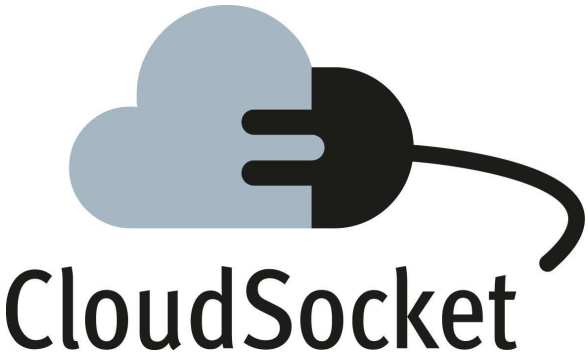

DATA MANAGEMENT PLAN

D1.2

Editor Name	Robert Woitsch (BOC)
Submission Date	June 30, 2015
Version	1.0
State	FINAL
Confidentially Level	PU

Co-funded by the Horizon 2020
Framework Programme of the European Union

EXECUTIVE SUMMARY

The research of CloudSocket focuses on conceptual and semantic models that can be applied to realise a smart alignment between domain specific business models and IT-Cloud deployable workflows. Hence, novel data set are expected in the domain of executable BPaaS bundles – for simplicity reasons can be seen as deployable workflows, and BPaaS business specifications – for simplicity reasons can be seen as business process models.

Those sample data from the business process domain as well as from the IT-Cloud domain provide a better understanding of the fundamental differences of those worlds, and demonstrate that although both models are expressed in one common standard called BPMN, the information that is provided within this standardised format is different.

Hence those data set may act for education or demonstration purpose, but are also intended to act as a start-up package for CloudSocket brokers that aim to create their own business process and IT-Cloud alignment platform.

The smart alignment of the two aforementioned domains – business and IT-Cloud – is another key outcome to be achieved. Hybrid knowledge management combines machine interpretable formats like RDFS, executable DMN or executable BPMN with human interpretable formats like concept maps, DMN or BPMN. Hence, a smart business and IT-Cloud alignment that is based on human and machine interpretable knowledge is introduced in form of a data set.

While additional publications and software are provided in separate distribution channels, it is obvious that in order to use the aforementioned data sets, relevant publication and software from the above will be provided. Most of the open access data sets are models, hence rely on corresponding modelling tools. Those models can be interpreted by using the openly distributed modelling tools on ADOxx.org or the OMiLAB, both in the form of community platforms for the project exploitation.

Hence five data sets have been identified:

- BPaaS Workflows: Executable and deployable workflows are called “BPaaS bundles”, which are sold via marketplaces as BPaaS. Samples and initial start-up configuration for a CloudSocket will be provided.
- BPaaS Business Process: Domain-specific and business compliant BPaaS business process models represent the business view to corresponding executable and deployable workflows. Samples and initial start-up configuration of a CloudSocket will be provided.
- Business and IT-Cloud Alignment Models: Hybrid knowledge representation that are partly human-readable and partly machine interpretable to identify the correct BPaaS Workflow for particular BPaaS Business Process.
- Mining of adaptation rules and deployment suggestions from semantic process logs: Raw data from business process logs and monitoring repositories are used and gathered in a semantic repository exploited via conceptual BPaaS analysis to generate adaptation rules that can drive the adaptation of a BPaaS as well as deployment suggestions that can be followed as important design knowledge for better allocating the resources to be used by a BPaaS.
- Meta Meta-Model enabling the incorporation of different script-based workflow dialects.

Following good practice, also CloudSocket will start distributing its data set via OpenAIRE.

The update of this Data Management Plan will be performed within the yearly review cycles.

Copyright © 2015 BOC and other members of the CloudSocket Consortium

PROJECT CONTEXT

Workpackage	WP1: Project Management
Task	T1.2: Project Management and Data Management
Dependencies	All WPs with respect to Data Management

Contributors and Reviewers

Contributors	Reviewers
Robert Woitsch (BOC), Julia Wels (ATOS), Antonio Gallo (FHOSTER), Vlad Mihnea (YMENS), Jürgen Jähnert,(BWCON), Yhongzeng Liang (BWCON), Emanuele Bellini (MATHEMA), Stefan Wesner (UULM), Jörg Domaschka (UULM),Dimitris Plexousakis (FORTH), Kyriakos Kritikos (FORTH), Sabrina Kurjakovic (FHNW)	All Contributors

Approved by: Robert Woitsch, as Project Coordinator (WP1 Leader)

Version History

Version	Date	Authors	Sections Affected
0.1	June 01, 2015	Robert Woitsch (BOC)	Initial setup of all sections
		Dimitris Plexousakis (FORTH), Stefan Wesner (UULM),	Overall templates, statements and guiding input
0.4	June 15, 2015	Kyriakos Kritikos (FORTH), Sabrina Kurjakovic (FHNW)	Contribution to data set templates
0.5	June 24, 2015	Vlad Mihnea (YMENS), Jürgen Jähnert,(BWCON) Yhongzeng Liang (BWCON), Jörg Domaschka (UULM), Kyriakos Kritikos (FORTH), Emanuele Bellini (MATHEMA), Antonio Gallo (FHOSTER), Julia Wels (ATOS)	Contribution and Review of all data sets
0.9	June 26, 2015	all	Final review
1.0	June 29	Robert Woitsch (BOC)	Finalisation

Copyright © 2015 BOC and other members of the CloudSocket Consortium
www.cloudsocket.eu

Copyright Statement – Restricted Content

This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

This is a restricted deliverable that is provided to the community under the license Attribution-No Derivative Works 3.0 Unported defined by creative commons http://creativecommons.org

You are free:

	to share within the restricted community — to copy, distribute and transmit the work within the restricted community
Under the following conditions:	
	Attribution — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
	No Derivative Works — You may not alter, transform, or build upon this work.

With the understanding that:

Waiver — Any of the above conditions can be waived if you get permission from the copyright holder.

Other Rights — In no way are any of the following rights affected by the license:

- Your fair dealing or fair use rights;
- The author's moral rights;
- Rights other persons may have either in the work itself or in how the work is used, such as publicity or privacy rights.

Notice — For any reuse or distribution, you must make clear to others the license terms of this work.
This is a human-readable summary of the Legal Code available online at:

http://creativecommons.org/licenses/by-nd/3.0/

Copyright © 2015 BOC and other members of the CloudSocket Consortium

TABLE OF CONTENT

Copyright © 2015 BOC and other members of the CloudSocket Consortium

LIST OF FIGURES

Copyright © 2015 BOC and other members of the CloudSocket Consortium

LIST OF TABLES

Copyright © 2015 BOC and other members of the CloudSocket Consortium

1 INTRODUCTION AND PROBLEM STATEMENT

The Data Management Plan is expected to describe the management of research data that are produced in CloudSocket according to certain rules that are laid out in a set of EC templates and guidelines.

In order to provide a complete document, those guidelines and templates are briefly introduced to explain the intention of this Data Management Plan. Then the specific aspects of the CloudSocket Data Management Plan are introduced before the structure of the document is introduced.

1.1 Introduction of Data Management Plan according to EC Templates

The Data Management Plan (DMP) is introduced in "Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020" (EC 2013 a) from page 10:

"*Data management and data management plans (DMP): A data management plan is a document outlining how the research data collected or generated will be handled during a research project, and after it is completed, describing what data will be collected / 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 (see separate Guidelines on Data Management in Horizon 2020 on the Participant Portal).*

…

The use of a detailed data management plan covering individual datasets is required for funded projects participating in the Open Research Data Pilot. A first version of the data management plan must be provided as an early deliverable within six months of the project.

Other funded projects that do not participate in the Pilot are also invited to submit a Data Management Plan if relevant for their planned research, but this is not compulsory.

Further information on data management, DMPs and a DMP template can be found in the Guide on Data Management which is available on the Participant Portal."

This Guide on Data Management introduces the DMP in the following way (EC 2013 b):

"Where relevant, applicants must provide a short, general outline of their policy for data management, including the following issues:

- *What types of data will the project generate/collect?*
- *What standards will be used?*
- *How will this data be exploited and/or shared/made accessible for verification and re-use? If data cannot be made available, explain why.*
- *How will this data be curated and preserved?"*

In order to detail those questions, there is a template provided and control questions to be answered for each data set in the Open Access Project OpenAIRE (OpenAIRE 2015).

Copyright © 2015 BOC and other members of the CloudSocket Consortium

1.2 CloudSocket Data Management

In order to apply the aforementioned templates and guidelines onto CloudSocket results, those data sets need to be identified and analyzed. Hence the relevant data sets are identified in this document based on the project result table as well as on the status of project work.

This initial Data Management Plan introduces the potential data sets and elaborates in more detail on which among those data can be provided. As this DMP is expected to be updated regularly, the last chapter introduces the proposal to update this DMP at every 12 month reporting period.

1.3 Document Structure

After this introduction, the document presents the list of identified data sets in chapter two. The third chapter applies the aforementioned data set description templates onto each of the identified data sets. The fourth and last chapter provides a recap and discusses how this DMP is planned to be continuously updated.

Copyright © 2015 BOC and other members of the CloudSocket Consortium

2 IDENTIFICATION OF DATA SETS

The identification of data sets is based on the project result and dissemination / exploitation strategy as well as on the list of exploitation items. Both lists are defined in the Description of Action (DoA) of CloudSocket (CloudSocket DoA, 2015).

2.1 Identification of Result Items

In the following sections a brief outline on planned activities and goals is outlined

WP	Research Results	Dissemination / Eploitation Management		Path	Result Items	Partner	Expected release time
WP2	Awareness Raising Cloud Transformation Framework	**Project internal Knowledge Management**	**Diessemination**	Publications **Gold Access**	1) Awareness Raising 2) Cloud Transformation Framework 3) BPaaS Design Specificaiton 4) BPaaS Allocation Specification 5) BPaaS Execution Specificaiton 6) BPaaS Evaluation Specification	1) all 2) 1,8,9 3) 1,9 4) 4,7 5) 2,7 6) 1,8	1) M 1 2) M 6 3) M 18 4) M 24 5) M 24 6) M 30
WP3	BPaaS Design Specificaiton BPaaS Allocation Specification BPaaS Execution Specificaiton BPaaS Evaluation Specification			Publication **Green Access**	1-6) Additional publications, which have not been able to publish under gold access.	1-6) all	
WP4	BPaaS Design Environment BPaaS Allocation Environment BPaaS Execution Environment BPaaS Evaluation Environment BPaaS Marketplace			Open Access on **Research Data**	7) Anonyminised CloudSocket Models 8) Anonyminised BPaaS Workflows	7) 1,4,5,6 8) 2,3,5,6	10) M 18 11) M 30
WP5	Use Case Challenges CloudSocket Models BPaaS Workflows CloudSocket Demonstration		**Exploitation**	**Commericalisation** through result usage within enterprise	9) BPaaS Marketplace - YMENS marketplace 10) CloudSocket for Business Incubator 11) CloudSocket for Robotics Cluster	9) 3 10) 5 11) 6	9) M36 + 10) M36 + 11) M36 +
WP6	Validation and Experience Financial Impact Feedback			**Coommercialisation** through incorporation in products and services	12) BPaaS Design Environment - ADONIS 13) BPaaS Allocation Environment - Livebase 14) BPaaS Execution Environment - yourBPM 15) BPaaS Evaluation Environment - ADONIS	12) 1 13) 4 14) 2 15) 1	12) M36 + 13) M36 + 14) M36 + 15) M36 +

Figure 1 CloudSocket Exploitation and Dissemination Items Management Template

Besides the expected project results in correspondence with the work packages and the partners, each partner formulated in more detail the expected exploitation item.

2.2 Identification of Data Sets

Analysing the identified results and open access items, this Data Management Plan proceeds by focusing on the sample and reference models as identified open data sets.

Publications are managed in the dissemination work package (WP7) and the corresponding deliverables and management tasks. Software is developed and managed in the implementation work package (WP 4) and its distribution and usage is handled in the execution work package (WP 8).

In the following model are elaborated in more detail according to the Data Management Plan template, whereas publications and software are for completeness reasons briefly addressed.

Copyright © 2015 BOC and other members of the CloudSocket Consortium

3 INITIAL SPECIFICATION OF DATA SETS

This section first introduces the data set template provided by the EC in (EC 2013a), which is then applied on to the aforementioned list of data sets.

Initial Data Set Template	
Data set reference name:	Identifier for the data set to be produced.
Data set description	Description of the data that will be generated or collected, its origin (in case it is collected), nature and scale and to whom it could be useful, and whether it underpins a scientific publication. Information on the existences (or not) of similar data and the possibilities for integration and reuse.
Standards and metadata	Reference to existing suitable standards of the discipline. If these do not exist, an outline on how and what metadata will be created.
Data sharing	Description of how data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling reuse, and definition of whether access will be widely open or restricted to specific groups. Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.).
Archiving and preservation (including storage and backup)	Description of the procedure that will be put in place for long-term preservation of the data. Indication of how long the data should be preserved, what is its approximated end, volume, what the associated costs are and how these are planned to be covered.
Additional Data Set Explanation	
Discoverable	Are the data and associated software produced and / or used in the project discoverable (and readily located), identifiable by means of a standards identification mechanism (e.g. Digital Object Identifier)?
Accessible	Are the data and associated software produced and / or used in the project accessible and in what modalities, scope, licenses (e.g. licencing framework for research and education, embargo periods, commercial exploitation, etc.)?
Assessable and intelligible	Are the data and associated software produced and / or used in the project assessable for and intelligible to third parties in contexts such as scientific scrutiny and peer review (e.g. are the minimal datasets handled together with scientific papers for the purpose of peer review, are data is provided in a way that judgments can be made about their reliability and the competence of those who created them)?
Usable beyond the original purpose for which it was collected	Are the data and associated software produced and / or used in the project useable by third parties even long time after the collection of the data (e.g. is the data safely stored in certified repositories for long term preservation and curation; is it stored together with the minimum software, metadata and documentation to make it useful,; is the data useful for the wider public needs and usable of the likely purpose of non-specialists.
Interoperable to specific quality standards	Are the data and associated software produced and / or used in the project interoperable allowing data exchange between researchers, institutions, organisations, countries, etc. (e.g. adhering to standards for data annotation, data exchange, compliant with available software applications, and allowing re-combinations with different datasets from different origins)?

Table 1 Data Set Description Template

Copyright © 2015 BOC and other members of the CloudSocket Consortium

3.1 Publication

Management and dissemination of publication is performed in the dissemination work package (WP 7) and described in the corresponding deliverables. For completeness reasons the following statements are provided.

Publications within the topics of (a) Awareness Raising, (b) BPaaS Allocation Specification, (c) BPaaS Design Specification, (d) BPaaS Evaluation Specification, (e) BPaaS Execution Specification, (f) Cloud Transformation Framework, (g) Design patterns to realize fault tolerance/high availability in Clouds as well as on (h) Knowledge-Based IT Alignment are published at networking events, conference, journals or on the Web to introduce the overall idea and explain certain aspects in more detail.

Publications will be stored in an Open Access Library (OAL). Currently the OpenAIREplus project with its data storage Zenodo is a potential candidate for such publishing. Beside OAL, such publications will be presented in research databases – such as ResearchGate or Academia.edu – as well as in corresponding literature database.

3.2 Software

Management and exploitation of software is performed in the exploitation work package (WP 8) and described in the corresponding deliverables. For completeness reasons the following statement is provided.

Distribution of software mainly depends on the business model. Within the consortium there are: (a) licenced software with a commercial licence model, (b) open access / source software that is distributed for a particular audience with a particular aim, as well as (c) open source software that is distributed according to the selected licence.

Commercially licenced software is announced via Web page and publications but distributed through bilateral distribution channel. Open Access software is distributed via corresponding community channels – like ADOxx.org. Open Source software could be published via OpenAIREplus or additionally via open-source repositories, like github.com, in order not only to provide the software but also facilitate its further development and extension and ensure a long-term availability of the code.

For all software stored in public repositories, the usual meta-data such as authors and commit history will be maintained by the repository. Further, dependencies to operating systems, run-time environments as well as third-party tools and libraries including their versions will be documented. Where possible the correct versions of these third-party tools will be stored with the source code. In addition, for all of the software in public repositories documentation or tutorials will be created that allows third parties to install and run the software.

Software, or software related publications published in OpenAIREplus, will be associated to a specific DOI. In addition, the software that is available in GitHub and similar public repositories will also be discoverable by any search engine, such as the one provided by Google. Through a Git URL, software could then be accessed/ retrieved. Particularly, software available on GitHub will be assigned DOIs via Zenodo (Github, 2015). The institutional repository of the University of Ulm automatically assigns URNs and - starting in 2016 - also DOIs to its resources.

Copyright © 2015 BOC and other members of the CloudSocket Consortium

3.3 Data Set as Models

This project realises smart Business Process and IT-Cloud alignment by applying conceptual models and semantic models. Hence, relevant data sets that represent core findings of the CloudSocket research are reflected in extensions to well-known domain (business process, case management, decision management) and / or IT (workflow, deployment, SLA) models, the integrative meta models (Concept Reference Ontology) as well as in models that enable the smart alignment (Semantic Mapping, Rule Models, Ontologies).

In the following the meta-models and sample models are listed as expected data set by filling all relevant details of the basic template exploited.

3.3.1 Meta-meta model enabling the incorporation of different script-based workflow dialects

Initial Data Set Template	
Data set reference name:	Workflow Meta-meta model
Data set description	A meta-meta model as well as a software to enable the incorporation and subsequent use of different script-based workflow dialects for workflow specification
Standards and metadata	The meta-meta model could be specified through known specification formats, such as ontologies or UML.
Data sharing	The meta-meta model will be provided in an individual manner with respect to the respective software supporting it. Open-access repositories like OpenAIREplus and github will be exploited for the sharing of the model and respective software.
Archiving and preservation (including storage and backup)	This will be the responsibility of the open-access repositories for free with no charges to FORTH.
Additional Data Set Explanation	
Discoverable	Both artefacts will be accessible through a DOI and a specific http URL.
Accessible	Both artefacts will be associated to an open-source licence to be determined in the near future.
Assessable and intelligible	
Usable beyond the original purpose for which it was collected	As part of open-access repositories, the model can be used even after the end of the project. In fact, they can also be further developed and extended.
Interoperable to specific quality standards	Not applicable

Table 2 Data Set Fact Sheet – Workflow Meta Model

Copyright © 2015 BOC and other members of the CloudSocket Consortium

3.3.2 BPaaS Bundle (deployable Workflows)

Initial Data Set Template	
Data set reference name:	BPaaS Bundle (deployable Workflow)
Data set description	BPaaS Bundle are executable workflows that are bundled with deployment information in order to be ready for Cloud deployment and execution. A critical mass of bundles is essential for attracting CloudBrokers to apply the CloudSocket idea. Hence, a set of ready to use bundles will be provided to ease the creation of so-called CloudSockets.
Standards and metadata	BPMN notation for the Activiti workflow engine will be used. Deployment information and meta data for BPaaS bundles will be similar to the format used in the PaaSage project (PaaSage, 2015), but with essential changes. The BPMN and deployment meta will be specified to document the format. It is expected that decision models as well as SLO information will be incorporated. In that case most common standards such as DMN and prominent quality description languages like OWL-Q will be used.
Data sharing	Open-access repositories like OpenAIREplus and git will be exploited for the sharing of the model and respective software. Modelling language relevant communities such as Open Models (OMiLAB.org) may be additionally be used.
Archiving and preservation (including storage and backup)	This will be the responsibility of the open-access repositories or the community provider.
Additional Data Set Explanation	
Discoverable	Artefacts will be accessible through a DOI and a specific http URL.
Accessible	Artefacts will be associated to an open-source licence to be determined in the near future.
Assessable and intelligible	N.A.
Usable beyond the original purpose for which it was collected	As part of open-access repositories, the model can be used even after the end of the project. In fact, they can also be further developed and extended.
Interoperable to specific quality standards	N.A.

Table 3 4 Data Set Fact Sheet – BPaaS Workflows

Copyright © 2015 BOC and other members of the CloudSocket Consortium

3.3.3 BPaaS Business Processes

Initial Data Set Template	
Data set reference name:	BPaaS Business Processes
Data set description	Those Business Processes are the counter part of aforementioned Workflows, as they describe the business view as well as business and domain related KPIs. Based on the business process classification frameworks, CloudSocket generates business processes for SMEs that can be transformed into the Cloud. This list of business processes is essential to ease the take up of new Business Process Brokers.
Standards and metadata	Business Processes are commonly represented in BPMN 2.0 format. In addition we expect Decision Model Notation to be the correct choice to introduce flexibility into the orchestration of a business process. Furthermore, KPI relevant information is required within the business process. Hence, standards like MathXML or prominent quality description languages like OWL-Q (extended towards specifying KPIs) or self-defined indicator extensions are expected.
Data sharing	Open-access repositories like OpenAIREplus and git will be exploited for the sharing of the model and respective software. Modelling language relevant communities such as Open Models (OMiLAB.org) may be additionally be used.
Archiving and preservation (including storage and backup)	This will be the responsibility of the open-access repositories or the community provider.
Additional Data Set Explanation	
Discoverable	Artefacts will be accessible through a DOI and a specific http URL.
Accessible	Artefacts will be associated to an open-source licence to be determined in the near future.
Assessable and intelligible	N.A.
Usable beyond the original purpose for which it was collected	As part of open-access repositories, the business process models can be used even after the end of the project. In fact, they can also be further developed and extended.
Interoperable to specific quality standards	N.A.

Table 5 Data Set Fact Sheet – BPaaS Business Processes

Copyright © 2015 BOC and other members of the CloudSocket Consortium

3.3.4 Business and IT-Cloud Alignment Models

Initial Data Set Template	
Data set reference name:	Business and IT-Cloud Alignment Models
Data set description	Semantic Lifting, textual annotation or Conceptual Reference ontologies are instruments that enable the alignment between two different conceptual model domains. This data set enables the alignment of domain specific business process with IT-cloud specific workflows. It is structured based on the Zachmann dimensions and completed with alignment concepts that had been worked out as a CloudSocket result.
Standards and metadata	Expected standard is RDFS.
Data sharing	Open-access repositories like OpenAIREplus and git will be exploited for the sharing of the model and respective software. Modelling language relevant communities such as Open Models (OMiLAB.org) may be additionally be used.
Archiving and preservation (including storage and backup)	This will be the responsibility of the open-access repositories or the community provider.
Additional Data Set Explanation	
Discoverable	Artefacts will be accessible through a DOI and a specific http URL.
Accessible	Artefacts will be associated to an open-source licence to be determined in the near future.
Assessable and intelligible	N.A.
Usable beyond the original purpose for which it was collected	As part of open-access repositories, the model can be used even after the end of the project. In fact, they can also be further developed and extended.
Interoperable to specific quality standards	N.A.

Table 6 Data Set Fact Sheet – Business and IT-Cloud Alignment

Copyright © 2015 BOC and other members of the CloudSocket Consortium

3.3.5 Adaptation rules and Deployment Suggestions mined from semantic process logs

Initial Data Set Template	
Data set reference name:	Adaptation rules and deployment suggestions mined from semantic process logs
Data set description	A set of generic and domain-dependent/application-specific adaptation rules which can be exploited for adapting services, workflows, BPaaS and components. A set of deployment suggestions for different types of BPaaS.
Standards and metadata	There are standards which can be used for the description of rules, including Production Rule Representation (PRP) proposed by OMG. Such standards can be embraced in order to transform the rules generated from the internal format of the platform to the standardized one. Deployment suggestions could be specified as DMN decision tables (e.g., indicating which deployment is best or worst for a specific BPaaS) or as CAMEL/CloudML deployment descriptions (based on the CAMEL meta-model proposed in the context of the PaaSage project) mapped to BPaaS
Data sharing	Some of the rules and deployment suggestions could be made accessible to the public. In this sense, they could be published in platforms/repositories such as OpenAIREplus.
Archiving and preservation (including storage and backup)	The management (archiving & preservation) of the public rules and suggestions will be delegated to the open-access repository with no charge. The management of the whole set of rules and suggestions (public and private) will be a complete responsibility of FORTH (as well as of course the update of public rules in the open-access repository).
Additional Data Set Explanation	
Discoverable	The public rules and suggestions will be associated to a specific DOI.
Accessible	The public rules and suggestions will be accessible via a particular open-source licence to be determined in the near future.
Assessable and intelligible	
Usable beyond the original purpose for which it was collected	As the public rules and suggestions will be accessible through the open-access repository, they can be extended and further developed by any interesting party even after the end of the project (so the rules and suggestions will still be usable beyond the original purpose especially as they can apply to different process/workflow/service management frameworks to drive the respective adaptations).
Interoperable to specific quality standards	Not applicable.

Table 7 Data Set Fact Sheet – BPaaS Process Monitor Logs

Copyright © 2015 BOC and other members of the CloudSocket Consortium

4 SUMMARY AND CONCLUSIONS

The research of CloudSocket focuses on conceptual and semantic models that can be applied to realise a smart alignment between domain specific business models and IT-Cloud deployable workflows. Hence, novel data set are expected in the domain of executable BPaaS bundles – called for simplicity reasons workflows, and BPaaS business specifications – called for simplicity reasons business processes.

Those sample data from the business process domain as well as from the IT-Cloud domain provide a better understanding of the fundamental differences of those worlds, and demonstrate that although both models are expressed in one common standard called BPMN, the information that is provided within this standardised format is different.

Hence those data set may act for education or demonstration purpose, but are also intended to act as a start up package for CloudSocket brokers that aim to create their own business process and IT-cloud alignment platform.

The smart alignment of the two aforementioned domains – business and IT-Cloud –is another key outcome to be achieved. Hybrid knowledge management combines machine interpretable formats like RDFS, DMN or BPMN with human interpretable formats like concept maps or DMN or BPMN. Hence, a smart business and IT-Cloud alignment that is based on human and machine interpretable knowledge is introduced in form of a data set.

In addition the workflow meta-model is also considered as an interesting aspect to accommodate for the heterogeneity in the workflow language space to be included in this data management plan.

It is obvious that in order to use the aforementioned data sets, relevant publication and software will be provided which may have been previously published and communicated via separate distribution channels. Most of the open access data sets are models, hence rely on corresponding modelling tools. Those models can be interpreted by using the openly distributed modelling tools on ADOxx.org (ADOxx.org, 2015) or the OMiLAB (OMiLAB. 2015), both in the form of community platforms for the project exploitation.

Following good practice, CloudSocket will also start distributing its data set via OpenAIRE.

The update of this DMP will be performed within the yearly review cycles.

Copyright © 2015 BOC and other members of the CloudSocket Consortium

5 REFERENCES

EC 2013 a; European Commission, Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020, Version 1.0, 11 December 2013, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf, access: 01.06.2015

EC 2013 b, European Commission, Guidelines on Data Management in Horizon 2020, Version 1.0, 11 December 2013, http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf, access 01.06.2015

OpenAIRE 2015, https://www.openaire.eu/, access: 01.06.2015

CloudSocket DoA 2015, CloudSocket Description of Action, 2015

OMiLAB 2015, http://www.omilab.org/, access: 24.06.2015

ResearchGate 2015, https://www.researchgate.net/, access: 24.06.2015

PaaSage 2015e, http://www.paasage.eu/, access: 24.06.2015

Github 2015, https://guides.github.com/activities/citable-code/, access: 24.06.2015

Copyright © 2015 BOC and other members of the CloudSocket Consortium