3.1 SHARING STATISTICAL PRODUCTION AND DISSEMINATION SERVICES AND SOLUTIONS IN THE EUROPEAN STATISTICAL SYSTEM (2016.06)

Type of Activity	Common services
Service in charge	ESTAT.B3
Associated Services	ESTAT.B1 ESTAT.B5

3.1.1 EXECUTIVE SUMMARY

European statistics are produced by the European Statistical System (ESS) which is a partnership between the Commission (Eurostat), and the public administration responsible for production and dissemination of official statistics in each Member State[24] (mainly national statistical institutes (NSIs) as well as other national authorities. Member States collect data and compile statistics for national and EU purposes. The ESS functions as a network in which Eurostat's role is to lead the way in the harmonization of statistics in close cooperation with the national statistical authorities. EU statistics focus mainly on EU policy areas - but, with the extension of EU policies, harmonization is growing and concerns nearly all statistical fields. Other statistics on more specific domains may be also produced by Commission DGs.

The ESS members broadly shared the same challenges and drivers: they have to embrace the digital transformation and to produce new innovative and high quality standard statistical products based on the new available data sources under the pressure of limited and decreasing resources.

Based on a long tradition of sharing information, standard and tools, Eurostat and the other members of the ESS jointly developed a common vision: the "ESS Vision 2020".
The ESS Vision 2020[25]

- aims at further developing the cooperation between ESS members;

- strives for setting up more collaborative production processes based on shared architecture in ESS, which should enable addressing new business requirements with a reduced time to market and strengthen quality while increasing efficiency at the same time;

- aims in particular at a future-proof dissemination and communication strategy that satisfies divergent and ever-changing user needs at both national and European level, is flexible enough to adapt to emerging technologies and offers a variety of output channels and services.

Eurostat and Member States will work together on the implementation of the ESS Vision 2020: a programme is defined consisting of concrete projects which all together should help realising the objectives of the ESS Vision 2020.

---

[24] and EEA and EFTA countries

[25] The statistical dissemination has received as well special attention at EC level as external communication is listed as a key domain for cross sector and cross policy IT rationalisation in "Communication from VP Šefčovič to the Commission: Follow up to the Communication "Getting the best from IT in the Commission" of 7 October 2010 - First decisions in the IT rationalisation process".

The objective of the action is to make the leap of sharing statistical services and solution in operational mode in the European Statistical System with in particular a focus on statistical dissemination.

It builds on three VISION 2020 implementation initiatives which are closely related:

1. ESS Enterprise Architecture (ESS EA) – a joint effort to create a comprehensive target state ESS architecture considering both Member States and Eurostat. This allows harmonisation of business processes and bridging the gap between business and IT;

2. "Shared SERVices" (SERV) – a project to create the conditions for sharing technical statistical services (including dissemination) and supporting their integration in the statistical production processes at national, ESS and Commission level;

3. Digital Communication (DIGICOM) – a programme to further develop and modernise key functions such as user analytics, communication, dissemination, data visualisation, mobile solutions etc.

Part of those initiatives are already funded by Eurostat and some other parts require additional funding – potentially from ISA^2. ISA^2 funding is necessary to:

1) Finalise and extend the current ESS statistical production reference architecture to get it closer to implementation integrating the information sharing and the interoperability aspects;

2) Develop a sustainable release of the common infrastructure elements such the ESS catalogue of shared services;

3) Perform a thorough benchmark of as is architecture in MS to identify components which can be readily transformed into shared services as well as mapping the needs and gaps and tentatively define roadmaps for benefiting from shared development;

4) Provide new reusable services and solutions based on existing components or certified open source statistical library/components and to allow statistical producer to upgrade their architecture. For dissemination, a reusable solutions will be derived from the "renovated Eurostat dissemination chain for statistical dissemination";

5) Set up reference implementations of processes using shared services suitable to various environments and to propose technical architecture patterns and open source environments suitable for integration of service in statistical production.

Beyond the Eurostat and the ESS dimensions, the project can serve *European Commission services that produce other statistics*;

- The statistical services for the production and dissemination of statistics delivered by the project will be publicly available and could thereby also be reused by any Commission DG and Agencies of the European Union or by any public administration, should they need so.

- Cost of integration should be relatively low as the reusable components should operate as well on the generic EC infrastructure.

- Packaged solutions will be produced for EC context allowing the plug in of dissemination solution in several EC Services.

3.1.2 OBJECTIVES

The overall objective of the proposed project is to realise the conditions and implement the sharing of statistical services among organisations contributing to the production and dissemination of European statistics building on the early developments initiated by two ESS Vision implementation initiatives, namely the ESS Enterprise Architecture initiative and the Shared Service project.

The objectives of the proposed project are the following:

1. Provide a detailed reference architecture for statistical production processes and information systems allowing the efficient integration of shared statistical services based on work conducted already;

2. Develop the common infrastructure necessary to support sharing of components such as the multi-tenant version of the ESS Service Catalogue;

3. Improved alignment to ESS Reference Architecture and adoption of Shared Services standards for at least 5 ESS members;

4. Support and guide statistical organisations to upgrade their architecture to align better to the target state architecture and to benefit from services sharing;

5. Identify and build shared services based on existing components in use or certified open source statistical libraries;

6. Implement new statistical production processes using shared services;

7. Deliver a reusable packaged solution for statistical dissemination.

3.1.3 SCOPE

The project will deliver:

- Extended and consolidated version of the Statistical Production Reference Architecture;

- Multi-tenant version of the ESS Service Catalogue and related common infrastructure;

- Support to statistical production organisation to upgrade and align their infrastructure to benefit from shared statistical services;

- A list of certified and existing components or libraries suitable for the compilation of shared services;

- A whitelist of architecture patterns and open source components for realising the integration and usage of shared services;

- Implementation of new statistical production processes using shared services providing reference implementations adapted to different contexts;

- A reusable packaged solution for statistical data dissemination.

Out of scope for the project *(what the project will not deliver):*

- Development of statistical methodologies, libraries and code (that is in the scope of separate business projects such as the "renovated Eurostat dissemination chain");

- Production of statistical data;

- Standardisation of metadata repositories used by EC statistical data producers.

3.1.4 PROBLEM STATEMENT

Historically, statistical organizations have developed their own business processes and IT-systems for producing statistical products. Therefore, although the products and the processes conceptually are very similar, the individual solutions are not (as represented by the different shapes in Figure 1). Every technical solution was built for a very specific purpose with little regard for ability to share information with other adjacent applications in the statistical cycle and with limited ability to handle similar but slightly different processes and tasks. This can be referred to as 'accidental architecture' as the process and solutions were not designed from a holistic view.
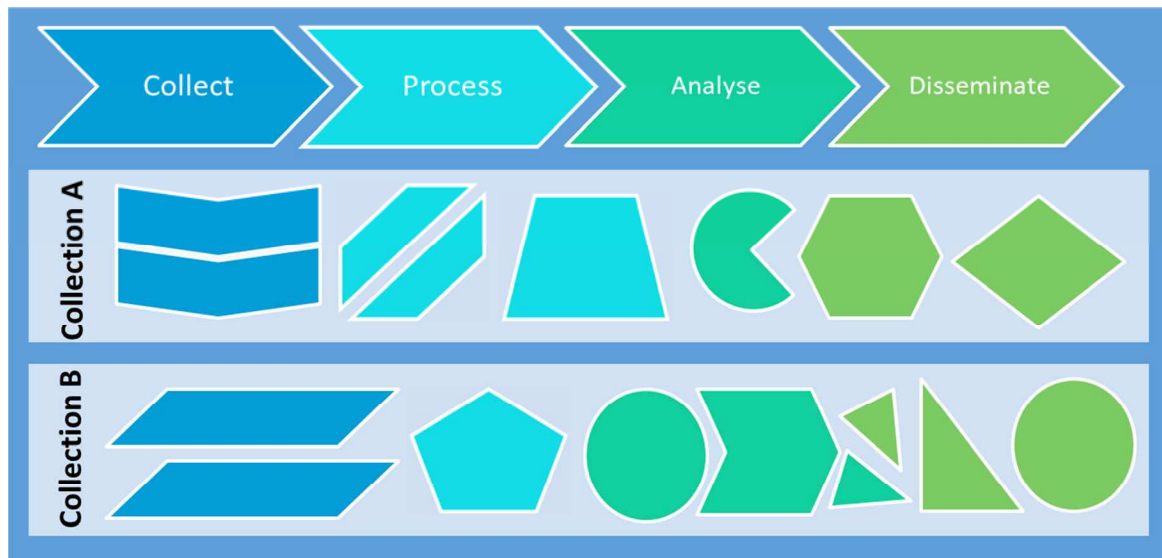

Figure 1: Accidental Architectures

Often it is difficult to replace even one of the components supporting statistical production. Use of these processes, methods and an inflexible and aging technology environment mean that statistical organizations find it difficult to produce and share data and information aligned to modern standards (for example, Data Documentation Initiative (DDI) and Statistical Data and Metadata eXchange (SDMX)). Process and methodology changes are time consuming and expensive resulting in an inflexible, unresponsive statistical organization.

Enterprise architecture is more and more used by statistical organisation to underpin their vision and change strategy. An enterprise architecture aims to create an environment which can change and support business goals. It shows what the business needs are, where the organization wants to be, and ensures that the IT strategy aligns with this. Enterprise architecture helps to remove silos, improves collaboration across an organization and ensures that the technology is aligned to the business needs. Enterprise architecture work enables to standardize organisation and processes. This is shown in Figure 2 where, as opposed to Figure 1, the countries have standardized their components and interfaces.
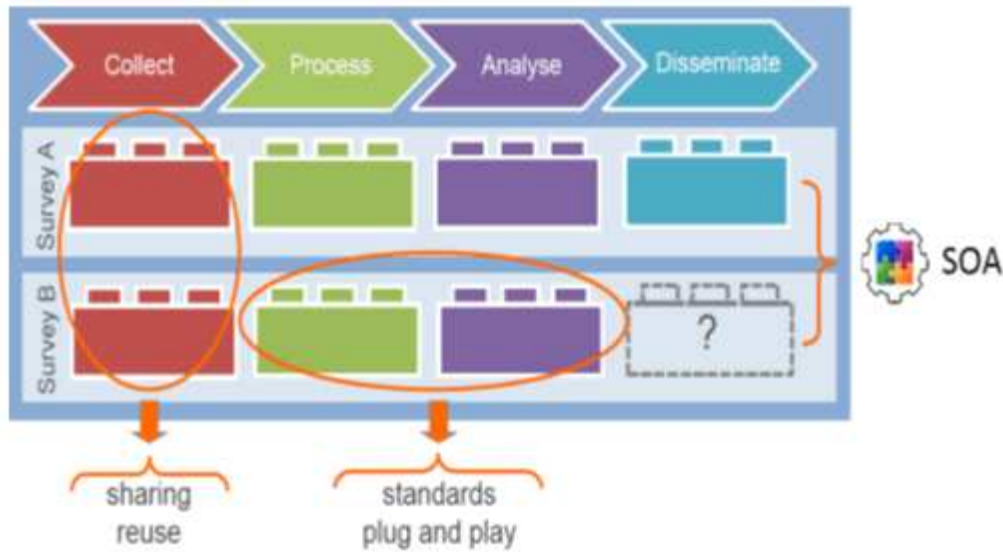
Figure 2: The result of standardization within an organization

The challenges of statistical production have a lot in common in different countries. The developed processes and IT tools do not correspond to the similarities; in most cases countries developed their own specific solutions. This model is highly costly and came under challenges as a result of new business needs and the need for cost savings.

A common reference architecture will allow the statistical organisation in the ESS to share development cost and to provide new statistical products in a cost efficiency manner.

3.1.5 EXPECTED BENEFICIARIES AND ANTICIPATED BENEFITS

Beneficiaries	Anticipated benefits
Eurostat	• Harmonization of processes and methodologies for the production of EU statistics • Capability to promote statistical standardisation and related best practices • Increased quality of statistical data • Technical interoperability, by providing standardized interfaces for statistical dissemination
European Statistical System – National Statistical Institutes in Member States	• Support the cost-efficient modernization of architectures • Improved statistical processes by sharing best practices and services • Delivery of new statistical products based on new data sources • Cost savings via reusable services and solutions (e.g. as the reusable dissemination solution to be made available through Joinup) • Capability to share further statistical services

Other European Commission services	• Reuse statistical services for the production of European-level statistics (e.g. data validation services)
	• Reuse architecture patterns for statistical production based on corporate EC platform
	• Reuse the statistical dissemination solution, achieving
	    ○ Cost-efficiency and economies of scale. To illustrate the scale of an investment on a single DG basis into a dissemination solution, a potential benchmark could be a recent statistical dissemination project launched by DG REGIO for the dissemination of cohesion policy data where the TCO for the period 2014 to 2018 was estimated to 252K EUR with the choice of a 3rd party cloud solution vendor. Economies of scale at Commission level can be achieved as soon as 4 DGs or services start to reuse the solution, as the investment into a reusable dissemination solution will then be lower than individual project expenditures on silo implementations
	    ○ IT rationalisation, i.e. systematic reuse of a set of systems for same business processes (statistical data dissemination)
	    ○ Integration with EU Open Data Portal, preparation for open data and semantic web
	    ○ Integration with NE-CMS
	
	Currently, as to the reuse of dissemination tools
	    ○ DG **COMP** is reusing legacy Eurostat dissemination systems for disseminating [State Aid Scoreboard statistical data](#), as part of the project DG COMP should be transitioned to use the future reusable dissemination solution
	    ○ DG **GROW** is about to start reusing legacy Eurostat dissemination systems for the dissemination of postal statistics, as part of the project DG GROW should be transitioned to use the future reusable dissemination solution
	    ○ DG **EMPL**[26] is reusing legacy Eurostat dissemination systems, for the dissemination of [Labour Market Policies](#), as part of the project DG EMPL should be transitioned to use the future reusable dissemination solution
	Then
	    ○ DG **ECFIN** expressed its interest[27] in a future solution for the scopes of the Business and Consumer Surveys, and

	the annual macro-economic database which is publicly available through [Ameco online](). Eurostat built a [prototype of future dissemination solution based on Ameco]()[28]
	o DG **TAXUD** is looking for a solution to disseminate effective tax levels. ESTAT.B3 is analysing dataset samples provided by TAXUD to investigate how these could be disseminated by means of the reusable dissemination tools
	o DG **EAC** would be interested to use the ESTAT provided dissemination tools for making student mobility data which is [ad interim hosted on the Open Data Portal]():
	o Agency **EACEA** would be interested to use the ESTAT provided dissemination tools for the scope of [Eurydice]() indicators
	o DG **COMM** is looking for a long term solution to disseminate survey data for the [Eurobarometer]() (survey data)
	o Agency **FRA** (and DG **JUST)** could be interested in the solution for the dissemination of [fundamental rights survey data]()

European citizens (data consumers)	• New and higher quality statistical products
	• Increased cost-efficiency of official statistics
	• Easier and smarter access to EU official statistic products through
	   o better interoperability of the multiple statistical data scopes
	   o multi-dimensional statistical datasets access by means of standardized interfaces, supporting cross-sector & cross-policy data reusability, consequently openness

The following measurements will be used to evaluate the success of the project:

- The Statistical Production Reference Architecture and the related architecture artefact are used for the development and governance of Shared Service in the ESS;

- A majority of ESS members have understood the ESS reference architecture and have benchmarked their local architecture against it identifying potential for reuse and needs and a possible transition roadmap for benefiting from shared services;

- A series of common resources are provided and managed for the Official statistical community in order to realise benefits from shared services:

  o The ESS Service Catalogue will contain conceptual description of a significant number of potential statistical services and resources candidate for sharing;

  o The catalogue will link to at least 5 new reusable services including dissemination services designed following guidelines based on existing component and library to be

hosted on a service oriented IS architecture in a Member State (e.g. by replication) or centrally on an ESS/Eurostat SOA infrastructure;

- o A whitelist providing reference and description of certified open source statistical library and routine to perform key ESS business functions;

- o A whitelist providing reference and guidance for the use of open source components for realising the SPRA in particular service exposure and integration.

- At least 5 ESS members have upgraded where necessary their architecture and re-use services of the service catalogue. This will result in a significant cost-saving for those ESS members – instead of developing a certain functionality they can re-use the functionality developed by another ESS member. A few standard implementations are referenced as good practices for other NSIs;

- A reusable dissemination solution is made available through Joinup and will be reused by at least 4 European Commission services. It allows to

- o Expose interfaces based on established standards such as the SDMX information model, and paves the way for open data / semantic web (statistical content made available in RDF);

- o Provide consistent interfaces for data consumers irrespective of the data publisher, the statistical domains and across sectors and policies (by means of the same standardized statistical dissemination (API) interface).

3.1.6 RELATED EU ACTIONS / POLICIES

Action / Policy	Description of relation, inputs / outputs
Commission decision on Eurostat	Commission Decision of 17 September 2012 on Eurostat (2012/504/UE), which appoints Eurostat as the leader for European Statistics within the Commission *Relation: policy, responsibilities* *Inputs: Article 1 subject matter, Article 2 definitions, Article 6 tasks of Eurostat* *Outputs: -*
Production of European Statistics	Regulation (EC) n° 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics and repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities, Council Regulation (EC) No 322/97 on Community Statistics, and Council Decision 89/382/EEC, Euratom establishing a Committee on the Statistical Programmes of the European Communities *Relation: policy, responsibilities, governance* *Inputs: Article 3 definitions, Article 4 the European Statistical System, Chapter 3 Production of European Statistics* *Outputs: -*
IT rationalisation	*Communication from VP Šefčovič to the Commission of 01/08/2012 : "Delivering user-centric digital services"*

	Relation: strategy, IT Vision, Enterprise Architecture *Inputs: Centre of Excellence on Enterprise Architecture* *Outputs: -*
IT rationalisation	Communication from VP Šefčovič to the Commission: Follow up to the Communication "Getting the best from IT in the Commission" of 7 October 2010 - First decisions in the IT rationalisation process *Relation: governance* *Inputs: 3.2 Domain-specific conclusions, "Rationalisation of external communication"* *Outputs: -*
Production of European Statistics	European Statistical System Vision 2020 *Relation:* *Inputs: 3.4 Efficient and robust statistical processes & 3.5 Dissemination and communication on European statistics* *Outputs: -*

3.1.7 REUSE OF SOLUTIONS DEVELOPED BY ISA, ISA[2] OR OTHER EU / NATIONAL INITIATIVES:

3.1.8 The project will be based on a significant number of solutions available

Solution	Description	Annotation
SDMX – Statistical data and Metadata eXchange	An initiative to foster standards for the exchange of statistical information, and is sponsored by BIS - ECB - EUROSTAT - IMF - OECD - UN - World Bank. SDMX is approved by ISO as an International Standard (ISO 17369:2013); SDMX information model can be used to standardize information exchange between standards	
Joinup	The reusable statistical services delivered by the project will be made available to anyone by means of Joinup. Joinup will serve as a repository for the ESS shared services catalogue where interfaces and service description will be make available in a standard way.	
EIRA - European Interoperability Reference Architecture	ISA programme of the European Commission for classifying and organising building blocks relevant to interoperability, which are used in the delivery of digital public services. The goal is to facilitate interoperability and reuse when developing public services. EIRA will be used to upgrade the ESS Statistical Production Reference Architecture.	

DCAT Application Profile for data portals in Europe	The DCAT Application Profile for data portals (DCAT-AP) provides a common specification for describing public sector datasets in Europe to enable the exchange of descriptions of datasets among data portals.	
EUPL - European Union Public Licence	The EUPL is the first European Free/Open Source Software (F/OSS) licence The expected applicable license scheme used for the delivery of this project is likely to EUPL (solution derived from existing open source software component, licensed under the EUPL)	
EIC - European Interoperability Cartography	The EIC is an instrument to map and analyse the interoperability landscape in Europe and to identify solutions that are available	
Open data Support	Open Data Support is a 36 month project of DG CONNECT of the European Commission to improve the visibility and facilitate the access to datasets published on local and national open data portals in order to increase their re-use within and across borders.	
Catalogue of Services - Service attributes	A2A Catalogue of Services- list of common service attributes or service descriptors to identify web services across different MS	
IPSIS	EC interoperability platform consisting of core tools and technologies, reference solutions and a competency centre for use by information system builder's to integrate IT services.	
CSPA	The Common Statistical Production Architecture developed under the coordination of the High Level Group on Modernisation of Official Statistics (HLG-MOS). It is an industry architecture for the official statistics industry i.e. set of agreed common principles and standards designed to promote greater interoperability within and between the Official Statistics producers	
ESS EA RF	The ESS Enterprise Architecture Reference Framework developed for the implementation of the ESS Vision 2020. It serves as a reference for describing and articulate project outcomes and contributions to ESS "to be state". The ESS EA RF will be publicly released by end 2015.	
GSBPM and GSIM	The Generic Statistical Business Process Model and the Generic Statistical Information model developed	

and maintained by UNECE. They provide reference frameworks (definitions, relations, attributes) for describing statistical production process and the piece of information used in these processes. It strive for greater interoperability and cooperation among statistical organisation	

3.1.9 EXPECTED RE-USABLE OUTPUTS (solutions and instruments)

Output name	Statistical Production Reference Architecture V1.0 and subsequent
Description	Based on the ESS EA RF developed by the ESS, this (ISA^2) project will provide a fully fledge and ready for implementation set of artefacts to standardise EU statistic production processes including information and interoperability aspects. Subsequent releases will be enriched by reference to standard solution and reference implementations. Organisation should use it to benchmark their production architecture, develop services to be shared and integrate shared service in their production of statistics. This reference architecture can be re-used by the ESS Members and also by other Commission DG's dealing with official statistics.
Reference	European Interoperability Reference Architecture (EIRA)
Target release date / Status	First release 31/12/2016

Output name	Multi-tenant version of the ESS Service Catalogue
Description	Multi-tenancy is an architecture in which a single instance of a software application serves multiple customers. A central ESS Service Catalogue is required to publish the statistical services that are available for re-use in the European Statistical System. This service catalogue should use the same solution as the global (UN sponsored) service catalogue of statistical services (CSPA Service Catalogue). These catalogues shall be based on the same system, but shall clearly indicate the level of availability of the offered statistical services (e.g. ESS level or global level). It shall be also analysed how the ESS Service Catalogue – that contains statistical services according to international standards – can be integrated with the Service Catalogue of the Joinup

	platform.
Reference	
Target release date / Status	31/12/2017

Output name	Reference implementations of statistical and dissemination processes using shared services adapted to different contexts
Description	The project will identify and prioritise development of shared services within ESS. 5 new shared services based on existing components or statistical libraries will be implemented during the project and reference point in the ESS catalogue. The shared service will be implemented in the statistical processes of multiple ESS members and bring them process improvements. Furthermore the implementation and integration of those statistical services in several ESS members will lead to improvement of those services, which will allow easier adoption by further organisations.
Reference	TF Shared Services Mandate
Target release date / Status	31/12/2017

Output name	White list of open source packages for statistical production business functions and for integration and orchestration of statistical productions
Description	Re-using services can be based on services developed by other statistical organizations and also on open source. This work package will leverage the open source solutions for statistical production and for process orchestration.

Commercial statistical production systems (e.g. SAS or ORACLE) have sophisticated functionality; however they require high license fees. On the other hand open source packages (e.g. packages in language R) are offering similar functionality. The price of open source is much lower, however it is not clear if the applied algorithms can be trusted at the same level and provide the same level of integration and interoperability of solutions as the ones in the commercial packages. Some NSI's are using the open source statistical packages, however some others |

	refrain using them with the rationale that verifying the open source packages would require at least as much resources as the licence fees of the commercial packages. However, if we consider the ESS as whole, and its economy of scale, the verification of open source software can already pay off. This deliverable will provide a verified set of open source packages that can be safely used for statistical production. Similarly the open source packages for process orchestration can facilitate the implementation of service oriented architectures in statistical organizations.
Reference	
Target release date / Status	31/12/2017

Output name	Technical architecture patterns for realising the ESS EA
Description	The analysis of open source software packages and the benchmark of MS architectures will produce a number of technical architecture patterns for realising the target sate architecture. This will enable MS's to make practical decisions to start implementing the architecture in full scale. These architecture patterns will also be available and beneficial to other producers of statistics such as parts of the Commission.
Reference	
Target release date / Status	31/12/2016

Output name	Fit/Gap Analysis and roadmaps for the transition to a target state architecture
Description	As part of the project, at least 5 ESS members use the benchmarks to evaluate their fit to the defined architecture. The members will perform a fit/gap analysis, define roadmap to target architecture, and implement measures to improve alignment with the target architecture. The roadmaps can be used by other organizations as examples for transitioning to a modernised architecture, which should lower barriers and increase the likelihood of successful realisation of the ESS EA.
Reference	
Target release date / Status	31/12/2016

Output name	Inventory of reusable software components for statistical production
Description	The benchmark of the Member State architectures will identify and qualify a various solutions and services that can be made available to the ESS community and potentially outside the ESS
Reference	
Target release date / Status	31/12/2017

Output name	Reusable solution for statistical dissemination
Description	1. CSV import The tool to load • Data Structure Definition (DSD)[29] defining a data collection • Statistical data that is part of a data collection to be disseminated 2. Data store 3. API i.e. web services for machine-to-machine automated data consumption in multiple formats (SDMX, RDF, JSON, tab separated …). 4. Web based data browser, i.e. interface for the access to the data and metadata, including the • Navigation in data collections, • Consultation of data and metadata, • Filtering on data dimensions, the visualisation of data (charts, maps), • Capacity to compare indicators and geo areas, • Capacity to download data in full or to query the API for filtered dimensions (data query)
Reference	SDMX Information Model Generic Statistical Business Process Model - GSBPM
Target release date / Status	**1st half of 2017** • The reusable dissemination solution should be available for use in production (cfr. Scope description above) by a first batch of Directorates General in the 1st semester of 2017. At the moment of writing this document, most probably this would be a group consisting of

---

[29] The Data Structure Definition (DSD) describes the structure of a particular set of data through a list of descriptor concepts. It defines which concepts are dimensions (identification and description - for example: Frequency, country, variable/topic, time period), and which are attributes (just description / qualification - for example the unit of measure, confidentiality flag or the data status flag). See also SDMX self-learning package No. 1 Student book **Introduction to SDMX**

	○ ECFIN ○ TAXUD ○ One or more of the following :EMPL / GROW / COMP ○ Possibly EAC /EACEA (to be assessed) - Sharing of the solution by means of Joinup, i.e. the solution can be reused by other institutions, public administrations, members of the European Statistical System, etc. **As from mid-2016 & beyond** - Further iterative extensions in terms of reuse at ESS level

3.1.10 ORGANISATIONAL APPROACH

3.1.10.1 Expected stakeholders and their representatives

Stakeholders	Representatives
Eurostat	Eurostat Architecture (extended) Team Eurostat B3, B4, B5
NSI's	Members for the Task Force Shared Service
Other Commission DGs	Other statistics - network of statistical correspondents
Information Resource Managers (IRMs)	• List of Information Resource Managers • Digital Stakeholder Forum - chaired by DIGIT and assures inter-service communication and coordination for all matters relating to IT in the Commission
DIGIT	• DIGIT.B1 (Architecture centre of excellence) • DIGIT Hosting services • DIGIT IPSIS team

3.1.10.2 Communication plan

Being part of the ESS Vision 2020 implementation the project will benefit from a broad communication plan designed and monitored by the ESSC and the VIG (ref). Specific component targeting Commission services producing statistics will be added.

The main list of stakeholders for ESS Vision 2020 and tentative related communication channels are:

European Commission	MyIntraComm

	Ad hoc seminar and workshops
	Digital Stakeholder Forum
	Leaflets
Eurostat staff	Cybernews
	Eurostat-Infos
	Lunchtime presentations
	Ad hoc seminar and workshops
NSI staff	European Statistical Training Program
	Ad hoc seminar and workshops
NSI management	Regular presentation (joint events)
	ESS Website
	Dedicated European Statistical Training Program courses
	Circabc
	Leaflets
General Public	Eurostat website (ESS vision dedicated section)
Official statistics Community	Conferences
	Leaflets

3.1.10.3 Governance approach

The project is

- owned by the Deputy Director General of Eurostat, Ms. M. Kotzeva;

- compliant with PM2 project management methodology.

The project will report about the generic tools for the reusable dissemination solution for internal Eurostat governance to the

- Eurostat Dissemination Chain Steering Group;

- Eurostat IT Advisory Committee, and finally to

- Eurostat Director's Meeting

The project will report to the related ESS governance bodies:

- The ESS Task Force Shared Services that will also act as a Steering Group for the project;

- ESS IT Director's Group (ITDG) will review project progress and its main deliverables;

- The Vision Implementation Group established by delegation of ESSC (The European Statistical System Committee) will provide strategic guidance for the project.

At EC level, involved governance bodies are the following

- IT Governance: the IT Board

3.1.11   TECHNICAL APPROACH

The project builds on a service-oriented paradigm to establish the cooperative architecture for ESS and for sharing statistical services among organisation. This approach leans on the SOA strategy of the Commission and on the Common Statistical Production architecture developed at industry level

facilitated by UNECE. The reusable solution for dissemination leans on the future renovated Eurostat dissemination chain (see task 8 description)

Task 1: Develop detailed ESS EA.
Activities: Detail the existing ESS EA to provide an operationalization of the sharing of services, the orchestration of these services in production processes and the management of data and metadata in the process. Communicate the ESS EA and facilitate the discussion and agreement in the ESS community on the ESS EA.
Deliverables: The ESS EA incorporating EIRA and operationalizing the sharing and orchestration of services and the management of metadata.

Task 2: Benchmark ESS architectures.
Activities: Analyse the ESS member architectures (as-is and target architectures) and benchmark the implementations and usage of technology to support the ESS EA. Identify potential components and services for sharing in the community.
Deliverables: Best practice architecture patterns, list of sharable services/solutions, sample roadmaps for realising the ESS EA based on specific as-is architectures and business requirements. List of possible candidates for shared services.

Task 3: Develop multi-tenant version of the ESS Service catalogue.
Activities: Develop an ESS Service catalogue for federated use in the ESS with requirements that support a flexible adoption and supporting easy discovery, test, and implementation of usage of a shared service which is preferably built on existing software.
Deliverables: An ESS service catalogue, which can be deployed in a federated manner including both service shared in the ESS as well as MS specific services.

Task 4: Develop statistical and dissemination services for sharing.
Activities: To select and develop existing functionality into shared services that can be used by the community.
Deliverables: Three services developed and made available to the community either at ESTAT or an NSI.

Task 5: Produce white-list of open source software.
Activities: Analyse existing open source software packages and produce a white-list of components to be used in the technical architectures
Deliverables: The white-list of open source packages and guidance on its usage in the ESS EA.

Task 6: Support architecture alignment.
Activities: Support the architecture alignment in ESS member to be carried out by a group of experts from ESS members e.g.: support for implementing an open source software package, detailed guidelines for exposing a shared service to the community.
Deliverables: Active support function to architecture alignment with established KPI (e.g. number of early adopters of the target architecture in the ESS) .

Task 7: Implement shared services in production processes.
Activities: Support the implementation of the shared statistical services within ESS members.

Deliverables: The integration of shared statistical services in 7 ESS members production processes.

Task 8: deliver a reusable statistical dissemination solution.

Activities: derive a simplified version from the future renovated Eurostat dissemination chain.

Deliverables: generic tools for reuse, i.e.

a) Software layers

This is derived from the architecture retained for the renovation of Eurostat's own dissemination chain (which is more sophisticated and not detailed in full here[30]).
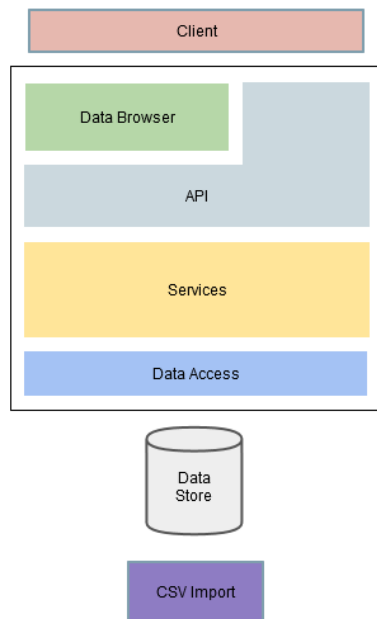

Figure 3: software layers of the reusable statistical dissemination solution

The client has access to both the data browser and the API. Note that the data browser is in itself a client of the API. The API delegates the work to the services which in turn use the data access layer to communicate with the data store. A separate (standalone) application provides the ability to import CSV data into the data store.

b) Data access

The default data store shall be a mix between flat files and a MySQL database. The default data access layer shall reflect this. However, if the data provider has the need for a more responsive and robust solution, an Oracle XML-based data store may be used and in this case a compatible data access layer shall be provided.

c) Packaging

The reusable solution for statistical dissemination will consist of a simple war file, containing the following:

- A web archive (war) containing the data browser and associated resources
- The individual services packaged as ejb files
- Any libraries which are not provided by the container
- A simple jar file which contains the data access layer.

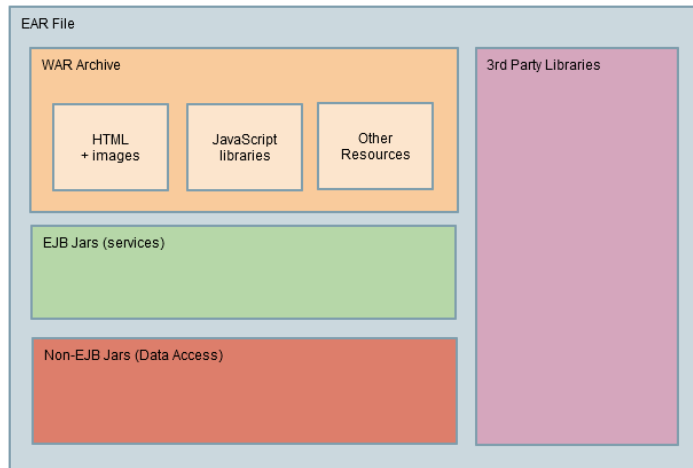[30] See the architecture documentation available in CircaBC

Figure 4: packaging of the reusable statistical dissemination solution

The CSV import module shall be packaged separately as an executable jar file.

d) <u>Infrastructure requirements and system configuration for the reusable statistical dissemination solution</u>

The delivered package is deployable on a LAMT (Linux, Apache, MySql, **Tomcat**) environment.
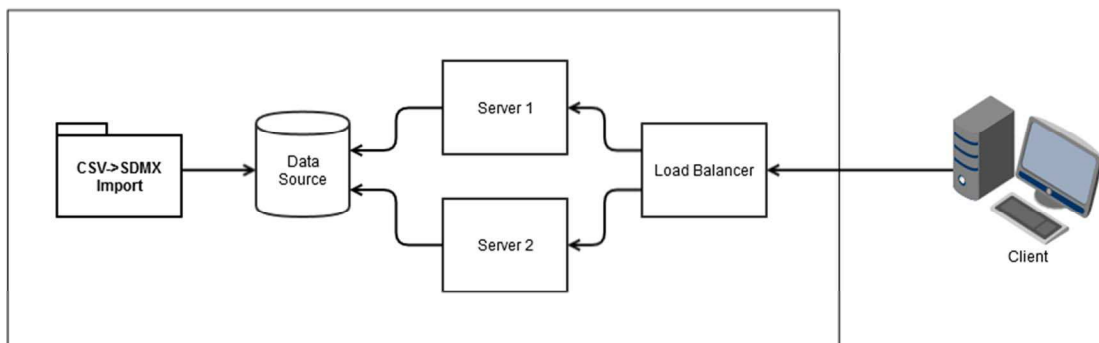

Figure 5: infrastructure requirements for the reusable statistical dissemination solution

The hosting requirements are thus:

- Two Tomcat servers (for failover)
- Data source, when necessary this is a XML enabled database
- A load balancer

Task 9: Upgrade of dissemination solution and services for the ESS in synergy with DIGICOM

Activities: extensions to the reusable statistical dissemination solution in a ESS context and further development of synergies with the Work Package 3 of DIGICOM[31], in particular

1. For Linked Open data (LOD)
   - Contribution to the DIGICOM WP3 stocktaking exercise on Linked Open Data at ESS level, i.e. what does the solution include in terms of support for RDF – Data Cube Vocabulary data delivery
   - Potential technical evolution(s) of the reusable solution for the dissemination of statistics to contribute to the development and maintenance of ontologies
   - Potential technical evolution(s) of the reusable solution for the dissemination of statistics in order to have seamless integration with the Pan-European Open Data Portal
   - Provision of the semantic format conversion service from SDMX 2.1 to RDF-Data Cube Vocabulary

---

[31] See also http://www.cros-portal.eu/sites/default/files//TFDIGICOM-2015-0910-02.pdf

2. For the facilitation and harmonisation of APIs to European data
   - Contribution(s) to the definition of ESS standard for APIs in the area of dissemination of statistics extending experience with standard API of the reusable dissemination solution

Deliverables: generic tools and services for reuse

3.1.12 COSTS AND MILESTONES

3.1.12.1 Breakdown of anticipated costs and related milestones

Phase: Inception Execution Operational	Description of milestones reached or to be reached	Anticipated Allocations (KEUR)	Budget line ISA^2/ others (specify)	Start date (QX/YYYY)	End date (QX/YYYY)
Inception	Project charter discussed and agreed with stakeholders	0	ESTAT	15/09/2015	31/12/2015
Execution	*Task 1: Develop detailed ESS EA.*	200	ISA^2	01/01/2016	31/12/2016
Execution	*Task 2: Benchmark ESS architectures.*	300	ISA^2	01/07/2016	30/06/2017
Execution	*Task 3: Develop multi-tenant version of the ESS Service catalogue.*	200	ISA^2	01/01/2017	31/12/2017
Execution	*Task 4: Develop statistical services for sharing.*	250 (50 per service)	ISA^2	01/01/2016	31/12/2018
Execution	*Task 5: Produce white-list of open source software.*	200	ISA^2	01/01/2017	31/12/2017
Operational	*Task 6: Support architecture alignment.*	500	ISA^2	01/07/2017	31/12/2018
Operational	*Task 7: Implement shared services in production processes.*	400	ISA^2	01/07/2017	31/12/2018
Execution & Operational	*Task 8: Deliver a reusable statistical dissemination solution.*	920	ISA^2	01/01/2016	30/06/2017
Execution & Operational	*Task 9: ESS extensions to reusable statistical dissemination solution.*	500	ISA^2	01/06/2016	31/12/2018

	Total	3470			

3.1.12.2 Breakdown of ISA^2 funding per budget year

Budget Year	Phase	Anticipated allocations (in KEUR)	Executed budget (in KEUR)
2016	Initiation & Execution	1.050	
2017	Execution & Implementation	1.720	
2018	Implementation & Operation	700	

3.1.13 ANNEX AND REFERENCES

Description	Reference link	Attached document
SERV Business Case	https://circabc.europa.eu/sd/a/0ffc64b1-5d5c-4a61-a030-4acd897779e0/SERV%20Business%20case%20v0.7.pdf	-
SERV TF Mandate	-	Version 1.1
ESS EA RF	-	Version 0.5 (draft)
ESS Vision 2020	http://ec.europa.eu/eurostat/web/ess/about-us/ess-vision-2020	-
DISSCHAIN RENOV Business Case	https://circabc.europa.eu/w/browse/a21ebeea-7491-4806-8306-2ace57894218	