



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Département fédéral de l'intérieur DFI  
Office fédéral de la statistique OFS

# Statistical Metadata System based on SDMX

Petko I. Yanev; Jean-François Fracheboud

Federal Statistical Office  
Switzerland



## Statistical Metadata System : Content

- Challenges
- Vision / Metadata Strategy
- Architecture and coverage
- Funtionalities
- Plug-in: Statistical Codebook



## Statistical Metadata System (SMS) : Challenges

- Centralized system for de-centralized metadata management
- Unique metadata creation : multiple usage
- An integrated system for the entire life-cycle of each statistic
- The memory of statistics



## Statistical Metadata System (SMS) : Challenges

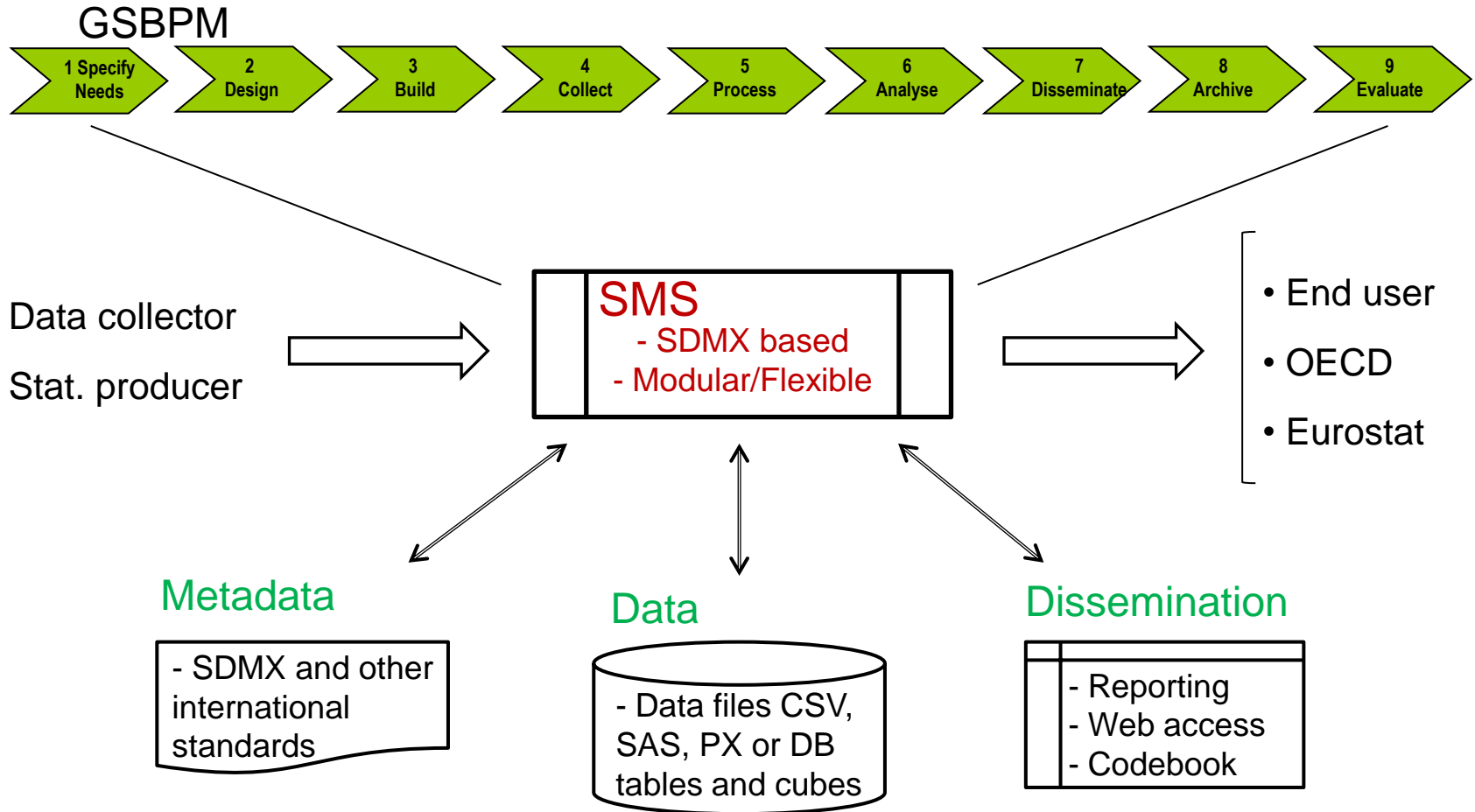
- Centralized system for de-centralized metadata management
- Unique metadata creation : multiple usage
- An integrated system for the entire life-cycle of each statistic
- The memory of statistics

**SMS MEETS ALL THESE CHALLENGES**

Building SMS since 2010; Operational since 2012

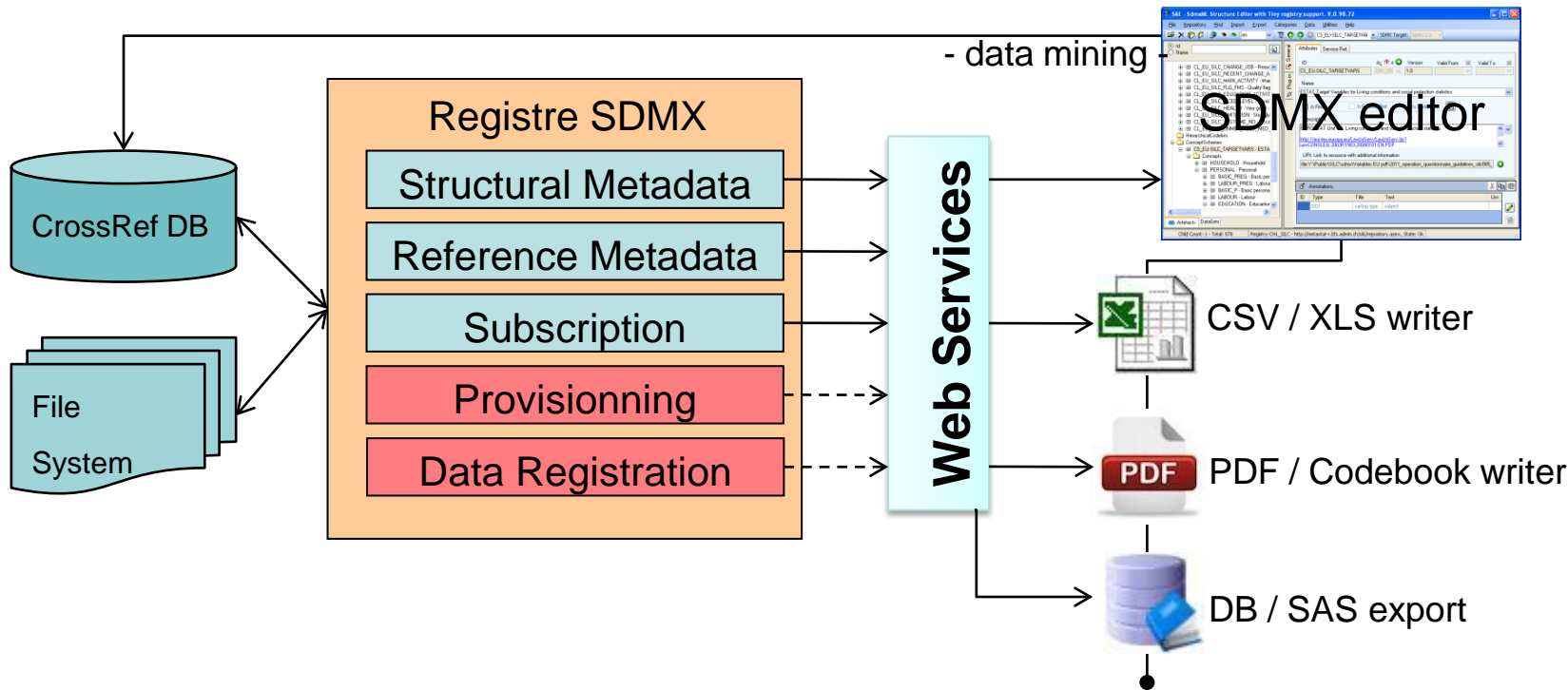


# Statistical Metadata System : Vision





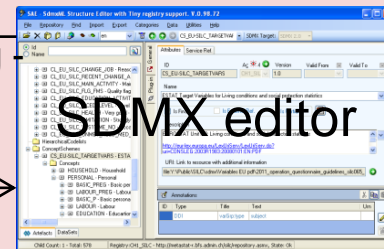
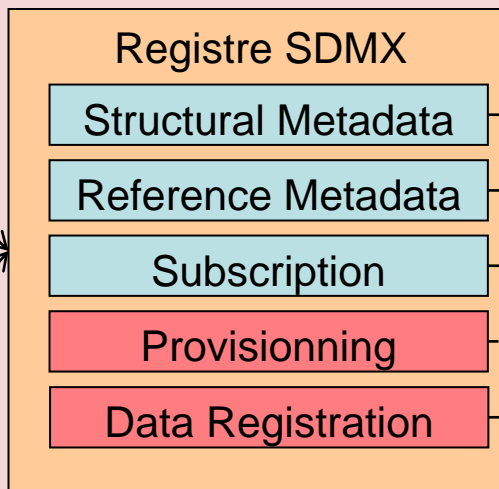
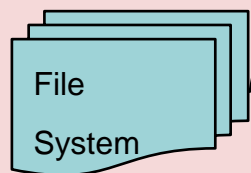
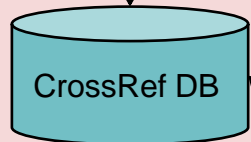
# SMS : Architecture and Coverage



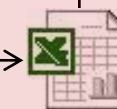


# SMS : Architecture and Coverage

SMS



- data mining

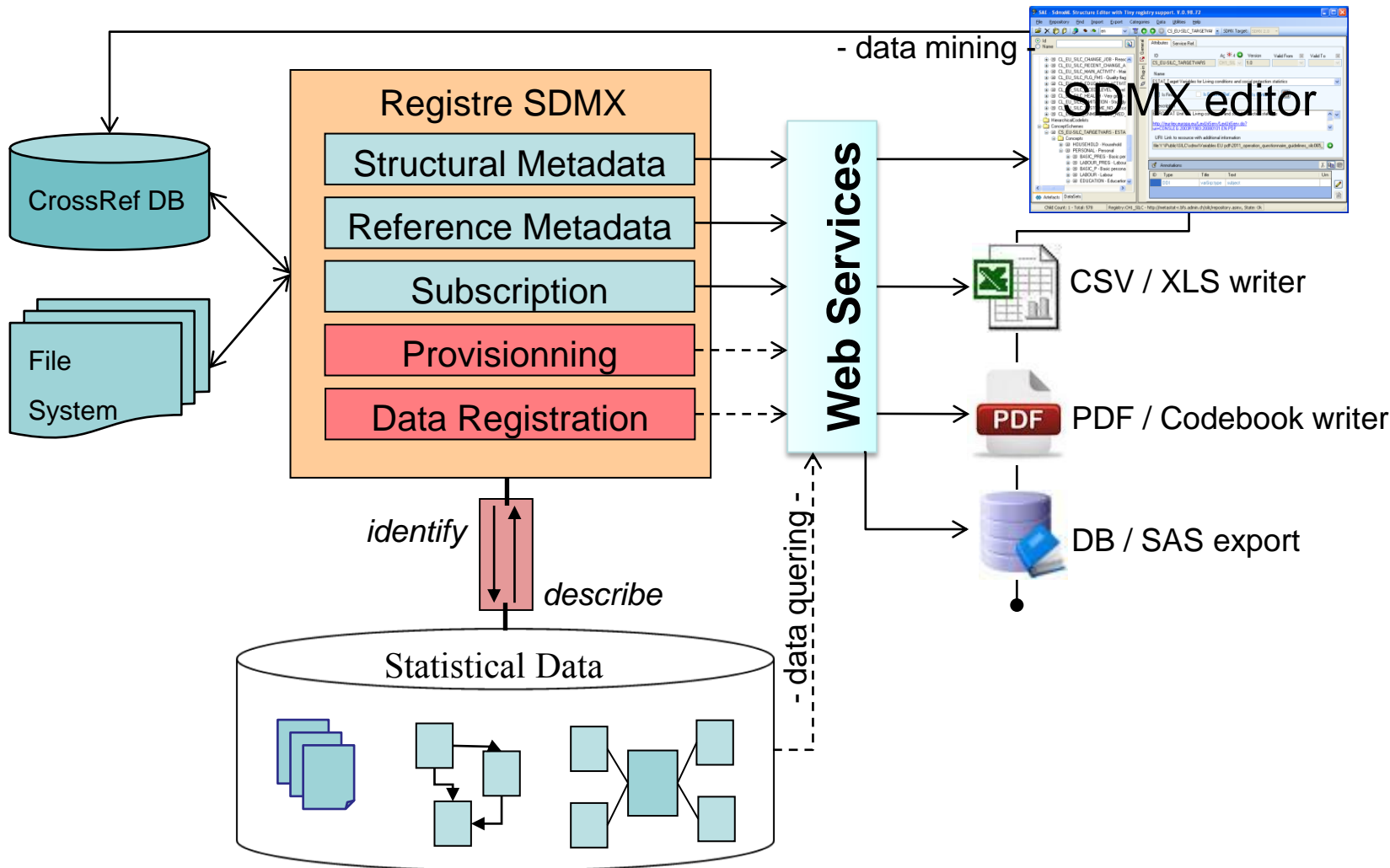


Statistique suisse



# SMS : Architecture and Coverage

Statistique suisse

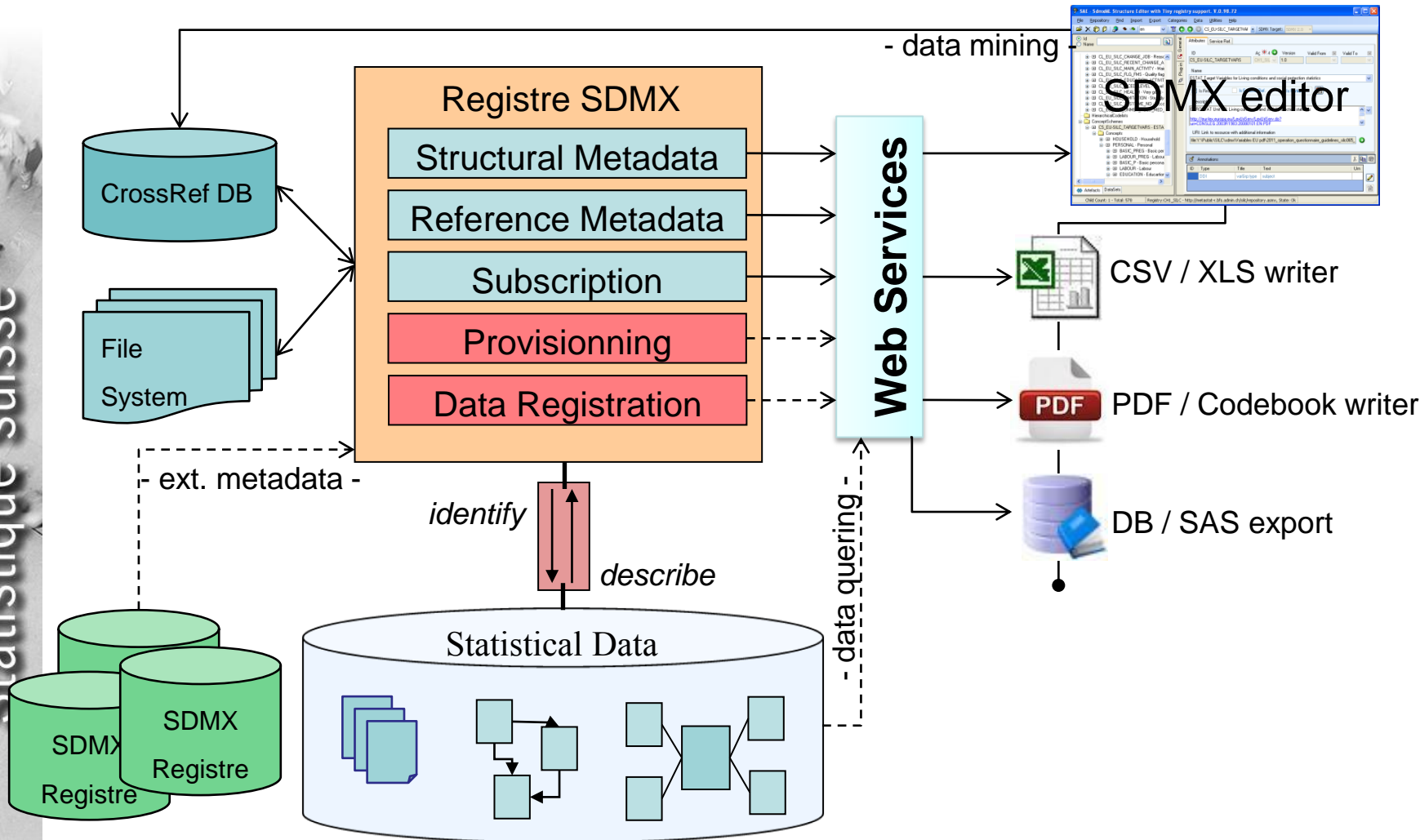






# SMS : Architecture and Coverage

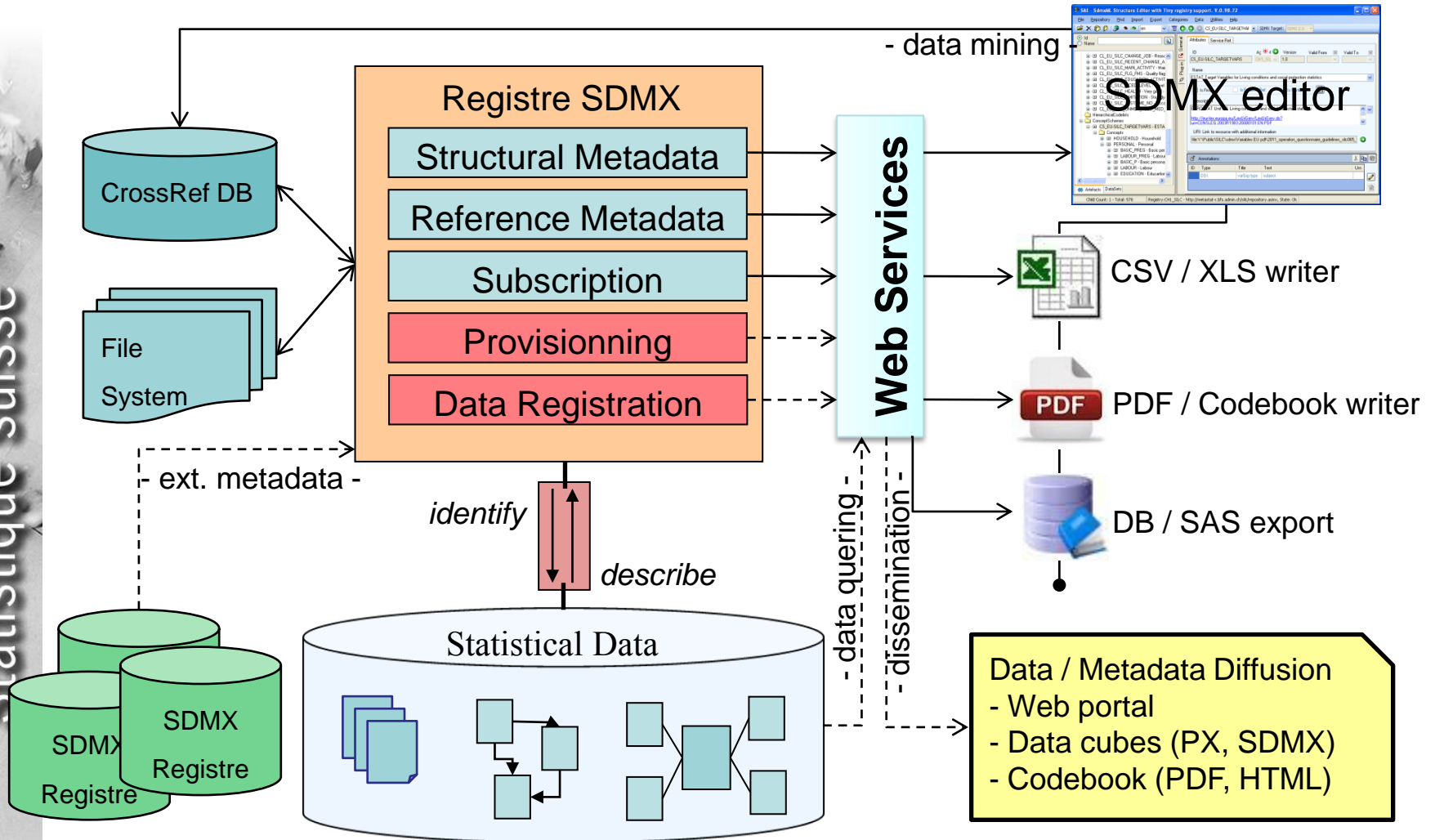
Statistique suisse





# SMS : Architecture and Coverage

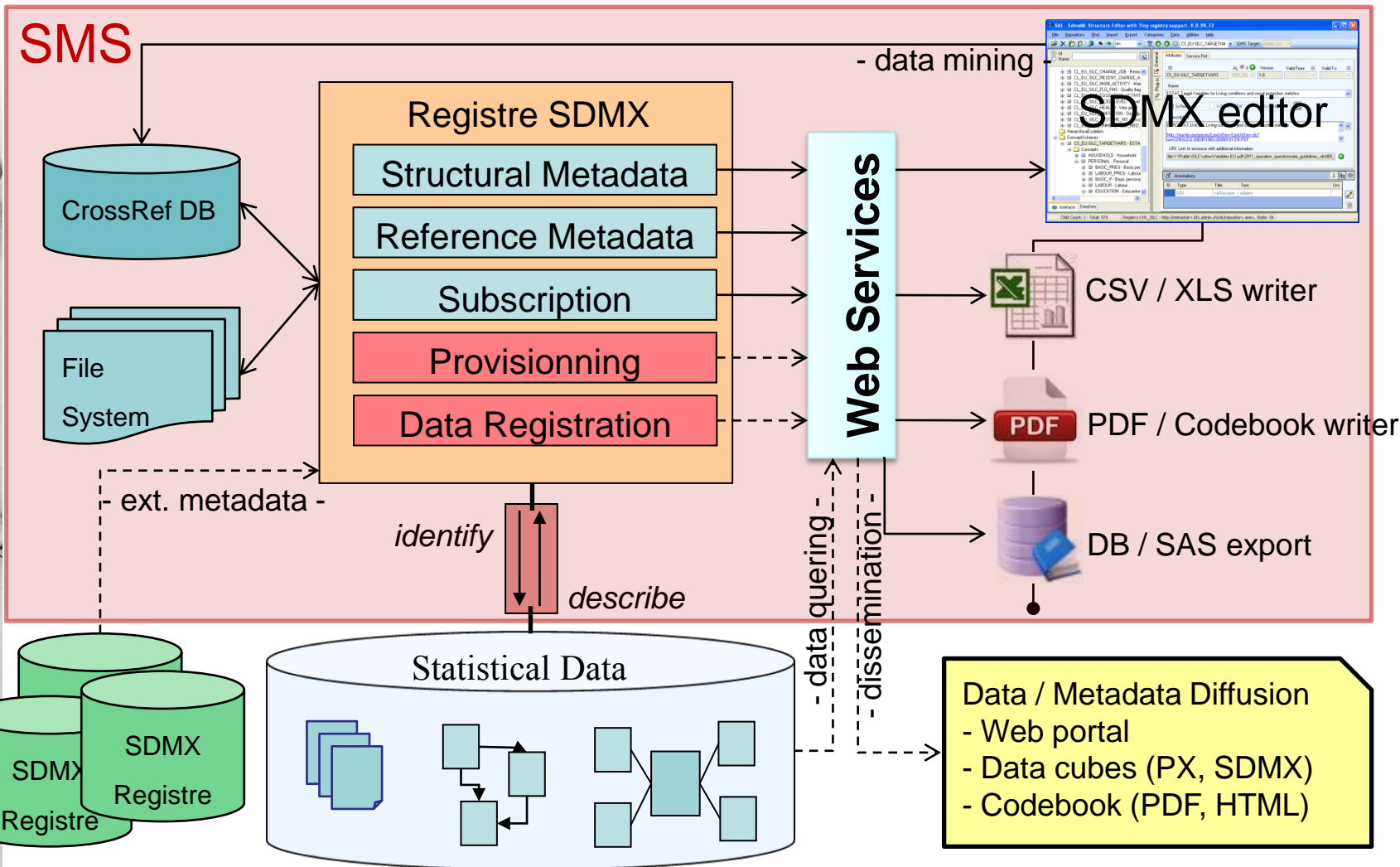
Statistique suisse





# SMS : Architecture and Coverage

Statistique suisse





## Statistical Metadata System : Functionalities

- Single Entry Point via Web Services
- Statistic life-cycle metadata management :
  - Management of statistical activities (national and European)
  - Creation / management of catalogues of variables per theme
  - Classifications and code lists with correspondence tables
  - Data description and registration / discovery
  - Metadata reports (quality, questionnaires, codebook)
- Harmonization of the structure and the metadata content
- Extensions: SAS plug-in; Questionnaire CATI; **Codebook**; etc.
- Metadata driven data access / monitoring
- Synergy, reutilization, redundancy reduction, integration



## Statistical Metadata System in Numbers

SDMX Artefact Type	Count
CategoryScheme	10
Codelist	1849
ConceptScheme	66
DataFlow	112
DataStructure	134
HierarchicalCodelist	116
MetadataFlow	48
MetadataStructure	26
MetadataSet	163

Registered users: 148



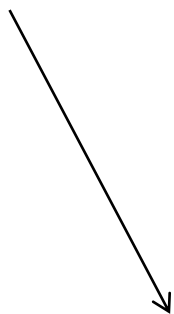
## Statistical Metadata System : Codebook

SDMX —————> DDI -----> DDI Codebook PDF / HTML

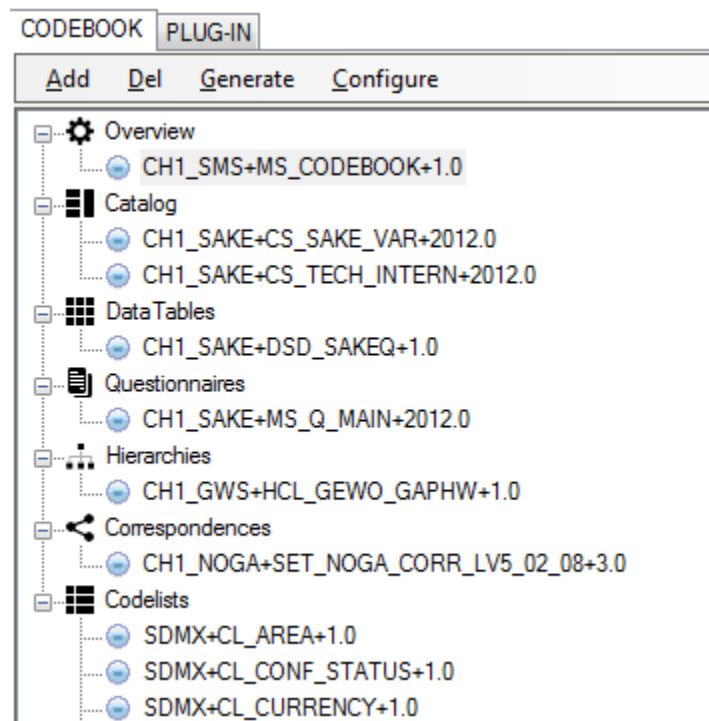


## Statistical Metadata System : Codebook

SDMX ~~→~~ DDI ..... DDI Codebook PDF / HTML



Generic Codebook





# Statistical Metadata System : Codebook

Statistique suisse

The screenshot shows the SAE - SdmxML Structure Editor with the following components:

- Menu Bar:** File, Repository, Find, Import, Export, Categories, Data, Utilities, Help
- Toolbar:** Includes icons for file operations and a dropdown menu set to 'fr'. The current project is 'CS\_SAKE\_VAR - - 2012' and the 'SDMX Target' is 'SDMX 2.0'.
- Left Panel (Tree View):** A hierarchical tree structure showing metadata categories such as 'OrganisationSchemes', 'Dataflows', 'ConceptSchemes', and 'Data Structure Definition'. The 'CS\_SAKE\_VAR - Vue d'ensemble' folder is selected.
- Central Panel (CODEBOOK PLUG-IN):** A 'General' tab with a 'Generate' button and a list of options: Overview, Catalog, Data Tables, Questionnaires, Hierarchies, Correspondences, and Codelists.
- Main Content Area:** Displays the Swiss flag and the text: 'Schweizerische Eidgenossenschaft', 'Confédération suisse', 'Confederazione Svizzera', 'Confederaziun svizra', 'Département fédéral de l'intérieur DFI', 'Office fédéral de la statistique OFS', and 'Codebook generator'.
- Status Bar:** Shows 'Child Count: 1 - Total: 1097' and 'Registry: CH1\_SAKE - http://metastat-r.bfs.admin.ch/sake/repository.asmx, State: Ok'.

SMS - Statistical Metadata System





# SMS : Codebook - Overview

The screenshot shows the SAE - SdmxML Structure Editor interface. The left pane displays a tree view of the codebook structure, including folders like OrganisationSchemes, Dataflows, Metadataflows, and various codebooks. The main pane shows the 'Overview' section for the 'Enquête suisse sur la population active (ESPA)' codebook. The overview includes a table of general information and a section for coverage.

Général	
<b>Titre</b>	Enquête suisse sur la population active
<b>Abbréviation</b>	ESPA
<b>No d'identification</b>	SAKE-2010-062012
	Office fédérale de la statistique, OFS
<b>Affiliation</b>	Département fédéral de l'intérieur, DFI
<b>Logiciel utilisé</b>	SMS, SDMX editor
<b>Version</b>	0.99
<b>Date</b>	2012-09-18
	OFS / RS / IT
<b>Affiliation</b>	Office fédéral de la statistique, OFS
<b>E-mail</b>	sms@bfs.admin.ch
<b>Type d'étude</b>	Enquête sur la force de travail
<b>Version</b>	ESPA 2010, fichier annuel.
<b>Date</b>	2011-10-18

Couverture	
<b>Information sur le sujet</b>	
<b>Mots clé</b>	Travail, Chômage, Sécurité sociale, Famille, Migration, Formation, Revenus, Déplacements, Culture et loisirs, Formation professionnelle
<b>Vocabulaire</b>	Vocabulaire ESPA
<b>Résumé</b>	L'Enquête Suisse sur la Population Active (ESPA) offre de précieux indicateurs sur le comportement en matière d'activité professionnelle de la population

Child Count: 1 - Total: 1097 Registry: CH1\_SAKE - http://metastat-r.bfs.admin.ch/sake/repository.asmx, State: Ok



# SMS : Codebook - Variables

The screenshot shows the SAE - SdmxML Structure Editor with the following details:

- Title Bar:** SAE - SdmxML Structure Editor with Tiny registry support. V.0.99.39
- Menu Bar:** File, Repository, Find, Import, Export, Categories, Data, Utilities, Help
- Toolbar:** Includes icons for file operations, navigation, and search. Shows 'fr' as the selected language and 'SDMX Target: SDMX 2.0'.
- Left Panel (Tree View):**
  - OrganisationSchemes
  - Dataflows
  - Metadataflows
  - CategorySchemes
  - Categorisations
  - CodeLists
  - HierarchicalCodelists
  - ConceptSchemes
    - CS\_CODEBOOK - DDI Co
    - CS\_SAKE\_VAR - Vue d'ensemble des variables** (selected)
    - Concepts
    - CS\_SMS\_OBS -
  - MetadataStructureDefinitions
    - MSD\_CODEBOOK - - 2.1
    - Metadata Targets
    - Reports
  - Data Structure Definition
    - DSD\_SAKEQ - Données
      - Dimensions
      - Measures
      - Attributes
      - Groups
  - StructureSets
  - ReportingTaxonomies
  - Processes
  - DataSet
  - MetadataSet
- Central Panel (Codebook View):**
  - General** / CODEBOOK / PLUG-IN
  - Buttons: Add, Del, Generate
  - Navigation: 4 / 353, 75.6% zoom, Find
  - Overview**
    - CH1\_SAKE+MS\_S
    - Catalog**
      - CH1\_SAKE+CS\_S
- Right Panel (Description):**

### Vue d'ensemble des variables ESPA

**Description**

L'enquête suisse sur la population active (ESPA) est une enquête auprès des ménages qui est réalisée chaque année depuis 1991. Le but principal de l'ESPA est de fournir des données sur la structure de la population active et sur les comportements en matière d'activité professionnelle. Grâce à l'application stricte de définitions internationales, les données de la Suisse peuvent être comparées avec celles des pays de l'OCDE et de l'Union européenne. Dès 2010, l'ESPA se déroulera en continu. Les interviews des 3ème et 4ème trimestres 2009 permettent le passage entre l'ancienne et la nouvelle ESPA.

**Information supplémentaire / Annotations**

No	Type : Title	Text
1	QUEST :	Q_2012_

No	ID: Name	Type
1	A: Informations concernant l'interview	Integer
2	B: Structure démographique du ménage	Integer
3	D: Questions filtres concernant le statut sur le marché du travail	Integer
4	E: Mobilité sur le marché du travail	Integer
5	F: Profession	Integer
6	G: Situation professionnelle	Integer
7	H: Type de contrat de travail	Integer
8	I: Lieu de travail	Integer
9	J: Migration	Integer
10	K: Volume de travail	Integer
11	L: Conditions de travail	Integer
12	M: Informations concernant l'établissement	Integer
13	N: Quête d'emploi	Integer
14	O: Caractéristiques du chômage	Integer
15	P: Propre situation auto-déclarée	Integer
16	Q: Formations achevées et formation actuelle	Integer
17	R: Formation continue	Integer
18	S: Variables socio-démographiques de la personne-cible	Integer

- Status Bar:** Child Count: 1 - Total: 1097 | Registry: CH1\_SAKE - http://metastat-r.bfs.admin.ch/sake/repository.asmx, State: Ok



# SMS : Codebook - Codelist

SAE - SdmxML Structure Editor with Tiny registry support. V.0.99.39

File Repository Find Import Export Categories Data Utilities Help

en CS\_SAKE\_VAR - - 2012 | SDMX Target: SDMX 2.0

CODEBOOK PLUG-IN

Add Del Generate

Overview  
Catalog  
Codelists  
SDMX+CL\_FREQ+1.0

**CL\_FREQ - Code list for Frequency (FREQ)**

**Description** It provides a list of values indicating the "frequency" of the data (e.g. monthly) and, thus, indirectly, also implying the type of "time reference" that could be used for identifying the data with respect time.

Code	Parent	Name
A		Annual
B		Daily - business week
D		Daily
M		Monthly
N		Minutely
Q		Quarterly
S		Half Yearly, semester
W		Weekly

**Codes détaillés**

**Code** A  
**Name** Annual  
**Description** It is typically used for annual data. This can also serve cases of multi-annual data (data that appear once every two, three or, possibly, five years). Descriptive information on the multiannual characteristics (e.g. frequency of the series in practice and other methodological information can be provided at the dataflow level, as long as these characteristics are applicable for the entire dataflow).

**Code** B  
**Name** Daily - business week  
**Description** Similar to "daily", however there are no observations for Saturday and Sunday (so, neither "missing values" nor "numeric values" should be provided for Saturday and Sunday). This treatment ("business") is one way to deal with such cases, but it is not the only option. Such a time series could alternatively be considered daily ("D"), thus, with missing values in the weekend.

Child Count: 1 - Total: 1097 Registry: CH1\_SAKE - http://metastat-r.bfs.admin.ch/sake/repository.asmx, State: Ok



## Statistical Metadata System : Next Steps

- Finalize the described SMS architecture
  - The new SDMX 2.1 compatible registry
  - The new Web services for registering and querying data
  - The Codebook extension
- Build an integrated metadata thesaurus



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Département fédéral de l'intérieur DFI  
Office fédéral de la statistique OFS

## **For all inquiries concerning the Statistical Metadata System @ FSO**

**Please contact :**

**[sms@bfs.admin.ch](mailto:sms@bfs.admin.ch)**

Statistique suisse