On the design of a reference architecture in Istat

**Keywords:** Enterprise Architecture, international standards, process standardization

# Introduction

In recent years, the National Statistical Institutes (NSIs) in the most advanced countries have been carrying out an in-depth analysis of their cultural, organisational, and technological context. The need of such analysis is connected to the new challenges NSIs are experiencing resulting from the considerable changes of the external (e,g. new statistical needs, new data sources) and internal context (e,g. budget cuts , self-contained organizational structures). To face these challenges NSIsare adopting an holistic approach in managing their business activities, and one of the key elements of the new approach is the Enterprise Architecture.

At international level, Eurostat’s Vision 2020 has defined a roadmap based on the development of an Enterprise Architecture to support the modernisation process and the sharing of information and experiences between different countries.

The following paragraphs provide a high level description of Istat’s Enterprise Architecture, implemented according to the international standards. Further, a use case is proposed to describe the application of EA principles to a strategic project launched in Istat.

# Enterprise Architecture

The main goal of EA is to create the best environment to support business goals, aligning IT strategy with business requirements. EA principles help to adopt a process oriented approach, helping to remove silos and to increase cooperation through an organisation.

## International reference standards

The ESS Enterprise Architecture Reference Framework (ESS EARF [1]) is a framework to support the implementation of the ESS Vision 2020. This framework has been developed by adapting to the statistical context the main concepts and definitions from the Open Group Enterprise Architecture Framework (TOGAF [2]).

ESS EARF also refers to the main standards adopted in official statistics and particularly:

* Generic Statistical Business Process Model (GSBPM [3]), for the classification of statistical processes, in terms of phases and sub-phases.
* Generic Statistical Information Model (GSIM [4]), for the information model that supports the definition of the information objects used in the statistical domain.
* Generic Activity Model for Statistical Organization (GAMSO [5]), for the activities that take place within a typical statistical organization generalizing GSBPM, with an extension in the areas of "Strategy", "Capability" and "Corporate Support".
* Common Statistical Production Architecture (CSPA [6]), as a reference for the design and implementation of statistical services, supporting production processes.

## Enterprise architecture domains

According to TOGAF, EA framework is based on the following domains:

* Business architecture (BA) domain, describing the organizational structures and addressing the questions WHAT and WHO.
* Application architecture (AA) domain, related to the different applications and their interactions, within the business processes, addressing the question HOW.
* Information architecture (IA) domain, describing the structure and interaction of logical and physical data assets and data management resources, addressing the question HOW.
* Technology architecture (TA) domain, referring to the software and hardware needed to implement data, and application services.

# Implementing Enterprise Architecture in istat

This work is mainly focused on the description of IA and AA domains in Istat.

## Information architecture

The implementation of EA principle in Istat has started focusing on the Information Architecture domain, compliant with a data centric approach. After the analysis of the current architecture (AS-IS), the revised framework is modelled in Figure 1. In the proposed IA, the data assets have been classified in three different layers, according to their status in the production chain and GSBPM core phases. The cross-cutting metadata layer drives all the transformations between and within the different data layers.



Figure 1. EA – Information Architecture

In the following IA layers are described:

* RAW DATA: layer that includes all ingested data, collected from different sources (e,g, administrative data, survey data, big data), considered as the initial input for the following transformation processes.
* WORKING DATA: layer that refers to the subset of data assets involved in the data editing and validation processes. More specifically, the application services accessing this layer perform data transformation and allow to track the main data changes.
* VALIDATED DATA: layer containing the final output, both at micro and macro level, of the statistical processes.

## Application architecture

The AA designed in Istat is based on the following main groups of applications, used in the core statistical processes or to support cross-cutting activities:

* UTILITY COMPONENTS: software systems performing cross-cutting functionalities that can be used by several production processes (e,g, data analytics, ETL, reporting).
* ENTERPRISE PLATFORM: software systems built to provide services at a corporate level, for several users, both internal or external (e,g, dissemination platform, data collection generalized software, metadata management corporate system).
* APPLICATION COMPONENTS: statistical services, implemented for specific users and for core business activities.

An overview of AA is displayed in Figure 2.

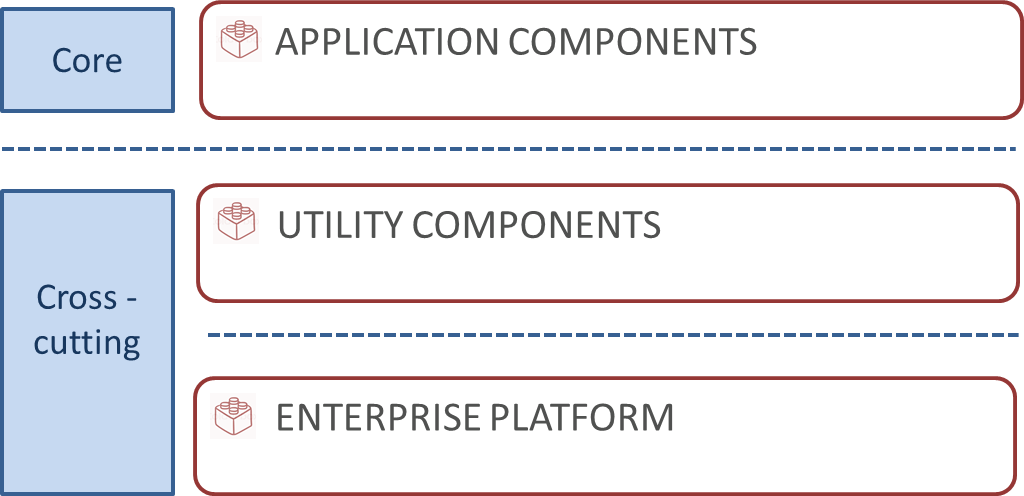


Figure 2 - EA – Application Architecture

# Conclusions

Istat has undertaken the way towards the adoption of an EA-based approach. The reference model proposed in this paper is a component of the overall strategy and can be used for a plurality of objectives, including:

* Ensuring the alignment of IT investments with Enterprise-level vision.
* Fostering standardization in statistical production processes.
* Enhancing reuse and sharing of solutions both within NSIs and at ESS-level.
* Supporting the governance of IT systems by decoupling logical architectural choices from specific technological solutions.

# References

1. ESS Enterprise Architecture Reference Framework (EARF):

https://ec.europa.eu/eurostat/cros/content/ess-enterprise-architecture-reference-framework\_en

1. TOGAF Standard, version 9.2: https://www.opengroup.org/togaf/
2. Generic Statistical Business Process Model (GSBPM), version 5.0:

https://statswiki.unece.org/display/GSBPM/GSBPM+v5.0

1. Generic Statistical Information Model (GSIM), version 1.1:

https://statswiki.unece.org/display/gsim/Generic+Statistical+Information+Model

1. Generic Activity Model for Statistical Organizations (GAMSO), version 1.1

https://statswiki.unece.org/display/GAMSO/GAMSO+v1.1

1. Common Statistical Production Architecture (CSPA), version 1.5:

https://statswiki.unece.org/display/CSPA/CSPA+v1.5