Working Paper 17.

UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Work Session on Statistical Data Editing
(The Hague, Netherlands, 24-26 April 2017)

Validation, Shared Services and Enterprise Architecture: How it fits

Prepared by Lucas Quensel-von Kalben, Federal Statistical Office, Germany

I. Introduction

1. Official statistics in the last decade is heavily investing in new approaches to its technical environment. National and international initiatives have been launched to streamline statistical production and standardize processes and application for quality gains and higher efficiency. This is altogether not new.[1] What is new is that the efforts in international organizations, national institutes and other national authorities that produce official statistics are better aligned than ever before. This alignment derived from realization that we all face similar problems and that a collaborative approach might save costs although it raises coordination efforts.

2. Collaboration, beside bilateral agreements, is mainly initialized by international organizations such as the UNECE, OECD or Eurostat. This is not driven by purely altruistic motives but should help to lower the burden in these institutions and increase the comparability (one key feature of data quality) of national statistics. Drivers are the High Level Group for the Modernisation of Statistical Production (HLG-MOS) and the ESS Vision 2020. Both initiatives identified similar instruments/concepts as framework for the standardization efforts as a common baseline for collaboration.[2] They have not “reinvented the wheel” but used ideas developed outside official statistics and adapted them to this domain (UNECE1 2015).

3. Two concepts are of more general interest to methodologists working in statistical data editing. These are Enterprise Architecture and Service-oriented Architecture. Enterprise Architecture (EA) is a common methodology for aligning business needs and technical solutions .

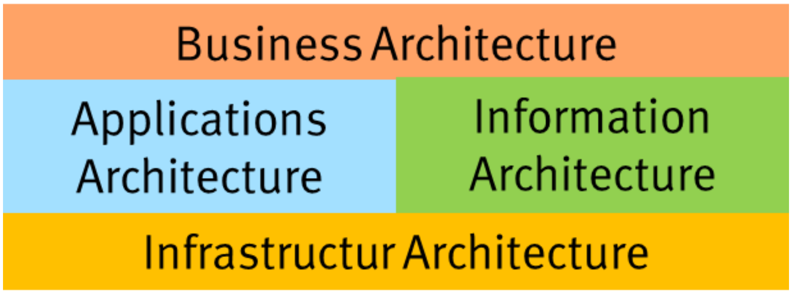

Figure 1

4. The specific Business architecture of official statistics is well known in the statistical community and is called GSBPM (Generic Statistical Business Process Model) (UNECE1 2016). The information architecture is less well known and has been defined as GSIM (Generic Statistical Information Model)

[1] The statistical data editing community and its work sessions aim for the same goal. The focus here is more on technical infrastructure and the application landscape and less on methodological issues.
[2] Collaboration can only start when all partners (i.e. international organisations, NSIs and ONAs) have agreed on a common business process, formats and IT-structures. This is what we call standardization

(UNECE 2017). The specific application architecture is known as CSPA (Common Statistical Production Architecture) (UNECE2 2016). Within the ESS similar concepts do apply.[3] The differences are very slight and for the argument not relevant. The level of abstraction from the general framework to individual application in a specific domain/ organization is not discreet but a continuum. On each level alternatives to the approaches chosen exist. In some cases they will be mentioned in the paper.

II. Business Architecture

5. Business Architecture is answering the question of what we do when we do official statistics. Everyone working in official statistics has a basic intuitive understanding of the statistical process. GSBPM and its national equivalents try to model and standardize this process to a level of abstraction that differentiate it from processes of other domains and is still applicable to statistics across domains. GSDEM (Generic Statistical Data Editing Models[4]) (UNECE 3 2015) is a level below this layer (i.e. more concrete) and probably more helpful for domain specialists and methodologists.[5] The methodological handbook of the ESSnet ValiDat Foundation, a project of the European Statistical System in 2015, is even more detailed and explores several methodological aspects involved in validation (De Zio et al. 2016).

6. This "drill-down" from the more abstract to rather pragmatic conceptual frameworks can be very helpful for identifying business functionalities that needs to be supported by IT-systems. Jumping directly to business functions could lead to ignore the "big picture".[6] Structural validation is a business function of validation (error detection), that is an activity strongly connected to other activities (error localization, editing, imputation). Structural validation can be applied in several sub-processes of the GSBPM (mainly, but not exclusively: 4.3 Run collection; 5.3 Review and validate; 6.2 Validate output). This splitting up of the validation process in different phases of the business process should be carefully planned.[7]

7. Another feature of the big picture is the dependencies of validation to other sub-processes. Validation will use the output of previous processes and be the input of following ones. And, to make it even more complicated, GSBPM never stated that the sequence of phases and sub-process is a linear one from 1.1 to 8.3. Loops and decisions in the process are not deviations from the pure model but an inherent feature.

8. This all seems straight forward and not altogether new, but violations of these principles are common. In the next layers of the EA-model, we will see how similar principles apply.

III. Information Architecture

9. Having identified what we do in official statistics, we can go on to identify what is the "material" we are dealing with. It is information or more materialistically speaking data and metadata. In terms of the GSBPM and other – more detailed - process models: What are input, throughput and output of each individual process. The equivalent in terms of level of abstraction to GSBPM is GSIM. The Generic Statistical Information Model is divided into 4 main groups of information objects (Business, Concepts, Exchange, and Structures). It is complete and coherent. I.e. it claims to identify all information objects (more than 170) necessary for GSBPM level 1 and 2.

---

[3] Enterprise Architecture Reference Framework (ESS 2015), Statistical Production Architecture (ESS 2016), SERV (ESS 2012)
[4] The author always wonders about the "models" instead of "model"
[5] Version 6 of GSBPM is trying to identify a further layer of the process (Level 3: Activities) that might lead in the same direction.
[6] The continuum of business perspectives is reflected by a simple analogy from real architecture: You would not build a family house in an industrial area,
[7] The ESS Task force on validation has proposed 6 very general principles for the validation process (TF Validation 2017). The most important one is probably the principle "the sooner, the better" that highlights on the efficiency of the validation process in early phases of statistical production or even earlier in the conceptual phase of the GSBPM (Phase 2: Design)

10. Comparable to the discussion of the layers in the statistical business architecture, GSIM is rather abstract. It can be mapped more or less successfully to more specialized information models that are in use in official statistics. Examples in the field of metadata are mappings to Neuchatel or ISO 11179. For information exchange physical implementations like SDMX or DDI are relevant.

11. For the construction of IT-systems these models are either to general or to restricting. Some intermediate level is necessary. Bridging the gap between conceptual and physical implementations another model has been developed, that is coherent with GSIM and can be transformed to models like SDMX or DDI. This so called Logical Information Model (LIM) (UNECE 4 2015) is of paramount importance to allow business processes/functions and IT-services to “speak-to-each-other”.

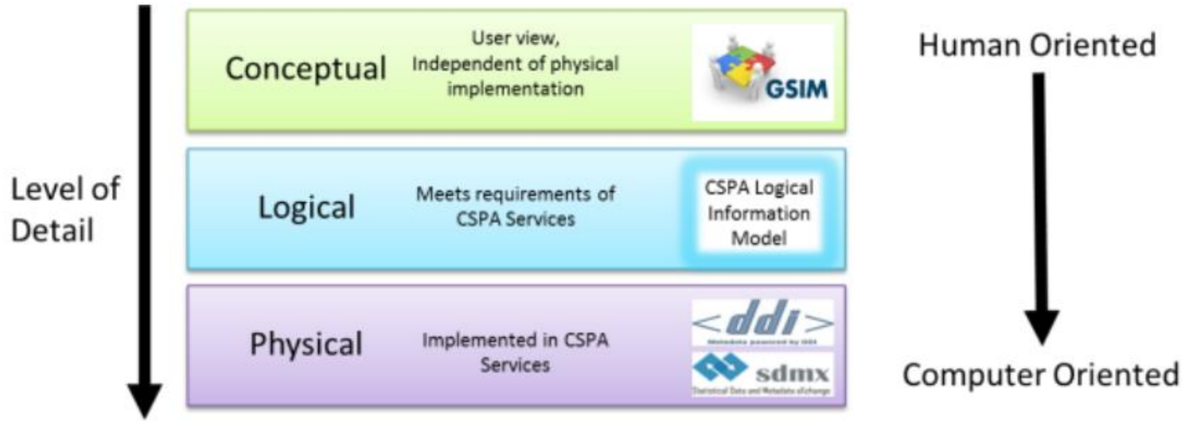

Figure 2 (UNECE 4 2015)

12. LIM is rather complex and not in the least complete yet. The CSPA Working group chose a rather pragmatic approach to elaborate the model in future. It will be incrementally extended when further information objects are needed in practice because a new IT-service has been identified and should be made available widely. Then the necessary information objects of GSIM will be refined and made adaptable to physical representations.

13. An example from validation might highlight this rather abstract presentation. An IT-service doing any kind of validation uses validation rules and raw data as input and returns validated data and a validation report as output. This is certainly an abstract version of the information objects needed in the process. From the business perspective this model is rather fine if not complete yet. Other information objects might be important for the process (e.g. classifications, validated data as comparative material, processual metadata and paradata etc.)

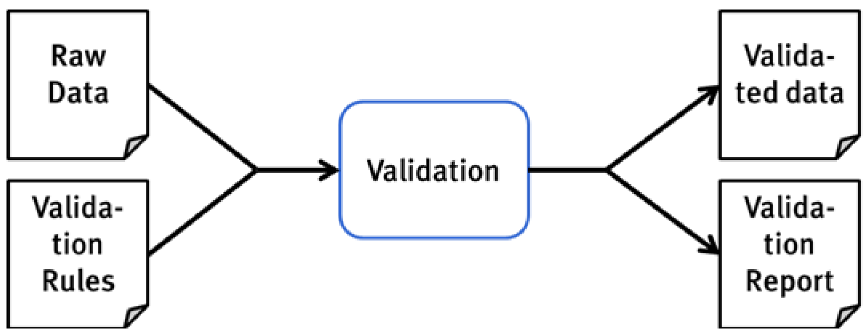

Figure 3 (Weichert 2017)

14. From an IT-point of view the model is also not specific enough. They do need a kind of technical interface description. What is the structure of the raw data?[8] How are validation rules represented?[9]

[8] Structure does reflect different concepts. On the one hand it is referring to a physical format of storage and exchange (relational tables in a SQL-database, flat files (CSV, XML others), specified XML-exchange formats (SDMX, DDI, national representations). On the other hand it does reflect the individual representation of a statistical entity. Here both meanings are applicable. GSIM is quite explicit on both meanings and helps to differentiate between these usages. In validation this distinction is expressed as structural versus content-based validation.

15. What is needed here is a common representation of the information model that can be talked upon by methodologists/subject matter specialists on the one side and IT-specialists on the other side. This representation must map to more abstract as well as more concrete presentations. This intermediate level is not only important as “lingua franca” between different groups but also for exchangeability of services between different institutions using different concrete data and metadata structures. A more abstract level would cause significant adaptations in the implementation of a service in a given (IT-)environment, a more concrete one would hinder organisations in its reuse. But here we have already reached the level of the application architecture.

IV. Application Architecture

16. The next perspective in Enterprise Architecture is the structure of the IT-applications and its integration into a wider system. This seems to be a straight forward thing to do. Solutions to the question of how business needs/requirements can be fulfilled by an IT-system are indeed manifold. The evolution of application architecture types though times and the current usage of these types in statistical organizations can be organized into different groups. Each of these solutions has its advocates and its individual advantages and disadvantages. The following types[10] are probably not a complete overview of possible solutions not even in the field of official statistics, but give an idea about its diversity. Two aspects are of major importance: Where are the data located and how is the control flow organized?

I) Monolithic application: Stemming from the period of the mainframe computer, the main idea is that different processes are organized in a batch job. The control flow is the sequence of jobs, data output of one job is data input of the next. A modern variant of this are national systems that cover most (at least in theory) of the complete production process of all or most statistics (often business statistics).[11]

II) Standard applications: Another approach is the usage of general purpose applications for the manipulation of data. These can be desktop or server systems. Typical examples are SAS or SQL-Databases, MS Office products or the usage of R. Data will be passed from application to application or manipulated in one location. The control flow is regulated by the users.

III) Specialized applications: Typical examples of this approach are applications in a high level programming language like Java or Microsoft .NET. They are usually organized as classical Client-Server architecture where the graphical user interface (GUI), business logic and persistence (Database) are formally divided. The applications do usually cover only one up to several statistical surveys in the same field (e.g. household surveys).

IV) Tool chain: Some organizations have split the process of statistical production in parts (even before the publication of GSBPM). They started by creating products for a specific business purpose and breaking the boundaries of silos (= statistical clusters). Examples are packages like CANCEIS or Tau-Argus, programming systems like Blaise or Manipula, informations systems for data dissemination like PC-axis and OECD’s .Stat. Data and control flow are

---

[9] The discussion of VTL as the new validation language in the ESS is here of interest (SDMX 2017; Gelsema et al. 2015)

[10] The UNECE and Eurostat are developing frameworks (Modernisation Maturity Model (UNECE 2 2017); MS Benchmark (Eurostat 2017)) that are similar to the proposed typology here. They do assume that the sequence of these types represent a kind of evolutionary process that ultimately leads to a “true” SOA-like Architecture.

[11] Many NSIs are currently (!) constructing such big systems that cover several sub-processes of phase 4, 5 and 6 of GSBPM, i.e. data collection, data procession and data analysis. The sub-processes are usually not separated technologically and dependencies between functionalities are huge. Sharing on this level is not possible or restricted to a complete usage of the whole system. Several of these have been presented in the Modernisation workshops of the UNECE (UNECE 2014 and UNECE 5 2015). In the authors view it is strange that these national systems are presented side-by-side to approaches arguing for international cooperation and sharing of services. These strategies are virtually excluding each other. It is an interesting question in itself why the national monolithic systems are still been developed and forces not joined for common solutions. Reasons might include lack of knowledge of this new application architecture and its merits, doubts about its feasibility in terms of speed of development and dependencies, subjective reasons (“not-invented-here” syndrome) and others.

handled very differently and the establishment of metadata systems for automation of processes is a great issue.

V) Service oriented architecture (SOA): This approach has been started in the 1990-ies and trickled into the domain of official statistics since early 2000-nds. Methodologically services are autonomous and stateless building blocks that represent generic functions in the business process. Technically they work in independent systems and use webservice-interfaces. The control flow between different services is organized by a process orchestrator. Data can be stored centrally or locally.

VI) Hybrid approaches: Most national and international systems for statistical production have been developed for decades. Legacy applications “that do the job” are widespread and coexist with other approaches side-by-side. Many institutions have created migration paths that itself will run for a long time. It is possible to integrate tools and services in a specialized application or wrap such applications as a service and integrate it into a SOA.

17. Common Statistical Production Architecture is based on SOA (type V). Building blocks are therefore the individual services and their integration into a whole system. Sharing services, one of the main goals of CSPA, is possible when services comply with some inherent and explicit rules. One issue is the question of granularity: How big should an individual service be? There is a lively discussion on this topic. Another major feature of a service is its statelessness. This means that a service does not know its position in the process chain and will not know anything of the control flow. The service does not know what other services are available and is not dependent on these. The process owner (orchestrator) is responsible for the correct usage of the service. The interface between orchestrator and each service is a contract saying given the right input you can rely on the right output. What happens inside is a black box. The essence of the contract is defined by the Logical information model (LIM).

18. Ideally CSPA-compliant services can be easily “plugged” into greater systems. In reality sharing of services is still in its infancy and first experiences show how difficult this can become. Some projects are currently running to integrate services into a foreign environment. Pilot implementations of services[12] in a European Commission sponsored project (ESSnet Sharing Common Functionalities in the European Statistical System) (ESS 2 2016) are very valuable and will be presented in early July 2017 in an international workshop in Wiesbaden.

V. Infrastructure architecture

19. This lowest layer of the Enterprise Architecture refers to the physical environment that is necessary to run the applications. From a methodological perspective this is rather independent of the business layer and only seldom of direct relevance. Almost all the above mentioned scenarios (I to VI) can be run on more or less the same physical infrastructure. This can be a local cluster of some medium sized servers up to mighty cloud services far removed from the customers. Looking specifically on a Service-oriented Architecture, non-functional requirements like performance, stability, availability, IT-security, data protection and confidentiality, political aspects like the dependencies on other institutions might influence decisions on the usage of foreign services and especially their location. The ESS (ESS 2015) developed four types of usage of services: Autonomous, interoperable, redundant and shareable. The different scenarios of usage will mainly be guided by the above mentioned requirements and not by pure functionality.

[12] Three different services are tackled: A service for seasonal adjustment based on JDemetra+, a service for the generation of questionnaires based on a DDI data structure definition and a third service on metadata dissemination.

VI. Technical approaches to validation in official statistics

20. In early 2015 the ESSnet ValiDat Foundation undertook a survey on validation approaches in the European Statistical System (Gießing & Walsdorfer 2015). The survey covered several aspects of the validation project.
21. From a technical perspective different types of validation procedures could be detected and classified in the above mentioned application architectures.[13] [14]

1) The first (and growing) type is the "Generic National Statistical Systems" (or monoliths) from above. They very similar to the special purpose applications below but claim to cover nearly all domains.
2) By far most organisations use general purpose tools (SQL, SAS, R, Excel and Co.) and apply them more or less directly to the data stored in files or individual databases. It requires a good understanding of the domain and a fair competence in writing validation rules. Let us call this approach the "ad-hoc or direct"-approach.[15]
3) A third type is the development of special purpose application for individual domains and surveys. Here the validation rules are usually specified by the subject matter departments and either directly implemented in the application or configurable by a multi-purpose (i.e. SQL) or specialized language.
4) Type 4 is the usage of special purpose validation tools cross cutting domains for typical business processes. Eurostat's EDIT tool is a typical example.
6) A last group consists of hybrid solutions. All kind of mixtures of approaches are possible. For example the data validation in the data collection phase is done jointly across domains by one system, data processing (GSBPM phase 5) individually by specialized applications and output validation (GSBPM sub-process 6.2) with general purpose tools.

22. All of the above mentioned approaches have their merits (and some major drawbacks[16]). They represent the different architectural patterns on the application layer mentioned above.

23. The solution propagated by the HLG and the ESS Vision 2020 is using CSPA, i.e. a service oriented architecture. Here building blocks allow a more fine-tuned adaptation to the specific requirements of individual domains. Currently no or very few organisations have made their application architecture CSPA compliant (or at least conforming to the more general SOA pattern). The contradiction between status-quo and state-to-be is obvious. The advantages of CSPA (see above) are very striking, but do not seem to be common knowledge yet.

24. Currently a project (ESSnet ValiDat Integration) tries to implement validation services into three different national infrastructures the "CSPA-way" (ESS 2017). The business focus of the project is validation of data prior to its transmission to Eurostat.[17] Business driver is avoiding data ping-pong

[13] This came as some kind of surprise when automatically assuming that every institute involved in official statistics follows a similar approach (the quite "natural" of once own office)!
[14] The numbering of the general and the validation specific application architectures is directly mapped. Leaving out type 5 (specialized architecture) is on purpose
[15] To be fair: The rules will often be developed only once and then applied many times to the material. The essence of this approach is still a trial-and-error approach.
[16] The direct approach (type 1) for example is not really production safe. Access control, logging and quality indicators cannot be centrally managed and monitored. Type 2 (specialized applications) might have the disadvantage of different methodologies applied in different domains. Problems with type 3 (Cross-cutting tools) could arise by trying to implement a validation strategy across different phases of GSBPM. The national monolithic systems are better equipped to these problems but are not particularly modification friendly. The pretension to cover all domains and all necessary functional requirements can usually not be fulfilled. Type 5 inherits nearly all of the disadvantages of the other types.
[17] Quensel-von Kalben 2015 explored the specific requirements of the member states in respect to this so called narrow focus as well as a wider focus ‚from-end-to-end'

between these organizations by using the same validation rules on either end. The business process is described in the "business architecture" (TF Validation 2017a). Eurostat and the ESS member states do agree on a set of common rule sets in a specific domain. These rules are expressed in VTL (and plain English). The rules are stored centrally in a common repository that can be accessed from systems within the ESS. IT-services will then be able to download rule sets and process them against nationally produced data. The services will provide a validation report that can be used to edit erroneous data manually or automatically.

25. This rather simple business architecture opens up a lot of questions. The business architecture can be supported by several forms of application and information architecture. In our case the application architecture is based on architecture of services (other solutions are possible). Eurostat (Gramaglia 2015) identified three main scenarios of the use of services. They do differ in respect to location of the control flow and individual services. In scenario 1 all control is on member state side. The process orchestration as well as all services are hosted in the member state premises. Scenario 3 is the complete opposite. Here data are transmitted to a central system (i.e. Eurostat) and processed there. Only the validation report and a subsequent data editing has to be done at the member state level. Scenario 2 is somewhere in between the other two. The process orchestrator is firmly in the hand of the member state. Individual services can be located on either side (local or remote). This is what we call Infrastructure Architecture.

26. These three scenarios have (again) benefits and drawbacks. The ESSnet is currently exploring which scenario suits which kind of organization most (based on the architecture the MS has implemented for its internal production chain) and will give some hints of implementing this architecture (Weichert 2017)

27. But are there any services available that can already be integrated into such kind of scenarios? Probably yes. Most of them are not yet CSPA-compliant or not even web-services (the technical solution behind CSPA), but could be used with some modifications.

28. Three examples (applications, not services) from the field of data editing and validation shall illustrate the current state in the member states of the ESS.

1. Autonomous validation tool: The R-package Validation (CBS)
2. Integrated validation infrastructure: eStatistik (Destatis)
3. Integrate-able autonomous services: StruVal and ConVal (Eurostat)

29. Mark van der Loo's Validation package (Van der Loo & de Jong 2016) has been developed in the years 2014/15 and is publicly available. The philosophy is to create an "embedded Domain Specific Language" that uses the R-specific language features and its graphical presentations (plug-ins). The application runs on desktop computers and requires not many resources (depending on the data to be processed).

30. eStatistik, the German system, is based on the above mentioned tool chain approach (Destatis 2017). It has been developed for several reasons. One is to ensure the usage of the same validation rules across the German Statistical System (a kind of structure similar to the ESS), the re-use of validations rules across different domains, and the integration of these rules in different phases of the GSBPM (4, 5 and 6). Three so called Editors are being used for the specification of surveys. In this context the most relevant one, is the PL-Editor[18] to specify logical data models, its mapping to a physical one, and validation rules. The PL-Editor can export its specification as Java code that can be integrated in specific Java-applications (the type 2 from above) or into a specific XML schema. This specification is uploaded into a survey registry together with the other specifications. From there it can be downloaded to cross cutting tools for data entry like webforms and modules for ERP (phase 4), for data processing tools in the offices (phase 5) or for output validation prior to transmission to Eurostat. The tools can interpret the specifications directly and apply to the appropriate data. All this is only part of the story but the one most relevant in this context.

[18] PL for Plausibility Checks

31. Eurostats StruVal and ConVal are services that are planned to be the core of data validation prior to transmission from ESS Member States to Eurostat (Peyronnel 2016). StruVal, a service for structure validation (or so called level 0 validation in the terminology of the methodological handbook), is based on the SDMX-converter. This has been wrapped as a web service and made available for some domains (National Accounts) already. It is currently only hosted at Eurostat but should make available for replication in other environments later in 2017. ConVal, content based validation or level 1+ validation, is a service in development. Currently it is based on an earlier solution of Eurostat called EDIT which was distributed as a desktop application for some years. The validations rules are specified in a format called EDIT-script. It is planned that VTL 1.1 should replace EDIT-script soon. Both services are restricted to the functionalities and the output of their former products (SDMX-converter and EDIT).

VII. Comparison

32. All three examples have their individual merits and drawbacks.

33. Validation is light weight, easy to use, easy to modify and still quite powerful. In its current state it can easily implemented and used in statistical environments. Its drawback is currently the integration into larger “industrial” production chains with its additional functional and non-functional requirements (performance, interoperability and automation, IT-security etc.). From a CSPA-point of view Validation can be used as a stand-alone solution for explorative validation research or – by wrapping it as a service – integrated into wider systems.

34. eStatistik is rather heavy weight and an ecosystem in itself. It works perfectly within this environment and is currently a good solution for the German Statistical System. Its Application Architecture is not strictly service oriented (decoupled) but rather strongly connected with each other. This makes its usage in another environment or the replacement of individual tools in the toolchain difficult. The flow of control is not localized in an independent orchestrator but distributed across the system. To make it shareable (one important goal of the CSPA-compliance) tools need to be isolated and made available as individual stateless services (wrapped).

35. StruVal and ConVal do almost already comply with the CSPA’s specification. They are constructed as CSPA services and addressed by a CSPA orchestrator. However, their use is rather restricted in terms of the proposed business sub-process and is currently still “experimental”. They rely on specific data structure definitions (SDMX) and are only available as “shared” services meaning they are located at Eurostat.

36. We are aware of the fact that many more tools and services are available that cover the field of validation and data editing extensively. Currently no systematic overview of these tools is available. The three examples provided here are representatives of different types of strategies in the field of official statistics. Probably other types could be detected.

37. A roadmap to concentrate forces for common solutions in the field of validation could start at either end of Enterprise Architecture and even in parallel: Elaborate on the business architecture and look for common patterns. Think about the necessary informations objects and create common information architecture. Identify existing solutions that can be made into services and develop best practice approaches to their integration into existing systems.

VIII. Conclusion

38. Enterprise Architecture is a useful framework for structuring alignment between business needs (including methodological improvements) and IT-solutions. This relates to national as well as international modernization activities. The last years have brought a much better understanding of the business architecture of official statistics. Conferences like the SDE provide valuable insight in the structure and processes of official statistics business needs. The underlying levels especially the information model and the application architecture are still in an initial stage and not as advanced as

sometimes optimistically thought. Having defined a sound theoretical model for these layers now is time for practical implementations.

39. Validation and data editing are core processes in statistical production. In the past each national statistical institute developed its own IT-solutions to this common problem. This is neither cost effective nor do these solutions comply with some aspects of data quality ("comparability" for example). Sharing IT-services could be a solution to these challenges. However, it is not an easy undertaking and requires a further standardization of the business architecture, as well as the underlying layers.

40. Interdisciplinary cooperative work and a common understanding of some concepts are key elements of success. This is altogether not new and applied for national initiatives before. The international focus on collaboration, where IT-solutions might not just be used by individual institutions, demands that kind of "joint venture" even more.

IX. Literature

Destatis 2017: eStatistik – A toolchain approach for statistical production in the German Statistical System (forthcoming in ESS 2017)

De Zio et al. 2016: Methodology for Data Validation (Handbook) (https://ec.europa.eu/eurostat/cros/content/methodology-data-validation-10-handbook-revised-edition-june-2016_en)

ESS 2012: Christine Wirtz: Shared Services (https://ec.europa.eu/eurostat/cros/content/ess-vip-shared-services_en)

ESS 2015: Enterprise Architecture Reference Framework (https://ec.europa.eu/eurostat/cros/content/ess-enterprise-architecture-reference-framework_en)

ESS 2016: Statistical Production Reference Architecture (https://ec.europa.eu/eurostat/cros/content/spra_en)

ESS 2 2016: Sharing Common Functionalities in the ESS (https://ec.europa.eu/eurostat/cros/content/sharing-common-functionalities-ess_en)

ESS 2017: ESSnet VaiDat Integration (https://ec.europa.eu/eurostat/cros/content/essnet-validat-integration_en)

Eurostat 2017: ESS Member State Benchmark Assessment (forthcoming)

Gelsema et al. 2015: A Study of VTL (http://ec.europa.eu/eurostat/cros/system/files/a_study_of_VTL-2015-11-06.pdf)

Gießing & Walsdorfer 2015: Validation in European Official Statistics: Results of an ESS survey (http://ec.europa.eu/eurostat/cros/system/files/2%20-%20ESSnet_ValiDat_survey.pptx)

Gramaglia 2015: Towards a European Validation Architecture (http://ec.europa.eu/eurostat/cros/system/files/1%20-%20Wiesbaden%20Workshop.pptx)

Peyronnel 2016: Eurostat Services (https://www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwib9_m0b_SAhXHaRQKHc75AIoQFggjMAA&url=http%3A%2F%2Fwww1.unece.org%2Fstat%2Fplatform%2Fdownload%2Fattachments%2F123144126%2F5.%2520Day%25202%2520%2520EurostatServices_Pierre.pptx%3Fversion%3D5%26modificationDate%3D1466683955531%26api%3Dv2&usg=AFQjCNFSaJpPi0lJamfKslz6XKGENxiPaw&cad=rja)

Quensel-von Kalben 2015: Mutual interests - Requirements from a member state perspective (http://ec.europa.eu/eurostat/cros/system/files/3%20-%2020151107%20Mutual%20interests.ppt)

SDMX 2017: Validation and Transformation Language (https://sdmx.org/?page_id=5096)

TF Validation: Validation Principles (forthcoming, part of a new version of De Zio et al. 2016)

UNECE 2014: MSIS-Workshop Dublin & Manila (http://www.unece.org/stats/documents/2014.04.msis.html)

UNECE 1 2015: Business Architecture (http://www1.unece.org/stat/platform/display/CSPA/Business+Architecture)

UNECE 2 2015: Developing a Logical Information Model for the Common Statistical Production Architecture (http://www1.unece.org/stat/platform/display/CSPA/Developing+a+Logical+Information+Model)

UNECE 3 2015: Generic Statistical Data Editing Model (http://www1.unece.org/stat/platform/pages/viewpage.action?pageId=117774163)

UNECE 4 2015: Logical Information Model (http://www1.unece.org/stat/platform/display/CSPA/CSPA+Logical+Information+Model)

UNECE 5 2015: Workshop on the Modernisation of Statistical Production (http://www1.unece.org/stat/platform/display/hlgbas/Workshop+on+the+Modernisation+of+Statistical+Production)

UNECE 1 2016: Generic Statistical Business Process Model (http://www1.unece.org/stat/platform/display/GSBPM/Generic+Statistical+Business+Process+Model)

UNECE 2 2016: Common Statistical Production Architecture (http://www1.unece.org/stat/platform/display/CSPA/Common+Statistical+Production+Architecture)

UNECE 2017: Generic Statistical Information Model (http://www1.unece.org/stat/platform/display/gsim/Generic+ Statistical+Information+Model)

UNECE 2 2017: Introduction to the Modernisation Maturity Model and its Roadmap (http://www1.unece.org/stat/platform/display/RMIMS/Introduction+to+the+Modernisation+Maturity+Model+and+its+Roadmap

Van der Loo & de Jong 2016: Data validation infrastructure: the validate package, (http://www.markvanderloo.eu/ files/statistics/vanderloo-useR2016.pdf)

Weichert 2017: The ESSnet ValiDat Integration (http://www1.unece.org/stat/platform/download/attachments/125436234/SDE_ESSnet_ValiDat_Integration.doc?version=1&modificationDate=1487958644088&api=v2)