ICS-TCS Integration Guidelines

Handbook for TCS integration: Level-2

Prepared by

EPOS-IP WP6 &WP7 teams

Project:	EPOS-IP (Horizon2020 – Project number: 676564)
Work Package(s):	WP6, WP7 and TCS WPs
Task(s):	Tasks regarding ICS-TCS integration and interoperability
Document type:	ICS TCS Integration Guidelines – Level-2
Document file name:	EPOS-ICS-TCS-Integration-Guidelines-Level-2-20150917.doc
Document title:	ICS-TCS Integration: Handbook for TCS Integration: Level-2
Purpose:	For internal use between Transverse WPs and the TCS WPs
Author(s):	EPOS – IP WP6 & WP7 Teams
Version:	1.0
Date:	20150917
Contact:	http://www.epos-eu.org/ http://eposwiki.bo.ingv.it/ daniele.bailo@ingv.it

September 2015

ICS-TCS Integration Guidelines

Handbook for TCS integration: Level-2

Prepared by

EPOS WP6 &WP7 teams

Table of contents

1 Why This Document

The purpose of this short document is to provide the TCS and other partners with the technical details and procedures required to achieve the integration between ICS and TCS. The ICS-TCS Integration Guidelines are intended as a handbook for TCS integration. A general introductory version (Level-1) has already been distributed within the EPOS community. This document is the second version (Level-2) of similar follow-up documents that will be distributed at later stages.

2 EPOS e-Architecture

Organization

The EPOS architecture has been designed to organize and manage the interactions among different EPOS actors and assets. To make it possible for the EPOS enterprise to work as a single, but distributed, sustainable research infrastructure, its architecture takes into account technical, governance, legal and financial issues. Four complementary elements form the infrastructure:

1. The **National Research Infrastructures (NRIs)** contribute to EPOS while being owned and managed at a national level and represent the basic EPOS data providers. These require significant economic resources, both in terms of construction and yearly operational costs, which are typically covered by national investments that must continue during EPOS implementation, construction and operation.
2. The **Thematic Core Services (TCS)** enable integration across specific scientific communities. They represent a governance framework where data and services are provided and where each community discusses its implementation and sustainability strategies as well as legal and ethical issues.
3. The **Integrated Core Services (ICS)** represents the e-infrastructure consisting of services that will allow access to multidisciplinary resources provided by the NRIs and TCS. These will include data and data products as well as, synthetic data from simulations, processing, and visualization tools. The ICS will be composed of the ICS-Central Hub (ICS-C) and distributed computational resources including also processing and visualisation services (ICS-D). ICS is the place where integration occurs.
4. The **Executive and Coordination Office (ECO)** is the EPOS Headquarters and the legal seat (ERIC) of the distributed infrastructure governing the construction and operation of the ICS and coordinating the implementation of the TCS.

The European Research Infrastructure Consortium (ERIC) has been chosen by the Board of Governmental Representatives as the legal model for EPOS and is used in designing the Governance Model. This includes a General Assembly of members and an Executive Director, supported by a Coordination Office. A funding model has been designed that will support the sustainable construction and operation of the whole EPOS enterprise. The model includes complementary funding sources for each of the key EPOS elements.

Figure 1 describes the EPOS technical architecture organised in three layers.

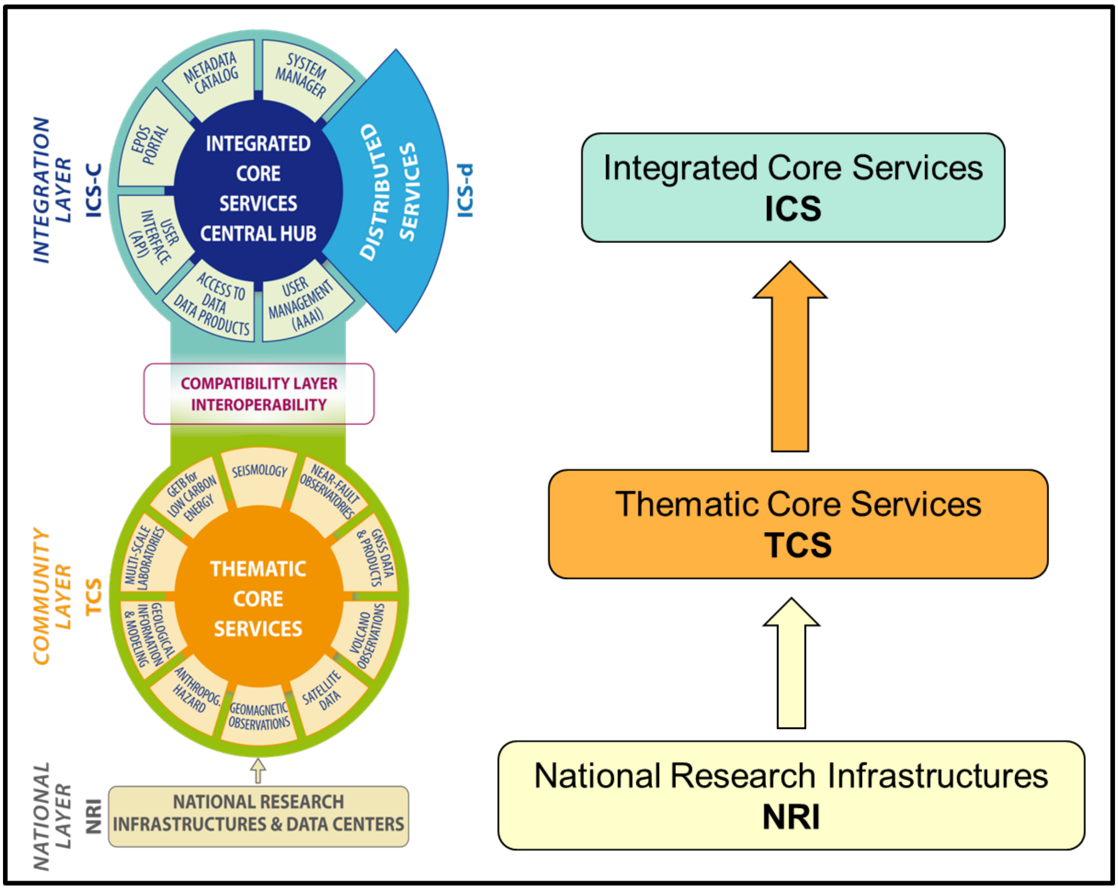

Figure 1: EPOS technical architecture. The diagram shows the three layers in which the EPOS components (institutions and services) have been organized: National Layer, Community Layer, Integration Layer including also an Interoperability Layer.

The main concept is that the EPOS TCS data and services are provided to the ICS (see Fig.1) by means of a communication layer called the *interoperability layer*, as shown in the functional architecture (Fig. 2). This layer contains all the technology to integrate data, data products, services and software (DDSS) from many scientific, thematic communities into the single integrated environment of the Integrated Core Services (ICS). The ICS represents the "core" of the whole e-infrastructure and those responsible for its implementation will provide the specification of the "*interoperability layer*". The ICS is conceptually a single, centralized facility but in practice is likely to be replicated (for resilience and performance) and localized for particular natural language groupings or legal jurisdictions.

Technical architecture

The ICS is made up of several, modular, interoperable building blocks (Fig. 2). The three layer structure adopted in the technical architecture of EPOS consists of the National Layer where the National Research Infrastructures provide the DDSS. Data providers in this layer are independent national institutions or organizations which have their own technical solutions that may (or may not) follow international standards in providing data and data products to the community. The second layer, Thematic Core Services (TCS) is the (European) Community Layer where community standards are applied to DDSS that are relevant to the specific thematic area of concern. The third (top) layer represents the integration of the DDSS that come from the TCSs, where high level international standards are applied. At this level metadata describing all DDSS need to be harmonized into a single metadata catalogue which is based on international standards. During the Preparatory Phase of the EPOS Project, a European metadata catalogue standard, CERIF (Common European Research Infrastructure Format), was tested and used for the prototype development. It has subsequently been adopted in the EPOS Implementation Phase. In order that the DDSS from the various TCSs can be converted into the chosen metadata catalogue standard (i.e. CERIF), there is a need for an additional layer where TCS data sets will be mapped and converted to CERIF. This is referred to as the "Interoperability layer". The various components of the ICS are now explained in more detail.

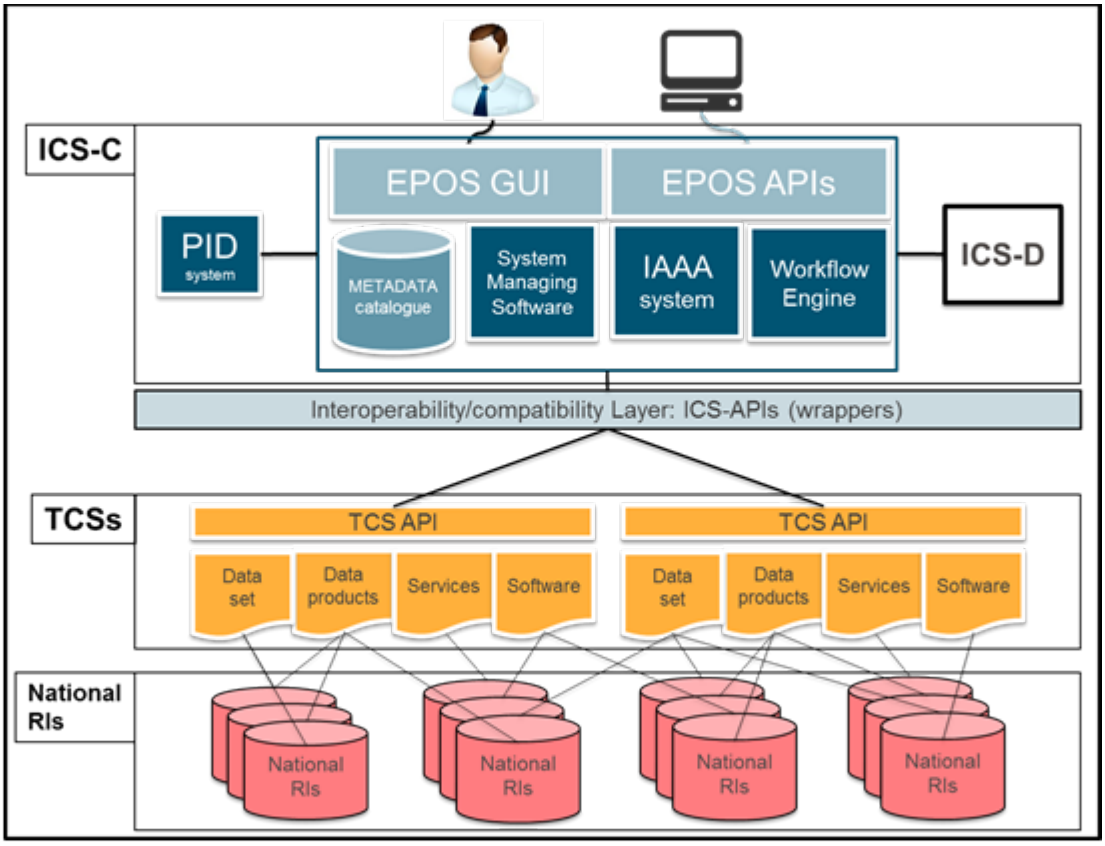

Figure 2: EPOS functional Architecture, describing the technical functional components of EPOS. It specifies for each layer the ICT modules and their function. At the ICS layer it describes the design of the integrating e-Infrastructure.

1) *Metadata catalogue*

Metadata describing the TCS DDSS are stored using the CERIF data model which differs from most metadata standards in that it (1) separates base entities from

linking entities thus providing a fully connected graph structure; (2) using the same syntax, stores the semantics associated with values of attributes both for base entities and (for role of the relationship) for linking entities, which also store the temporal duration of the validity of the linkage. This provides great power and flexibility. CERIF also (as a superset) interoperates with widely adopted metadata formats such as DC (Dublin Core), DCAT (Data Catalogue Vocabulary), CKAN (Comprehensive Knowledge Archive Framework), INSPIRE (the EC version of ISO 19115 for geospatial data) and others.
The metadata catalogue will also manage the semantics, in order to provide the meaning of the attribute values. CERIF stores the semantics in a 'semantic layer' referenced from the syntactic layer thus providing a single integrated semantic environment which is efficient because it uses standard IT (usually relational but all other database/processing environments may be used). CERIF interoperates with OWL (Web Ontology Language), SKOS (Simple Knowledge Organization System) and other semantic representation languages.
Metadata from the communities will be mapped to the metadata catalogue in order to create appropriate links between common concepts in different disciplines. This process involves the harmonization and interoperability of the various DDSS from the different TCSs through dedicated software modules. It requires TCS APIs for converting DDSS to the TCS specific metadata standard. It also requires ICS APIs (wrappers) to map and store this in the ICS metadata catalogue (i.e. CERIF). These TCS APIs and the corresponding ICS APIs collectively form the "interoperability layer", which is the link between the TCSs and the ICS.

2) System Managing Software

The system managing software will manage the metadata catalogue and all other modules (e.g. workflow engine and generally all the resources involved to satisfy the user requests).

3) Workflow engines and provenance

A key aspect of the semi-automatic composition of software to meet a user request is the provision of a workflow to link together the software services as they access appropriate data. Workflow engines available are many and each of them fits different use cases and architectures. We will take into account the computational models (i.e. from the Computational Earth Science community) to be supported and the communities' requirements. We can anticipate that, following the experience gained by the EPOS partners in initiatives such as VERCE[1] and ongoing work in RDA (Research Data Alliance), particular attention will be dedicated to cross platform streaming libraries.
CERIF can also provide provenance information since the linking entities associate with the role have, as attributes, both date/time start and date/time end. This handles versioning and – via the linking entity record –the relationship of one base entity instance (e.g. a dataset) to another. On the other hand, for a comprehensive traceability of the processes and agents that contributed to the generation of the research product, we foresee the integration in CERIF of the

---

[1] Virtual Earthquake and seismology Research Community in Europe e-science environment http://www.verce.eu/

W3C-PROV ontology. This will guarantee interoperability with other institutional data archives, fostering data preservation and curation across domains.

4) IAAA to data and computational resources (cloud, grid, HPC)

This module will manage and interoperate with all the major 'common' IAAA (Identification, Authentication, Authorisation, Accounting) services and standards from AAAI (Authentication, Authorisation, Accounting Infrastructure) such as SAML, OAuth, OpenID, X.509 and related products such as EduGAIN, Shibboleth, Kerberos and others and also for user directory services such as Microsoft Active Directory and LDAP. Addressing the IAAA in a satisfactory way is a challenge at the present stage, and is also being faced in other projects and initiatives following AAAI (e.g. AARC[2], EGI-Engage[3]), with which EPOS is collaborating. *The goal in this collaboration is to implement a smart* IAAA *mechanism which is able to hide from the user all the complexity of delegation-based* AAAI *mechanisms*.

5) ICS-D

As already described, Integrated Cores Services – Distributed (ICS-D) will include services from external computing facilities. These will include HPC (High Performance Computing) machines for modelling and simulation according to the requirements of the Computational Earth Science community, and HTC (High Throughput Computing) clusters for data intensive applications such as data mining. The data workflow will be managed by EPOS ICS-C in order to provide the end user with appropriate computational services, even though actual computations will be provided by ICS-D. Additional ICS-D services will provide visualization and processing capabilities. ICS-C will have to develop provisions for communicating with these external services in a seamless manner.

6) Web services / APIs

EPOS-IP, wherever possible, will use web services[4] as the main vehicle for software services, defining APIs, and implementing the best practices for a sound microservices architecture, so that workflows can be composed semi-automatically. Web services and APIs, together with appropriate mapping of metadata which will drive data convertors, will also be the driving technology for the "connection" of TCS with "ICS".

The detailed description of the whole EPOS e-infrastructure is out of the scope of this document. For further details and information follow the link to **EPOS-ICT summary** at:
http://www.epos-eu.org/assets/documents/WG7/EPOS-ICT-summary.pdf

3 Principles

In the following the high-level principles for the ICS development and the ICS-TCS interactions are explained.

[2] Authentication and Authorization for Research and Collaboration, https://aarc-project.eu/
[3] EGI-Engage project (Engaging the Research Community towards an Open Science Commons, https://www.egi.eu/about/egi-engage/
[4] Web services also act as GRID services or CLOUD services

Co-development (ICS and TCS)

The development of the ICS depends on end-user requirements and the DDSS provided by the TCS. The TCS are at different stages of maturity. For some – with little infrastructure to date - the adoption of the EPOS architecture is straightforward. For others – with several years (decades) of infrastructural investments already existing - a jointly agreed evolutionary plan to converge to interoperability with ICS will be adopted. The EPOS approach in this context is neither top-down nor bottom up: the main idea is to use the general architecture and follow a cooperative approach in the designing and development of the software to build the compatibility layer, which is the place where harmonisations and communications are achieved.
Such an approach also encourages EPOS to focus more on an architectural design for the system that is capable of being engineered and implemented collaboratively rather than using a top-down reference model approach.
This would facilitate cross-disciplinary access and utilization of the data for the scientific purpose.

Do not reinvent the wheel

a. *Reuse local technologies*
The co-development philosophy maximises re-use of existing software services, data availability and resources.
b. *Do not build a supercomputer: build an integration with an ICS-D service provider*
EPOS will have some of its own computing resources in order to provide the 'uniform view' over EPOS entities. However, to facilitate intensive processing, EPOS will provide – subject to authorisation – access to appropriate computing facilities including HPC (High Performance Computing) machines for modelling and simulation, and HTC (High Throughput Computing) clusters for data intensive applications such as data mining. Other ICD services will include visualization and processing services. These facilities are known collectively as ICS-D (Integrated Core Services – Distributed) and will be provided by external service providers.

Microservices approach

The Microservices architecture approach envisages small "micro" services dedicated to the execution of a specific class of tasks, which have high reliability. EPOS will either take existing software services and 'wrap' them for EPOS use or build new services complying with EPOS architectural standards. The aim is to have micro services with defined interfaces which can be composed together to form a software stack able to address unpredictable user request.

Clear long-term technical goals, but iterative short-term approach

The overall architecture of EPOS is clearly defined and agreed through EPOS-PP. However, its implementation in EPOS-IP would require a step-by-step approach to build a reliable system environment to meet the requirements of end-users and their communities. To this purpose iterative work cycles both for ICS developments and for ICS-TCS communication have been set up.

4 Recommendations to TCS

R.1 Each type of DDSS delivered must be associated with descriptive metadata that enables efficient data discovery and contextualisation (which enables a user to determine relevance and quality). The metadata should be, when possible, an approved (e.g. OSI) or *de facto* standard and ideally already in CERIF.

R.2 Each type of DDSS delivered (and its associated metadata) should be accessible via web services and/or APIs.

R.3 Each web service or API to access DDSS should be, when possible, based on approved (e.g. OSI) or *de facto* standards. Before building any new web services, a preliminary communication with the ICS team should be carried out in order to verify that it is suitable. The ICS development team will happily help and provide advice. In order to optimize efforts, TCS should also carry out a preliminary check with other communities to ensure the proposed web service has not already been implemented.

R.4 If a TCS needs to develop a metadata standard for its community, a preliminary check should be carried out with the ICS team and other communities, in order to verify that elements of their data has not already been described by another community.

R.5 The integration among TCS should be as deep as possible: ideally, every single community should adopt (intra-community) a common strategy to share tools, data formats, API, data management rules, data analysis and processing frameworks to a common cross-EPOS standard. As a consequence, (a) efforts to build news services or to improve old ones will be minimized and (b) it will be simpler to integrate diverse data products at ICS level.

R.6 Recommending IAAA standards is premature at this stage but will be communicated when appropriate solutions are found. It is important however that all TCSs reflect upon their needs in terms of IAAA.

5 TCS Integration

TCS generic architecture

The different Thematic Core Services (TCS) have varying degrees of maturity in their development. Therefore, it is not possible to deal with TCS as if they are all equal and homogeneous. Some TCS have a very specific services architecture based on years of experience in that specific domain. Other TCS don't have any history of developing services. Some TCS rely on a federated architecture, others on a single sited data centre.

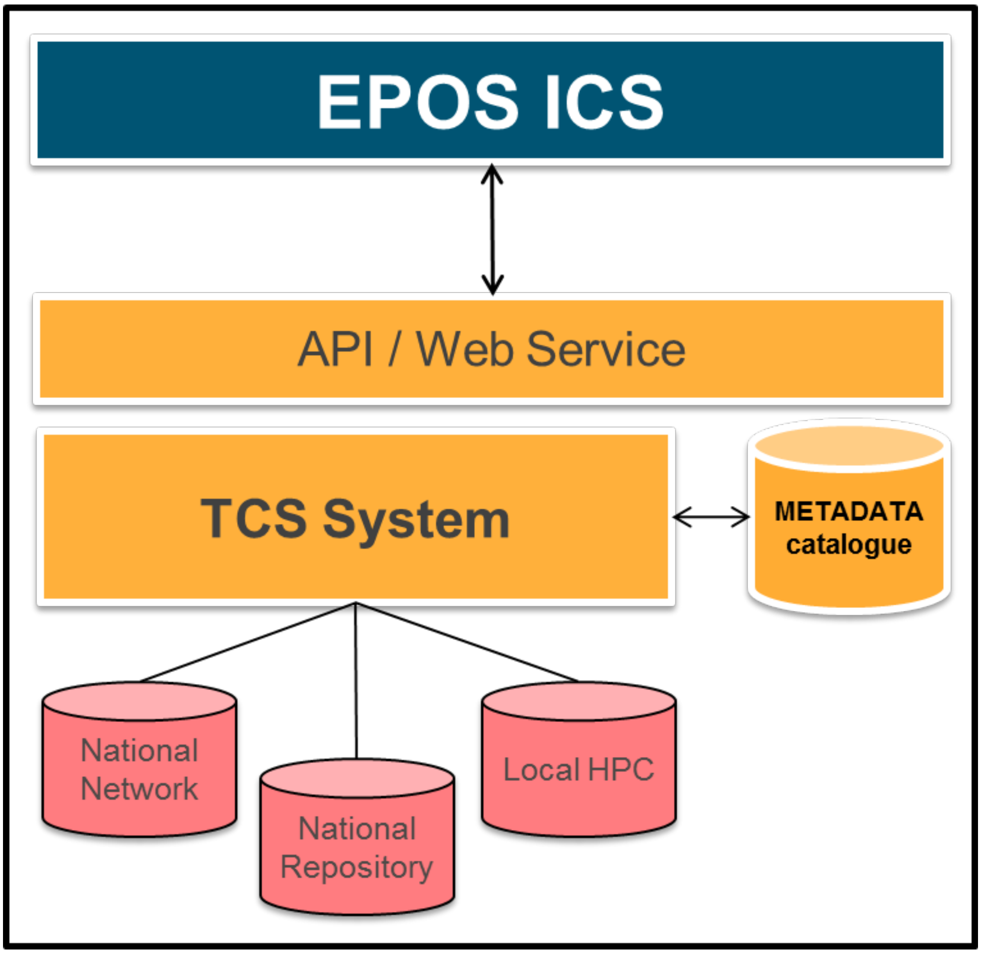

Figure 3: generic TCS architecture. Independently of the community technical choices to federate and make the data accessible, all TCSs should have web services / APIs which enable ICS to access the data.

Some have already done the effort of defining metadata standards and web services to disseminate the data. Others are still in the process of undertaking such work.
As a consequence the scenario is very heterogeneous and includes many services with different levels of "technical" maturity.
Nevertheless, it is still possible to define a very generic architecture which, independent of the TCS level of maturity, can be applicable to any TCS.

The architecture is very "general purpose", simple, and presents the following elements:

a) *National Layer Research Infrastructures*: these can vary in different countries within a specific scientific community. They represent the existing DDSS.
b) *TCS system*: this represents the e-Infrastructure for a specific scientific community. It may include the software used to federate National RIs, or the software to present results on the web (web portal).
c) *Metadata Catalogue*: this is usually a database where each data object (e.g. file, in case of non-streamed data) is referenced and described by the metadata. It can be used to drive the Web Services

d) *Web Services / API*: this is the entry point where users can access the data. EPOS ICS has no particular recommendation with respect to the software used to build the TCS system. What matters is the way ICS and other stakeholders can access data. API based Web Services ensure data and metadata are accessible and discoverable by humans and machines (e.g. ICS system)

With respect to the metadata and related web services, there are two possible strategies to follow in order to achieve the ICS –TCS interoperability:

1. *Metadata dump:* the metadata from TCS is fully copied to the ICS metadata catalogue. It guarantees that the metadata is fully managed by the ICS, and it lowers the burden of TCS in providing a highly efficient and robust system providing access to the metadata. However, it requires periodic (e.g. daily) polling/copying procedures and synchronization mechanisms must be put in place to guarantee consistency.
2. *Metadata Runtime Access:* The access to metadata is done at runtime by querying web services with the defined APIs. The APIs specification must be stored in the ICS Metadata catalogue (to enable ICS to access the system in an autonomic way). It avoids the error-prone procedure of dumping the metadata. However it requires that TCS build very reliable and robust systems, able to manage a high number of concurrent queries as generated by users of the ICS.

In either case the conversion to CERIF needs to be done. Since this could take some computing resources (1) is likely to be more efficient unless the data structures (not the data themselves) of a particular TCS are very fast changing.
It is however likely that we need to use both strategies described above due to the high degree of variability in the TCS DDSS.

ICS-TCS communication

The main requirements of each TCS were collected during EPOS-PP. There will be an additional requirements elicitation process during the initial phase of the EPOS-IP project in order to ensure that the requirements from the TCS communities are properly taken into account. In EPOS-IP interactions between ICS and TCS will be handled through cyclic teleconferences (every 3-6 months) with each TCS. In such meetings the current status of the ICS and its implementation will be summarized and feedback from the TCS community will be provided. EPOS meetings and assemblies will be also exploited to discuss ICT issues, plan the work and improve synergies.

Expected work-cycle

Communication and collaboration between the ICS and TCS teams will be key to building a common understanding of the work we are required to undertake. To initiate this process the ICS development team will collect requirements and use cases from TCSs in order to start to outline and develop the system for the integration layer.

Based on the information collected during the requirements elicitation process, the ICS development team will prepare a suggestion for the harmonisation process between the TCSs. Both the requirements elicitation process as well as the harmonisation of the requirements between the various TCSs, will take place in an iterative manner allowing both ICS developing team and TCS communities to contribute both in the preparation phase but also later in the implementation and validation phase of this information.

Timeline

The detailed timeline for the work-cycle described above will be developed and presented during the kick-off meeting and will be included in the next version (Level-3) of the ICS-TCS Integration Guidelines document.

ICS-TCS common work

The joint work is carried out in the "interoperability layer", a software/technical layer to enable communication among ICS and TCS. Once the TCS-ICS communication is established, data, metadata and services coming from the various communities must be managed.
This is accomplished in the ICS Central Hub (ICS-C) by means of a metadata catalogue, namely the facility which, together with system manager software, manages and orchestrates all resources required to satisfy a user request. By using metadata, the ICS-C can discover data requested by a user, gain access to them, send them to a processing facility (or move the computation to the data), and perform other complex tasks. The catalogue contains: (i) technical specifications to enable autonomic ICS access to TCS discovery and access services, (ii) metadata description of the digital object (DO) with direct link to DO, (iii) information about users, resources, software, and services other than data services (e.g. rock mechanics, geochemical analysis, visualization, processing).
In the roadmap to enable data integration and ICS-TCS communication, three steps are required (figure 4):

(i) Metadata definition,
(ii) web services/APIs definition,
(iii) match and map with ICS

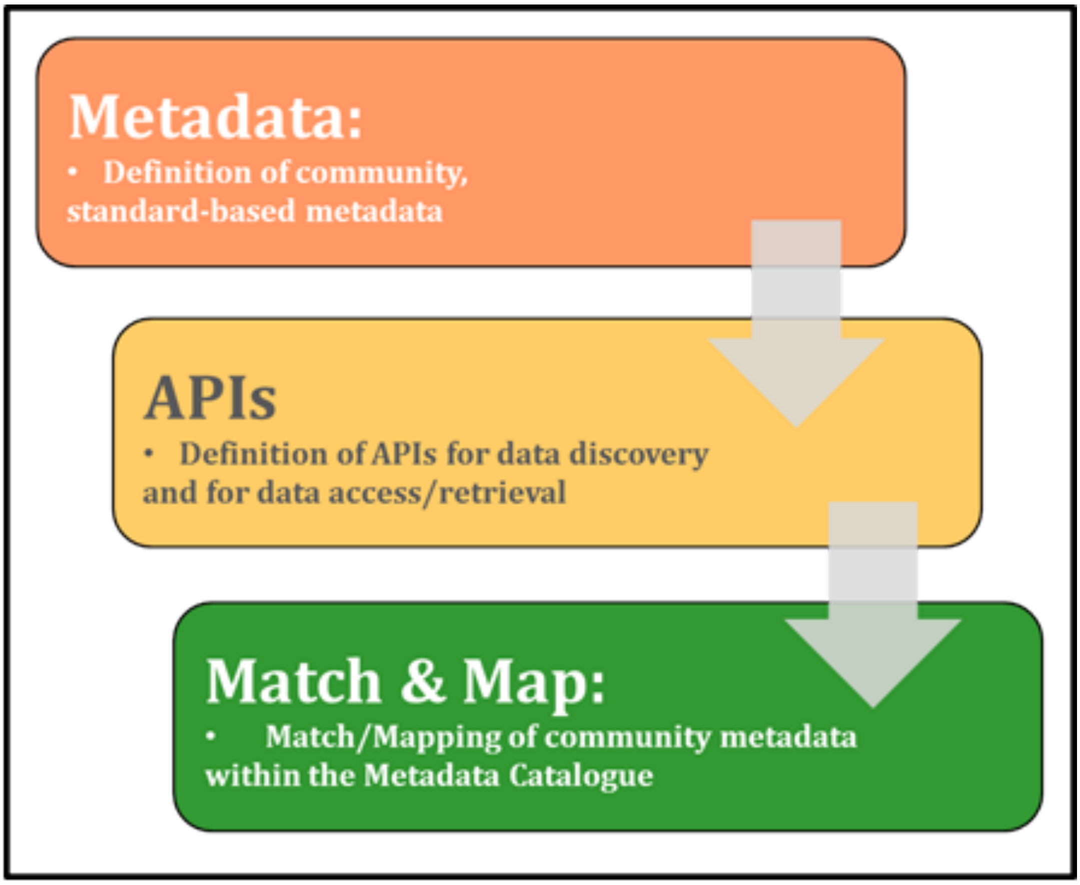

Figure 4: The three steps required in TCS-ICS common work and interaction through the interoperability layer.

Their description follows:

STEP 1: Metadata
As anticipated in the recommendations, it includes the definition and usage of metadata describing the community data (e.g. seismic waveforms, GPS time-series, geological maps), software (e.g. an analysis or visualisation application), services (e.g. use of specialist equipment), resources (e.g. computers, instrumentation, detectors) and users (with their roles, responsibilities and authorities used for AAI).

STEP 2: Web services for access
The heterogeneity of the EPOS community implies that TCS have different maturity levels. Some communities may only give access to metadata attached to raw Digital Objects (e.g. ftp repository containing files with metadata in the header). However the best option is to provide a software layer which implements services for (a) "Data discovery"; (b) "Data Access and Retrieval" (e.g. APIs, RESTful); (c) "data analysis/visualisation/modelling/mining", possibly based on existing standards.

a. *Data Discovery services*
   are used to discover, through metadata, the data of interest in remote repositories. EPOS ICS is indifferent with respect to

topology (single-sited, distributed or federated repository) and access points (single or multiple) as long as clear definitions about how to access to the services are provided.

b. *Data Access and Retrieval services*
   are used to access a digital object (i.e. a file, a dataset, a document) and use it. The term access used can be thought also as a download process. These services may include also IAAA .
   For both services the use of already existing, international, community accepted standards is highly recommended. OGC-services standards, INSPIRE based standards, Dublin Core and other international standards can be relatively easily integrated because a number of tools to manage data and metadata already exist. Reusing existing software tools will save resources for software maintenance in the long term on both TCS and ICS side.
c. *Data analysis/ visualisation/ modelling/ mining services*
   are used to extract meaning from the data using the software services provided. EPOS ICS will discover appropriate software services for the kinds of data requested by the user and (semi-automatically) compose the software stack to process the dataset(s) as required.

STEP 3: Metadata match & map
For each metadata standard from a single TCS a metadata Matching/Mapping procedure is required. This is also called the metadata assisted canonical brokering.
The purpose is matching the metadata elements of each dataset (but also for software services, users and resources as computers, detectors) at TCS level and metadata entities in the ICS metadata catalogue, and generating mappings between matched elements so that any data instances (values of attributes within a record) under one schema can be converted to instances under the other schema.
Matching involves finding corresponding attributes in the two schemas even if they have different names (especially a problem in multilingual environments) or – in the worst case – even different data types. As an example, in two different datasets both concerning observations on the earth surface, there may be attributes describing geospatial positions having different names (Latitude/Longitude versus Northing/Easting) and different accuracy and precision attributes. In order to properly match metadata elements, a common ICS and TCS technical staff shared work, is required. Such work implies that:

- A subset of the most relevant community metadata elements is selected by TCS technical staff in cooperation with TCS scientists
- Such subset is then mapped into the EPOS catalogue by the ICS technical staff.

NEXT STEPS:
Following issues will be addressed in the next steps:

- Suggested technologies/standards.

- ICS will collect from TCS requirements, use cases, assets (requirements elicitation process).
- Details of the timeline and plans showing how to share the work between ICS and TCS. It will be up to TCSs to determine what they expose to ICS.

6 ICS-D Integration

This short remark is intended to organize and provide mechanisms for interaction between those groups that are either interested in following the developments with regard to the Integrated Core Services Central Hub (ICS-C) and the Integrated Core Services Distributed (ICS-D) or those who have a future interest in hosting the ICS-C or declaring an official in-kind contribution for any of the future ICS-D's, and those who are directly involved in developing ICS-C and ICS-d in the EPOS Implementation Phase (EPOS-IP) (i.e. WP6: ICS-TCS interactions and interoperability and WP7: ICS design and development).
For the sake of simplicity the first group of actors is called "interest groups" and the second group of actors is called "WP6 and WP7 developers" throughout the following text.
The difference between the interest groups and the WP6 and WP7 developers is that the latter are the beneficiaries of the WP6 and WP7 as developers, whereas the first group represents those that may either be involved in other WP's of EPOS-IP as beneficiaries, may be involved in one or several of TCS's without being a beneficiary of EPOS-IP or they may be third-parties completely independent from the EPOS-IP.

In order to provide a forum for interactions between the interest groups and the WP6 and WP7 developers, we propose the following mechanisms:

1. During the EPOS-IP Proposal writing stage (i.e. until the January 14, 2015):

   We have established an open interaction and discussion forum in the EPOS Collaborative area (ICS discussion area)where this was used by both interest groups and the WP6 and WP7 developers. Based on the above, WP6 and WP7 developers wrote up the individual tasks and the descriptions in WP6 and WP7, taking into account the requests and declarations provided by the interest groups.
2. During the EPOS-IP Project execution stage (i.e. 2015-2018) there are mechanisms provided and provisions given in the various tasks of WP6 and WP7 to integrate declarations made by interest groups during the EPOS-IP project duration. This includes:
   a. A formal procedure for selecting the ICS-C hosting country during 2015.
   b. A formal application procedure for integration of the various ICS-D's will be prepared through a template. Their further evaluation and subsequent selection process will be made available through an open, inclusive and transparent approach.

In addition we should mention here the continuous evolving nature of EPOS through time beyond the EPOS-IP and the possibility of adding further relevant and interesting ICS-D's during the EPOS Operational Phase (EPOS- OP) (i.e. 2019-....).

Appendix 1 - Acronyms

AAAI	Authentication, Authorisation, Accounting, Infrastructure (in the context of architecture)
AARC	Authentication and Authorization for Research and Collaboration, https://aarc-project.eu/
APIs	Application Programming Interface
CERIF	Common European Research Infrastructure Format
CES	Computational Earth Science
CLOUD	*Cloud* computing is a model for enabling ubiquitous network access to a shared pool of configurable computing resources. Remote computer servers providing storage and computational services
CKAN	Comprehensive Knowledge Archive Framework
DC	Dublin Core Metadata Standard
DCAT	Data Catalogue Vocabulary
DDSS	data, data products, services and software
DO	Digital object
ECO	Executive and Coordination Office
EduGAIN	International interfederation service interconnecting research and education identity federations
EGI-Engage	Engaging the Research Community towards an Open Science Commons, https://www.egi.eu/about/egi-engage/
EPOS-IP	EPOS Implementation Phase
EPOS- OP	EPOS Operational Phase
ERIC	European Research Infrastructure Consortium
GRID	A network of interconnected large scale high performance computational facilities
HPC	High Performance Computing
HTC	High Throughput Computing
IAAA	Identification, Authentication, Authorisation, Accounting (in the context of services)
ICS	Integrated Core Services
ICS-D	Distributed Integrated Core Services
INSPIRE	INSPIRE is a EU Directive in (May 2007), establishing an infrastructure for spatial information in Europe to support Community environmental policies, and policies or activities which may have an impact on the environment
ISO-19115	International standard for geospatial data

Kerberos	A computer network authentication protocol
LDAP	Lightweight Directory Access Protocol
NRIs	National Research Infrastructures
OAuth	An open protocol to allow secure authorization in a simple and standard method from web, mobile and desktop applications
OGC	Open Geospatial Consortium
OpenID	*OpenID* is an open standard that allows users to authenticate to websites without having to create a new password. AAAI standard
OSI	Approved API or Web service standard
OWL	Web Ontology Language
SAML	Security Assertion Markup Language. AAAI standard
Shibboleth	A computer network authentication protocol
SKOS	Simple Knowledge Organization System
RESTful web service	Web Services based on the REST architecture (web)
TCS	Thematic Core services
VERCE	Virtual Earthquake and seismology Research Community in Europe e-science environment http://www.verce.eu/
W3C-PROV	World Wide Web Consortium standards on provenance information
X.509	The *X.509* specification defines a standard for managing public keys through a Public Key Infrastructure (PKI). AAAI standard

Appendix 2 - references

Metadata

Bailo, D., & Jeffery, K. G. (2014). EPOS: A novel use of CERIF for data-intensive science. In *Procedia Computer Science* (Vol. 33, pp. 3–10). doi:10.1016/j.procs.2014.06.002

Jeffery, K., & Bailo, D. (2014). EPOS: Using Metadata in Geoscience. *Metadata and Semantics Research*, 170–184. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-13674-5_17

Jeffery, K., Houssos, N., Jörg, B., & Asserson, A. (2014). Research information management: the CERIF approach. *International Journal of Metadata, Semantics and Ontologies*, *9*(1). Retrieved from http://inderscience.metapress.com/index/VL5422N2U7112669.pdf

Jeffery, K., Asserson, A., Houssos, N., & Jörg, B. (2013). A 3-Layer Model for Metadata. *Proc. Int'l Conf. on Dublin Core and Metadata Applications 2013*, 3–5. Retrieved from http://wiki.dublincore.org/images/d/da/CAMPpapers.pdf

e-Infrastructures

Bailo,D., Jeffery, K. G, Spinuso, A., Fiameni G. Interoperability Oriented Architecture: The Approach of EPOS for Solid Earth e-Infrastructures, *in 2015 IEEE 11th International Conference on eScience eScience 2015 (in press)*

Atkinson, Malcolm, et al. (2015), VERCE delivers a productive e-Science environment for seismology research, *in 2015 IEEE 11th International Conference on eScience eScience 2015 (in press).*