interface, providing service filtering capabilities and access to the descriptions of individual services.

### 3.4.5 Aquaculture Knowledge Mechanisms

Any knowledge framework requires mechanisms to handle its information. The aquaculture knowledge framework has three different sets of mechanisms. The Semantic Reasoning Mechanisms are services that make use of the knowledge contained in these various described ontologies to apply reasoning techniques able of infer logical consequences from any set of asserted facts. The Semantic Enrichment Mechanisms are mainly services that use the ontologies to enrich knowledge sources as documents or training courses. Finally, the Knowledge Management Mechanisms are
services that use appropriate semantic queries to retrieve or formalise knowledge from/to the ontologies.

### 3.5 Target groups for Big Data Standards

#### 3.5.1 General

The creation of sector specific standards for big data in the aquaculture business will bring benefits to a number of stakeholders. The target stakeholders are listed below.

- Fish Farm Owners
- Fish Farm Managers
- Fish Farm Operatives
- Fish Farm Veterinarians
- Fish Farm Suppliers
- Researchers
- Government organizations (like environmental agencies)

### 3.6 Sector specific approach to defining standards for big data

This CWA deals with standards for Big Data in the Aquaculture sector. There is potential for similar approaches to be used in other sectors for instance;

- Manufacturing
- Healthcare

### 3.7 Project research outcomes

#### 3.7.1 Introduction

The following section of the report outlines the outcomes of the project research.

aquaSmart is an innovative, multi-lingual cloud based tool that uses state of the art technologies and global data access to help the aquaculture sector to 1) lower production costs, 2) improve profitability, 3) improve operational efficiency and 4) carry out their business in a sustainable, environmental friendly way.

aquaSmart enhances innovation capacity of the aquaculture sector by addressing the problem of global knowledge access and seamless data exchange for reuse between aquaculture companies and their stakeholders.
aquaSmart will enable aquaculture companies to perform data mining at the local level and get actionable results, and to further benchmark these results on the global scale through the availability of multi-lingual Open Data.

3.8 Recommendations

User requirements are the key to producing a system that is functional and useful to the end users. Extensive involvement of the end users was critical in the process of deriving the results in the project. It is recommended that the implementation of Big Data standards should always pay particular attention to the requirements of the end user.

Micro-service architectures have proven to provide highly scalable fault-tolerant solutions, and as such have been chosen as an architectural model for AquaSmart. Since they give (almost) complete autonomy to how the services are developed, they offer flexibility in choosing the most appropriate tool for a particular job rather than prescribing a monolithic framework or environment.

Careful consideration should be given to the nature of the data and a proportional approach to handling it proposed with the flexibility adjusting the architecture to incorporate a more “big data” batch oriented approach should this be warranted.

3.9 The work programme and future planning

1. The work on standards commenced at the beginning of the project in February 2015.
2. Initial meetings and discussions were held with project team members to ascertain the source of the various inputs to the proposed CEN Workshop Agreement.
3. Discussions also took place with representatives of CEN regarding the process and potential for a CEN Workshop on Big Data. Arising from these early discussions investigations with ISO/IEC and NSAI were initiated.
4. The work of ISO/IEC JTC 1 WG 9 was researched and it became clear during mid 2015 that there was potential for duplication with this work. The project plan for WG 9 became clear and this allowed this project to plan for a sector specific approach to Big Data Standards while ISO/IEC JTC 1 WG 9 concentrated on overarching standards.
5. Meetings were initiated with NSAI (National Standards Authority of Ireland). NSAI is Ireland’s official standards body and a member of CEN and ISO. Both Tom Flynn and Dudley Dolan participated in a number of meetings with the NSAI to ensure the reaching of a mutual agreement on the approach to the creation of a CEN Workshop on Big Data.
6. A meeting was held with the Irish ICTSCC (ICT Standards Consultative Committee). This cleared the way for inclusion of a sectoral approach to big data standards through CEN in conjunction with and subject to liaison with ISO WG 9.
7. The resources for the secretariat for a CEN Workshop on Big Data were discussed and as a result it was decided to meet with the ICS (Irish Computer Society), which is the representative body for ICT Professionals in Ireland. The ICS agreed in principle that hosting services for the proposed CEN Workshop would be within its mandate. The ICS expressed interest in the potential for the health sector to be included in the work at an early stage.

8. During this process, led by Dudley Dolan, there was continued contributions/input from both Tom Flynn and also by representatives from UNINOVA, particularly Ricardo Goncalves and João Sarraipa (Uninova).

9. In November 2015, a meeting was held with Ray Walshe, Lead Editor for Big Data Standards with ISO/IEC JTC 1 WG 9.

10. The first version of the proposed Business Plan for the CEN Workshop was submitted to CEN in December 2015 and as a result of feed-back a number of changes were incorporated. In particular, due to change in procedures in CEN, the Business Plan will now be called a Project Plan. In addition, further recommendations from CEN were incorporated into a revised document.

11. Following the revision of the Project Plan, the NSAI submitted this to CEN so that the process of creating a CEN Workshop could be continued. The first draft of the CEN Workshop Agreement was created and this was used as a starting point for the CEN Workshop which was held on 24th January 2017. At this meeting the document was ratified by the members and is now being revised prior to being circulated to the community for final approval. The next meeting of this CEN Workshop on Big Data is planned for June 2017.

12. At the kick-off meeting held in the CEN Office in Brussels on 24th January 2017 it was agreed that it would be most appropriate that Mr. Gary McManus of WIT/TSSG should act as Secretary to the CEN Workshop on Big Data in view of his deep involvement with the project team and ongoing interest and commitment to the Big Data concept. Secretarial responsibility remains with NSAI (National Standards Authority of Ireland) and WIT/TSSG will provide the service on their behalf.

3.10 Methodology
This draft CEN Workshop Agreement was developed through interviews, meetings and discussions with the aquaSmart project team. Where possible, the content was taken from already prepared, relevant deliverables. The input was assembled by Q-Validus in consultation with CEN and NSAI. The input was prepared by the aquaSmart project team.

3.11 Communication and dissemination
As the CWA is developed the content will be disseminated publicly through the project web site and through members of the CEN Workshop on Big Data. It is planned to use the process of developing a
CEN Workshop Agreement (CWA) to communicate and disseminate the results of the aquaSmart project.
4 CONCLUSIONS

The objective of developing a CEN Workshop Agreement as a deliverable for this project resulted in wide ranging discussions and negotiations with both European and World-wide Standards Bodies, namely CEN and ISO. During the course of the project agreements were reached between CEN and ISO which ensured that there would not be any unnecessary duplication of effort between the two bodies. This led to discussions with ISO IEC JTC 1 WG9 and also with CEN. It had been made clear by CEN that the proposed CEN Workshop must not duplicate work being carried out by other CEN activities and also that it does not duplicate and ISO/IEC JTC 1 activities. In order to ensure that this is the case, wide ranging consultation has taken place and it has been concluded that a sectoral approach to Big Data standards will meet the needs of this project and also meet the requirements and restrictions of CEN. CEN has introduced changes to their processes and through engagement with CEN, we have adjusted our work in standardisation accordingly.

It was also considered that the setting up of a CEN Workshop for Big Data for Aquaculture would be rather limiting and to this end the scope of the Workshop was increased to include Manufacturing and Healthcare.

The successful creation of a CEN Workshop on Big Data will ensure that the Aquasmart Project will provide a sustainable outcome which can be used by other projects and other sectors.
5 REFERENCES


- CEN: European Committee for Standardization - [https://www.cen.eu/CEN's mission is to promote voluntary technical harmonization in Europe in conjunction with worldwide bodies and its European partners.](https://www.cen.eu/)
6 APPENDIX A – CEN Workshop ‘Big data’ Kick-off Meeting

Kick-off meeting CEN/WS
CEN Workshop ‘Big data’ (WS BDA)

Draft agenda

Date & Time: 24 January 2017, 10:30 – 16:30 CET
Location: CEN – CENELEC Management Centre
17, Avenue Marnix
1000 Brussels

<table>
<thead>
<tr>
<th># Item</th>
<th>Topic</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Opening and welcome (10:30)</td>
<td>Alina IATAN (CEN-CENELEC)</td>
</tr>
<tr>
<td>2</td>
<td>Roll call of participants</td>
<td>Alina IATAN</td>
</tr>
<tr>
<td>3</td>
<td>Adoption of the agenda</td>
<td>Alina IATAN</td>
</tr>
<tr>
<td>4</td>
<td>Presentation of CEN workshop concept</td>
<td>Alina IATAN</td>
</tr>
</tbody>
</table>
| 5      | Background of the Workshop proposal  
1. Background of the proposal  
2. Expression of user needs | Workshop proposers |
| 6      | Establishment of the CEN/WS Big Data  
1. Appointment of Chairman and Vice Chair  
2. Confirm the Workshop Secretariat  
3. Discussion and approval of the Project Plan  
4. Organization of the work  
5. Planning of the meetings, follow-up actions | Alina IATAN / WS Chairman/ WS secretary |
| 7      | Any other business | Chairman |
| 8      | Closing of the meeting (16:30) | Chairman |
7  APPENDIX B – CEN Workshop ‘Big data’ Project Plan
Project Plan for the CEN Workshop on Big Data WS number or WS
Acronym: WS Big Data (WS BDA)

Workshop
(approved during the Kick-off meeting on 2017-01-24)

1. Status of the Project Plan

This Project Plan was approved by the Workshop participants during the Kick Off meeting held on 24 January 2017 in Brussels.

2. Background to the Workshop

The creation of this CEN Workshop on Big Data was conceived following the identified need for standardisation in the domain of Big Data. It is motivated by a number of published European Policy documents including the EU 2016 Rolling Plan for ICT Standardisation and the Digital Single Market Strategy for Europe. The H2020 funding of projects also stresses the importance of standards. In this context the funding of the Aquasmart project related to Big Data in aquaculture was also a factor in the need for a Workshop.

The AQUASMART project responds to the EU’s Blue Growth Strategy for marine and maritime sustainable growth and the Commission’s Europe 2020 Strategy. Aquaculture industry, which comprises mainly SME organizations, represents a significant source of protein for people. Globally, nearly half the fish consumed by humans is produced by fish farms. Global production is forecasted to increase from 45 million tons in 2014 to 85 million by 2030, making the aquaculture industry the fastest growing animal food producing sector in the world. The European Union needs an innovative aquaculture industry to meet rising seafood demand and to enhance its commercial stocks. The proposed CEN Workshop will not only provide a focus for the research work of the Aquasmart Project but will also provide

\[1\] Here the date of updating by the last editor
an excellent platform for dissemination of the key results of the project in the area of analytics.

**Market environment: the challenge of Big Data for Organisations**

The proposed CEN Workshop contributes to reaching policy objectives under Digital Single Market Strategy for Europe (published in May /June 2015) which sets out that “maximizing the growth potential of our European Digital Economy – this requires investment in ICT infrastructures and technologies such as Cloud computing and Big Data, and research and innovation to boost industrial competitiveness as well as better public services, inclusiveness and skills.”

“Big data, cloud services and the Internet of Things are central to the EU’s competitiveness”. “The Big Data sector is growing by 40% per year, seven times faster than the IT market.” “We need to define missing technological standards that are essential for supporting the digitisation of our industrial and services sectors (e.g. Internet of Things, cyber-security, big data and cloud computing) and mandating standardisation bodies for fast delivery.”

Data sets grow in size in part because they are increasingly being gathered by cheap and numerous information-sensing mobile devices, aerial (remote sensing), software logs, cameras, microphones, radio-frequency identification (RFID) readers, and wireless sensor networks. The world's technological per-capita capacity to store information has roughly doubled every 40 months since the 1980s; as of 2012, every day 2.5 Exabyte’s (2.5×1018) of data were created; The challenge for large enterprises is determining who should own big data initiatives that straddle the entire organization.

However, in spite of the relevance of the Big Data today, there is a clear lack and need for regulation concerning its reference Architectures, Technologies, Methods and Applications.

**Policy Environment**

The proposed CEN workshop is in line with the European Policy in a number of areas. Below is an extract from the EU 2016 Rolling Plan for ICT Standardisation (Published December 2015)

“With the continuously growing amount of data (often referred to under the notion Big Data) and the increasing amount of Open Data, interoperability ever more becomes a key issue for leveraging the value of this data. Standardisation at different levels (such as metadata schemata, data representation formats and licensing conditions of Open Data) is essential to enable broad data integration, data exchange and interoperability with the overall goal to foster innovation on the basis of data. This refers to all types of (multilingual) data, including both structured and unstructured data, as well as data from different domains as diverse as geospatial data, statistical data, weather data, Public Sector Information (PSI) and research data (see also the Rolling Plan contribution on ‘e- Infrastructures for Data and Computing-Intensive Science’), to name just a few”
Benefits to all stakeholders

- **Health Care**

Big Data Standards in Healthcare will ensure that information generated by electronic health records, payer claims, pharmacy data and mobile health technologies will offer intriguing possibilities to utilise “big data” technologies in the service of health care. Entities of all types are actively integrating and analysing disparate streams of data to improve the efficiency of everything from drug R&D to care coordination. The ultimate success of both predictive and preventive healthcare resides within this ability to build and analyse large repositories of genetic, phenotypic, prescribing, health outcomes, population and other types of data. There is a massive opportunity to extend the use of modern technologies in healthcare and deliver real benefits to patients, staff and healthcare organisations. With the introduction of Electronic Health Records and e-Referrals, patients can see their medical records on their phone or tablet in addition to their caregivers, enabling patients and clinicians to access and update medical records remotely across all settings.

- **Aquaculture**

Big Data Standards in AQUASMART will enable innovation in the aquaculture sector, turning the available local and heterogeneous large volumes of data in a universally understandable open repository of data assets, globally accessible through the services of the AQUASMART cloud and semantically interoperable. Thus, users in this industry can innovate taking the novel capabilities for seamless and holistic access of multilingual data products and services in the Aquaculture value chain, bridging across borders, languages, industries and sectors, removing barriers both technical and organizational. This brings a set of capabilities never achieved before in this industrial sector, which addresses a capability for data access in cross-sectoral domain, centred in the Aquaculture industry and embracing transportation, logistics and food retailers in a cross border and cross-lingual setting. AQUASMART’s success will attract further participation of other EU aquaculture companies to contribute their data into the incubator established by AQUASMART. The development of the AQUASMART multi-lingual open data framework (i.e. product, service and e-Training) enables an innovative technology and service to be delivered to the aquaculture industry for economic and commercial impact.

- **Other industry sectors**

Many other industry sectors have shown interest in standards for Big Data. These include sections of the Manufacturing sector as well as Tourism interests.

Existing standards and standards related activities and documents

Big Data is a very new area and only a few standardization (or pre-standardization) initiatives have been started at international level. ISO/IEC JTC 1/WG 9 Big Data is
developing an architecture framework completed by some terminology work. ISO/TC 69 has initiated a work item on application of statistical methods for Big Data. A new CEN Workshop, CEN WS ISAEN, will focus on the definition of a so called ISÆN identifier and it will work out a structure of associated metadata that may be used in different use cases. Preliminary discussions have been held with Mr. Wo L Chang, Convenor of ISO/IEC JTC 1/WG-9. The outcome of these discussions indicated that WG-9 could be interested in a sectoral approach to Big Data standards based on CWAs which could be used by WG-9 as use cases in their work.

3. Workshop proposers and Workshop participants

The Workshop proposer is Q-Validus, a member of the AQUASmart consortium responsible for the standardization aspects. Some other organisations have expressed their commitment to support this standardization activity as follows:

- **UNINOVA – Instituto de Desenvolvimento de Novas Tecnol**
- **Irish Computer Society**
- **Insight Centre for Data Analytics**
- Waterford Institute of technology
- Andromeda Group
- Ardag
- Grammos
- Jožef Stefan Institute
- Integrated Information Systems SA

Following approval of the CEN Workshop Project Plan, the registration will be open at the CEN Workshop Secretariat. Participation in the CEN Workshop will remain open to additional interested parties until the end of the public consultation phase. The workshop is open to any other interested party in this subject matter.

4. Workshop scope and objectives

The key objective of this CEN Workshop is to develop three CWAS setting out guidance documents that aid in the uniform understanding and promote reference implementations of Big Data based solutions. The first document to be produced by the Workshop will be a CEN Workshop Agreement (CWA) for use of Big Data in the Aquaculture industry. It is anticipated that in the future the Workshop will create appropriate Big Data CWAs for the Manufacturing Sector and Health Sector.
Workshop programme

The language of the CEN Workshop and its documentation will be English. The CWAs shall be drafted and published in English only.

Meeting plan:
- The Kick-off meeting and first CEN Workshop Plenary Meeting are held in Brussels on 24 January 2017.
- The second plenary meeting - conf call in 21 March 2017
- The third plenary meeting will be held in Madeira in 27 June 2017.
- The following plenary meetings will be agreed upon in the upcoming meetings.

The Workshop participants will work electronically to draft the content of the CWAs. The first Big Data CWA for Aquaculture will be finalized by end of March 2017.

Following the kick off meeting the workshop will be opened to other projects which have shown an interest in producing sectoral standards for big data. It is anticipated that the CWA for HealthCare will commence in March 2017.

The estimative duration of the workshop is a maximum of three years to complete the CWAs for the three sectors.

In this context the proposed workshop should complete its work in three years with the publication of:
- CWA for Aquaculture 2017 (already being worked on)
- CWA for Healthcare 2018
- CWA for Manufacturing 2019

Timetable (may be amended depending upon progress). Any significant change in the timetable will lead to an update of the Project Plan.

5. Workshop structure

The CEN Workshop on Big Data will operate under the CEN rules for the production of CWAs. The members of the Aquasmart project team will take part to the workshop.

The Chairman will be appointed at the kick-off meeting. The responsibilities of the Workshop Chair include the following tasks:
- To chair Workshop plenary meetings;
- To ensure that the Workshop delivers in line with its Project Plan;
- To manage the consensus building process.
• To draft the first draft CWA’s and incorporating the comments to produce revision 2 documents.

NSAI will provide the Workshop secretariat, subject to formal acceptance of the business plan at the Kick-Off meeting. The NSAI will submit the project plan and provide the secretariat through the Irish Computer Society. The meetings will be called by the secretariat and all minutes and documentation will be maintained by the Secretariat.

The responsibilities of the secretariat cover the co-ordination of administrative duties involved in the organization of the Workshop including:

• To maintain the meeting schedules and arrange the physical meeting locations;
• To distribute the agenda and other documents for the Plenary Sessions;
• To prepare draft minutes of the Plenary Sessions;
• To maintain the web site and ensure the timely availability of the working documents for meetings;
• To ensure that the CWA is available in time on the appropriate CEN format.
• To interface with the CCMC regarding strategic issues, problems arising, external relationships, etc.

The development project runs mainly as a gathering of seasoned experts responsible for commenting the drafts and suggesting improvements via electronic mail. All communication shall be copied to Secretariat and all participants to ensure transparency, openness and equal treatment of all stakeholders.

The CEN Workshop Secretariat will support the agreed upon CEN Workshop activities. The Secretariat provides a professional management support in the form of administrative, operational and technical services to the Workshop. (Preparation of agendas, approvals, communication, arrangement of meetings, registration of decisions and minutes.

6. Resource requirements

All costs related to the participation of interested parties in the Workshop’s activities have to be borne by themselves. It is anticipated that no fee per registered member will be charged.

Registered members will have voting rights and will effectively run the workshop. The workshop Plenary Meetings will be open to non-registered members on the invitation of the Chairman.

7. Related activities, liaisons, etc.
At an early date, it is intended to seek liaison arrangements with ISO/IEC JTC 1 WG 9. The CEN Workshop will make use of the upcoming output of WG 9 in a number of areas as soon as information is available. See below:

ISO/IEC 20546 Information technology Big data – Overview and vocabulary
ISO/IEC 20547 Information Technology – Big data Reference architecture (5 Parts)
8. **Contact points**

<table>
<thead>
<tr>
<th><strong>Chairman</strong></th>
<th><strong>Secretariat</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name:</strong> Dudley Dolan</td>
<td><strong>NSAI- National Standards Authority of Ireland</strong></td>
</tr>
<tr>
<td><strong>Company:</strong> Q-Validus, Dublin 4 Ireland</td>
<td><strong>Secretary:</strong> Gary McManus</td>
</tr>
<tr>
<td><strong>E-mail:</strong> <a href="mailto:dudley.dolan@eircom.net">dudley.dolan@eircom.net</a></td>
<td><strong>Digital technology</strong></td>
</tr>
<tr>
<td><strong>Vice-chairperson</strong></td>
<td><strong>E-mail:</strong> <a href="mailto:gmcmanus@tssg.org">gmcmanus@tssg.org</a></td>
</tr>
<tr>
<td><strong>Name:</strong> Ricardo Jardim-Goncalves</td>
<td></td>
</tr>
<tr>
<td><strong>Company</strong> UNINOVA</td>
<td></td>
</tr>
<tr>
<td><strong>E-mail:</strong> <a href="mailto:rg@uninova.pt">rg@uninova.pt</a></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>CEN-CENELEC Management Centre</strong></th>
<th><strong>Contact CEN Workshop Proposers</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name:</strong> Alina IATAN</td>
<td>Dudley Dolan</td>
</tr>
<tr>
<td><strong>Program manager</strong></td>
<td>Q-Validus</td>
</tr>
<tr>
<td><strong>E-mail:</strong> <a href="mailto:aiatan@cencenelec.eu">aiatan@cencenelec.eu</a></td>
<td><strong>E-mail:</strong> <a href="mailto:dudley.dolan@eircom.net">dudley.dolan@eircom.net</a></td>
</tr>
<tr>
<td><strong>phone:</strong> +32 2 255 0816</td>
<td></td>
</tr>
</tbody>
</table>
Annex A- List of Workshop proposers

List of organisations supporting the Workshop proposal

1. **Waterford Institute of Technology (WIT)**: WIT is the major provider of higher education in the South East region of Ireland and one of the largest Institutes of Technology's in Ireland. The Institute's range and scope reflect the changing needs of the southeast as well as the country as a whole. In particular, WIT hosts the Telecommunications Software & Systems Group (TSSG), which is an internationally recognized centre of excellence for ICT research and innovation. TSSG carries out a wide spectrum of industry informed research in Information and Communications Technologies (ICT), particularly technologies enabling communications and information services. TSSG’s four key prioritised technical research areas include: Mobile Platforms and Services, Security Privacy and Identity, Data Analytics and Social computing, Adaptive Networks and services. They create economic impact by translating their knowledge base and innovation into leading edge products and services by continuing their engagement with Industry in collaborative R&D, knowledge generation and transfer. Over the past five years, TSSG has delivered innovative solutions to over 110 Irish companies, and has created 11 spin out companies in the South East including the award winning FeedHenry, a groundbreaking mobile cloud platform company, and ZolkC, a leading provider of mobile technology for international visitor attractions.

2. **INTEGRATED INFORMATION SYSTEM SA (I2S)** is a Small to Medium Sized organization currently employing eleven people. The company was established in 1997 with offices in Rhodes and Athens. I2S is a high-tech company specialized in the area of software technology. Our mission is to develop and deliver innovative software technology that helps businesses become more competitive. Our consultancy and training services assist our clients to put this technology to work. The main activities of the company are development of software for the aquaculture sector. We are the developers of the aquaManager system (www.aquaKmanager.com), which is an integrated solution for production planning, management, cost analysis and financial forecast. It supports all stages of the fish farming production, from hatchery to harvest. It is used by some of the largest fish farming groups in the world from Europe to America to Middle East and Africa.

3. **The Instituto de Desenvolvimento de Novas Tecnologias (UNINOVA)** is a multidisciplinary, independent and non-profit research institute employing around 180 persons, located in the metropolitan area of Lisbon. It was formed in 1986 by the Faculty of S&T of the University Nova de Lisboa (FCTKUNL K www.fct.unl.pt), a group of industrial associations, a financial holding, and up to 30 companies. It is an active partner of Madan Parque (www.madanparque.pt), a business
facilitator and accelerator, incubating Micro and SME’s through several layers of support to entrepreneurial activity. The main aim of UNINOVA is to pursue excellence in scientific research, technical development, advanced training and education. By working closely with industry and universities, technological innovations are transferred into profitable business concepts and, existing products further developed to match new industrial requirements. Due to its tight connection with the University and Madan Parque, UNINOVA has, since its foundation, hosted and supported the development several PhD thesis, as well as the creation of several successful spin-offs.

4. **GRAMMOS S.A.** was founded on May 11, 1999. It is located in the municipality of Filiata, prefecture of Thesprotia in western Greece. It operates two sea parks, of a total sea surface of 1.5 acres and 2.25 acres respectively. GRAMMOS is producing Sea Bass (*DICENTRARCHUS LABRAX*), Sea Bream (*SPARUS AURATA*) and Sharpsnout Sea Bream (*DIPLODUS PUNTAZZO*). Every year, the company provides the market with thousands of tons of fresh fish that reach the consumers in the greatest possible speed and safety. Production is very well organized and supported by specialized software systems.

5. **ARDAG** is Israel’s oldest and largest mariculture group in Israel. Based in Gulf of Aqaba/Eilat, ARDAG is owned by five kibbutzim (collective settlements) in the Southern Arava Valley. Ardag is a small-medium size limited company, currently employing about 70 people in hatchery, packing hose and 2 farms. Ardag was founded by the mid. 1980’s in the red sea, and during the years was spreading throughout Israel towards the Mediterranean sea. The farms are now located in the Mediterranean Sea – one is in Ashdod port, the other site is located 6 miles off-shore Ashdod. The 2 sites produced about 2,000 tons of sea-bream per year, and the hatchery produces about 10,000,000 fingerlings of sea-bream and other species of sea-water fish. There are plans to add more marine species to the existing “fish basket” in the future to improve and increase the overall production and profits of the company. From 2008, there is a joint venture between ARDAG, Magan Michael and NCM, to hatch and grow the Mediterranean White Grouper (*Epinephelus anaenous*). The Ashdod farm is holding a HACCP protocol for the last 7 years. All products are selling domestic as fresh chilled fish. Part of Ardag group is Karat Caviar growing Russian sturgeon (*Acipenser gueldenstaedtii*) since 1992. Karat Caviar has succeeded in improving the environmental conditions in which the fish develop. The Farm combines traditional methods of production with modern sturgeon breeding techniques, resulting in our distinctively delicious Karat Caviar.
6. **Andromeda Iberica**: is a limited company and a medium sized organization currently employing people in farms and packaging plant and in the hatcheries. It was founded in 2001, in Burriana (Spain) as Acuicola Marina. In 2007, the company created the packaging installation in Burriana, Frescamar. In 2009 increased the presence in Spain incorporating 2 farms (Niorseas SL and Industrias Pesqueras Balmar) located in Alicante and one hatchery (Alevines del Sureste) located in Murcia. In 2008, it became part of the group Andromeda. In 2010 increased the hatchery production volume with the acquisition of Piscimar allocated in Burriana. Andromeda Iberica established all the production in the east southeast of Spain, concretely in the regions of Comunidad Valencia and Murcia. The actual production per annum is around 6.200 ton, divided in 2.037 ton of bass, 2.732 ton of bream and 1.361 ton of meagre.

Actually, our production arrives with the maximum freshness to all European countries, and we are working closely with our customers to satisfy the necessities of the final consumer. The company is certified to ISO 9001:2008 and ISO 14001:2004, both from 2009 and ISO 22000 from 2014, in addition to the achievement of GLOBAG.A.P in 2012.

7. **Q-VALIDUS** An Irish SME headquartered at Nova UCD the Innovation and Technology Transfer Centre at University College Dublin. The company was founded in 2007 to meet the growing demand for certification of skills, training and knowledge across a wide range of markets and occupations. The demand for certification is being driven by regulatory factors, requirements for continuing professional development, as well as related industry compliance requirements. Other drivers include advances in automated technology solutions, globalisation, and increased workforce mobility and accessibility and training requirements. The production of valid, credible and legally defensible testing and assessment solutions is a complex task requiring specialist skills. By combining decades of experience in assessment and certification with best industry practice and proven psychometric principles and standards, Q-Validus can offer clients the latest technology solutions and advisory services to design, develop, market, manage and deliver the very best in assessment and certification programmes. In addition to technical services delivery, Q-Validus also develops and deploys its own IT certification programmes. Q-Validus has implemented many on-site training course, created new IT certification programmes that are designed to improve productivity in the workplace through more efficient and effective use of modern day office IT applications. The certification programmes focus on role based activities and are designed to improve functional and collaborative skills within the workplace. The Q-Validus certification programmes provide a ‘total solutions’ offering combining syllabus definitions, courseware materials, e-learning support tools, tutor ware, automated testing and candidate management
solutions across global markets with flexible delivery options. Customers include the ECDL Foundation, BCS, Coca Cola, NAMA, IPA, Oasis Ventures Dubai and many others.

8. **The Jožef Stefan Institute (JSI)** is the leading Slovenian scientific research institute, covering a broad spectrum of basic and applied research. The staff of more than 930 specializes in natural sciences, life sciences and engineering. The subjects concern production and control technologies, communication and computer technologies, knowledge technologies, biotechnologies, new materials, environmental technologies, nanotechnologies, and nuclear engineering. The mission of the Jožef Stefan Institute is the accumulation K and dissemination K of knowledge at the frontiers of natural science and technology to the benefit of society at large through the pursuit of education, learning, research, and development of high technology at the highest international levels of excellence. The Artificial Intelligence Laboratory is concerned mainly with research and development in information technologies with an emphasis on artificial intelligence. The main research areas are the following: (a) data analysis with an emphasis on text, web and cross modal data, (b) scalable real-time data analysis, (c) visualization of complex data, (d) semantic technologies and (e) language technologies.

9. **Irish Computer Society** was founded in 1967 as the professional body representing information and communication technology professionals in Ireland. Its objective is to promote professional information and computer skills. The ICS is a member of the Council of European Professional Informatics Societies. The Irish Computer Society hosts seminars, workshops and conferences on current topics of interest and also conducts surveys of trends in the industry. National events include the National Data Protection Conference, the Public Sector IT Conference, the Leaders Conference, and in 2014 they ran the first Tech Week Ireland with 42,000 people taking part all over Ireland. The ICS has also made representations and submissions to government committees and forums where ICT knowledge is beneficial and on areas of concern for ICT professionals. The ICS is a nomination body for the Industrial and Commercial Panel for Seanad Éireann. The ICS promotes ICT skills in schools by information campaigns and sponsoring competitions.

10. **The Insight Centre for Data Analytics** is one of Europe’s largest data analytics research organisations, with 350 researchers, more than 40 industry partners and over €88 million of funding. Insight is made up of four main centres: Insight@DCU, Insight@NUI Galway, Insight@UCC and Insight@UCD as well as a number of affiliated bodies. Each of Insight’s main centres has a long track record of data analytics research. In November 2013 they came together under Science Foundation Ireland as Insight. The size of the centre allows for cooperation on a large scale, which
enables the organisation to compete for funding and opportunities at a much higher level than was previously possible.
Annex B- List of Kick-off meeting participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Organisation</th>
<th>E-mail</th>
<th>address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jardim-Goncalves Ricardo</td>
<td>UNINOVA, New University of Lisbon</td>
<td><a href="mailto:rg@uninova.pt">rg@uninova.pt</a></td>
<td></td>
</tr>
<tr>
<td>Mary Cleary</td>
<td>Irish Computer Society</td>
<td><a href="mailto:mary@ics.ie">mary@ics.ie</a></td>
<td></td>
</tr>
<tr>
<td>Mats Ahlin</td>
<td>SIS, Swedish Standards Institute</td>
<td><a href="mailto:Mats.ahlin@sis.se">Mats.ahlin@sis.se</a></td>
<td></td>
</tr>
<tr>
<td>Dudley Dolan</td>
<td>Q Validus/ AQUASmart</td>
<td><a href="mailto:dudley.dolan@eircom.net">dudley.dolan@eircom.net</a></td>
<td></td>
</tr>
<tr>
<td>Terry Landers</td>
<td>Microsoft</td>
<td><a href="mailto:Terry.Landers@microsoft.com">Terry.Landers@microsoft.com</a></td>
<td></td>
</tr>
<tr>
<td>Aikaterini Sylla</td>
<td>ESCO</td>
<td><a href="mailto:aikaterini.sylla@esco-secretariat.eu">aikaterini.sylla@esco-secretariat.eu</a></td>
<td></td>
</tr>
<tr>
<td>Garry Cleere</td>
<td>Q-VALIDUS</td>
<td><a href="mailto:garry.cleere@q-validus.com">garry.cleere@q-validus.com</a></td>
<td></td>
</tr>
<tr>
<td>Alina Iatan</td>
<td>CEN CENELEC</td>
<td><a href="mailto:aiatan@cencenelec.eu">aiatan@cencenelec.eu</a></td>
<td></td>
</tr>
<tr>
<td>Gary McManus</td>
<td>TSSG</td>
<td><a href="mailto:gmcmanus@tssg.org">gmcmanus@tssg.org</a></td>
<td></td>
</tr>
</tbody>
</table>