II. CSPA 2.0 Common Statistical Production Architecture

13. As part of the modernization effort, the High Level Group for the Modernization of Statistical Production and Services (HLG-MOS) wants to take action in order to address the problems and issues described in the previous section. For this reason, HLG-MOS has put priority on the development of the Common Statistical Production Architecture (CSPA) and its implementation.

14. If the official statistical industry had greater alignment at the business, information and application levels, then sharing would be easier. CSPA assists statistical organizations to address these problems by providing a framework, including principles, processes and guidelines, to help reduce the cost of developing and maintaining processes and systems and improving the responsiveness of the development cycle. Sharing and reuse of process components will become easier - not only within organizations, but across the industry, where industry is defined as a set of organizations with similar inputs, processes, outputs and goals (in this case official statistics), as a whole.

15. The value proposition of CSPA, in providing statistical organizations with a standard framework, is to:

   - Facilitate the process of modernization and support the modernization efforts within statistical organizations
   - Provide guidance for transformation within statistical organizations
   - Apply a consistent enterprise architecture approach within and across statistical organizations to respond to the challenges of emerging information sources such as big data
   - Facilitate the reuse / sharing of solutions and services and the standardization of processes, and thus a reduction in costs of production
   - Encourage interoperability of systems and processes
   - Provide a basis for flexible information systems to accomplish their mission and to respond to new challenges and opportunities
   - Leverage the wider statistical community to more rapidly develop capabilities in areas of emerging need such as the ability to harness alternative data sources
   - Enable international collaboration initiatives for building common infrastructures and services
   - Provide the ability to supplement internal capability by drawing on skilled resources from across the statistical community
   - Foster alignment with existing industry standards such as the Generic Statistical Business Process Model (GSBPM), the Generic Statistical Information Model (GSIM) and the Generic Activity Model for Statistical Organizations (GAMSO)

16. CSPA is the industry architecture for the official statistics industry. An industry architecture is a set of agreed common principles and definitions designed to promote greater interoperability within and between the different stakeholders that make up an industry.

17. A number of industry standards focusing on specific areas have already been developed. CSPA builds on and uses these existing standards, notably the GAMSO, GSBPM and GSIM. Adoption of these standards by organizations in the industry will improve the common understanding and alignment necessary for joint development, sharing and reuse of components.

18. CSPA complements and uses these pre-existing frameworks by describing the mechanisms to design, build and share components with well-defined functionality that can be integrated in multiple processes easily. CSPA focuses on relating the strategic directions of the HLG-MOS to shared principles, practices and guidelines for defining, developing and deploying Statistical Services in order to produce statistics more efficiently.

19. CSPA gives users an understanding of the different statistical production elements (i.e. processes, information, applications, services) that make up a statistical organization and how those elements relate to each other. It also provides a common vocabulary with which to discuss implementations, with the aim to stress commonality. It is an industry architecture enabling the vision and strategy of the statistical industry, by providing a clear, cohesive, and achievable picture of what is required to get there.

Scope of Architecture

20. CSPA is a reference architecture for the statistical industry. The scope of CSPA is statistical production across the processes defined by the GSBPM (i.e. it does not characterize a full enterprise architecture for a statistical organization). It is understood that statistical organizations may also have a more general Enterprise Architecture (for example an Enterprise Architecture used by all government agencies in a particular country) or specific Enterprise Architecture elements linked to government initiatives (e.g. the European framework for sharing IT services in public administrations, the European Interoperability Framework (EIF) and eGovernment Action Plans).

21. CSPA is descriptive, rather than prescriptive. Its focus is to support the facilitation, sharing and reuse of Statistical Services both across and within statistical organizations. CSPA is not a static reference architecture; it is designed to evolve further over time.

22. An important concept in architecture is the "separation of concerns". For that reason, the architecture is separated into a number of "perspectives". These "perspectives" are:

   - Business Architecture which defines what the industry does and how it is done (statistics in our case),
   - Information Architecture which describes the information, its flows and uses across the industry, and how that information is managed,
   - Application Architecture which describes the set of practices used to select, define or design software components and their relationships, and
   - Technology Architecture which describes the infrastructure technology underlying (supporting) the other architecture perspectives.

Service Oriented Architecture

23. The value of the architecture is that it enables collaboration in developing and using Statistical Services, which will allow statistical organizations to create flexible business processes and systems for statistical production more easily.

24. The architecture is based on an architectural style called Service Oriented Architecture (SOA) 1. This style focuses on Services (Statistical Services in this case). A service is a representation of a real world business activity with a specified outcome. It is self-contained and can be reused by a number of business processes (either within or across statistical organizations).
25. A Statistical Service will perform one or more tasks in the statistical process. Statistical Services will be at different levels of granularity. An atomic or fine-grained Statistical Service encapsulates a small piece of functionality. An atomic service may, for example, support the application of a particular methodological option or a methodological step within a GSBPM sub-process. Coarse-grained or aggregate Statistical Services will encapsulate a larger piece of functionality, for example, a whole GSBPM sub-process. These may be composed of a number of atomic services.

26. The granularity of Statistical Services should be based on a balanced consideration between the efficiency of the Statistical Service and the flexibility required for sharing purposes - larger Statistical Services will usually enable greater efficiency, whereas a finer granularity will allow greater flexibility for supporting sharing and reuse. Services, regardless of their granularity, must meet the architectural requirements and be aligned with CSPA principles.

27. CSPA will facilitate the sharing and reuse of Statistical Services both across and within statistical organizations. The Statistical Services that are shared or reused across statistical organizations might be new Statistical Services that are built to comply with CSPA or legacy/existing tools wrapped to be Statistical Services, which comply with the architecture. This is shown in Figure 4 by the shapes inside the building blocks.

![Figure 4: Making sharing and reuse easier](image)

**Using CSPA**

28. CSPA also provides a starting point for concerted developments of statistical infrastructure and shared investment across statistical organizations. CSPA is sometimes referred to as a "plug and play" architecture. The idea is that replacing Statistical Services should be as easy as pulling out a component and plugging another one in. There are a number of ways in which CSPA may be used by statistical organizations. These are outlined in the sections below.

**Strategic Planning**

29. If statistical organizations are creating and using an industry strategy ("Industry Architecture") and this leads to projects/work programs, they could also integrate/streamline their investment strategies. Where a statistical organization plans to contribute to and use CSPA in the future, they should modify and integrate their road maps to align with the CSPA framework. Each statistical organization needs to define a strategy to move from its current state to the common future state defined in their roadmap.

**Development within statistical organizations**

30. When a statistical organization identifies the need for a new Statistical Service, there are a number of options they can pursue. In order to fill the gap, the statistical organization can look for Statistical Services that are available in the collaborative space (that is, in the Global Artefact Catalogue).

31. If an appropriate Statistical Service is not found in the CSPA Global Artefact Catalogue, the statistical organization can either:

- advertise their need by including this information in the Service and/or Global Artefact catalogue to start collaborating, or
- start designing and developing a new Statistical Service, or
- modify an existing Statistical Service to meet new functional and/or non-functional requirements.
32. This could be done independently or in collaboration with other statistical organizations. This development work should be done in alignment with CSPA to ensure that the new Statistical Services can be added to the CSPA Global Artefact Catalogue for sharing and reuse by other statistical organizations.

33. **Sharing** means exchanging concepts, designs or software, where each user of a service creates and operates its own implementation of that service. There are levels of sharing. A limited form of sharing would be to provide another participant with the means to replicate (make a copy of) the asset (for example give the source code) (i.e. they share an aspect of the asset only). A more involved form of sharing would entail that the asset is made entirely common (in this case the asset is also reused).

34. **Reuse** means common use of a single implementation of a service, with only one organization acting as the service provider (the one who runs the service).

35. It is thought that in the current environment, it is more likely that statistical services will be shared across organizations rather than reused. Sharing between statistical organizations still leaves the option of reusing the various implementations internally within the environment of the individual statistical organization. However, as shared data platforms are being developed, reuse might become more viable in the near future.

**Vendors**

36. A statistical organization may choose to have a vendor develop a Statistical Service. A vendor in this case means either a third party commercial vendor or a statistical organization that is selling a product. In the case of a new Statistical Service, the statistical organization should request that it is built in accordance with CSPA. When the product already exists, statistical organizations should verify together if the product meets relevant community requirements. If it does not, statistical organizations can try to influence the vendor to meet requirements. If it meets the industry requirements, statistical organizations would ask the vendor to register the Statistical Service implementation to the CSPA Services Catalogue.

**Impact on organizations**

37. There will be a number of required changes and investment for an organization to modernize. Adoption of CSPA will support modernization processes with a view to generating the long term benefits identified in the value proposition (see paragraph 10).

38. The main changes required at the organization level can be grouped as:

A. **People Changes**
   - Openness to international cooperation
   - Building trust in international partners (especially as they may be building services for your organization)
   - Sense of compromise (acceptance that nothing will be optimized for local use, rather it will be optimized for international or corporate use)
   - Development of new functional roles to support use of the architecture (e.g. Assembler, Service Builder)

B. **Process Changes**
   - Adoption of an industry wide perspective
   - Different approach to business process management and design
   - Commitment to service (contract between different functional units)

C. **Technology Changes**
   - Setting up an adequate middleware infrastructure (messaging, repositories)
   - Uplift of physical network capabilities (bandwidth, etc.)
   - Management of security features

39. In addition to the costs and the targeted benefits, an organization adopting CSPA will benefit from:
   - A sustainable and efficient strategy to cope with legacy and phasing out of existing applications
   - A cycle that enables cost saving from reduction in production costs to be reinvested in further infrastructure transformation
   - A positive image both on national and international/industry scene

---

1. CSPA Guidance recommendation: To assess the SOA readiness of a statistical organization the following maturity assessment is available: The Open Group Service Integration Maturity Model (OSIMM) (accessed 22\textsuperscript{nd} December 2014) \\[link\]