

SDMX Model d'information

Dany Ghafari, SDG and Environment Statistics Unit, UNEP

Chiffres vs données

Number of Touristic establishments in Italy, Annual data			
Indicator Time	A100 Hotels and similar	B010 Tourist Campsites	B020 Holiday dwellings
2002A00	33411	2374	61479
2003A00	33480	2530	58526
2004A00	33518	2529	56586
2005A00	33527	2411	68385
2006A00	33768	2510	68376
2007A00	34058	2587	61810

- Les chiffres en eux-mêmes n'ont aucun sens.
- Pour que les données soient utilisables, elles doivent être correctement décrites. Les descriptions permettent aux utilisateurs de savoir ce que les données représentent réellement.

Développement d'un modèle de données pour l'échange de données

- Le modèle de données est développé pour fournir des descriptions de toutes les caractéristiques pertinentes aux données à échanger
- Dans certains aspects, le développement d'un modèle de données est similaires au développement d'une base de données relationnelle
- Dans SDMX, le modèle de données est représenté par une définition de structure de données (DSD).
- La "forme" de SDMX DSD est à peu près similaire au schéma en étoile.
- Pour concevoir un DSD, nous devons d'abord trouver des concepts qui identifient et décrivent nos données.

Concept

- "Unité de pensée créée par une combinaison unique de caractéristiques"*
- Chaque concept décrit quelque chose à propos des données.
- Les concepts doivent exprimer toutes les caractéristiques pertinentes des données.

* Source: [SDMX Glossary](#)

Identification des concepts

Country - Pays	1980	1985	1990	1995	1999	2000	2001	2002	2003
Angola.....	6993	8754	9194	11072	12692	13134	13533	13942	14366
Botswana.....	906	1083	1276	1487	1529	1541	1549	1552	1565
Lesotho.....	1339	1538	1792	2050	2037	2035	2050	2065	2080
Malawi.....	6183	7340	9667	11129	11270	11308	11554	11806	12064
Mauritius - Maurice.....	966	1020	1057	1117	1151	1161	1169	1178	1187
Mozambique.....	12095	13711	14187	16004	17808	18292	18616	18946	19283
Namibia - Namibie.....	1030	1518	1349	1540	1711	1757	1787	1817	1848
South Africa.....	29170	33043	37066	41465	42902	43309	43634	43966	44306
Swaziland.....	560	664	744	855	910	925	933	942	950
Zambia - Zambie.....	5738	7006	8152	9456	10218	10421	10639	10683	11092
Zimbabwe.....	7126	8292	9903	11261	12333	12627	12843	13065	13292
Southern Africa, Total - Afrique de australe, totale.....	72106	83969	94387	107436	114561	116510	118305	119962	122033

Ref. Zone:
Pays

Indicateur

Période

Multiplicateur d'unité

Valeur d'obs.

Schéma conceptuel de SDMX

- "Ensemble de concepts utilisés dans une définition de structure de données ou une définition de structure de métadonnées."*
- Le schéma conceptuel place les concepts dans une unité maintenable.

Concept name	Concept ID
Indicateur	INDICATOR
Zone de référence	REF_AREA
Période	TIME_PERIOD
Multiplicateur d'unité	UNIT_MULT
Valeur d'observation	OBS_VALUE

* Source: [SDMX Glossary](#)

Dimension

- Lesquels des concepts sont utilisés pour identifier une observation ?
 - Indicateur
 - Zone de référence
 - Période de temps
- Lorsque les 3 sont connus, on peut situer sans ambiguïté une observation dans le tableau.
- C'est ce qu'on appelle les **dimensions**.
 - Une dimension a une signification similaire au champ de clé primaire d'une table de base de données.

Attribut

- Dans notre exemple, le **multiplicateur d'unité** représente des informations supplémentaires sur les observations.
- Ce concept n'est pas utilisé pour identifier une série ou une observation.
- Ces concepts sont appelés **attributs**.
 - A ne pas confondre avec les attributs XML !
 - Similaire aux champs de clé non primaire d'une table de base de données.

Mesure Primaire

- La valeur d'observation représente un concept qui décrit les valeurs réelles transmises.
- Dans SDMX, un tel concept est appelé **mesure primaire**.
- La mesure principale est généralement représentée par un concept avec l'**OBS_VALUE** comme identification.

Dimension ou Attribut ?

- Le choix du rôle d'un concept a de profondes implications sur la structure des données.
- Les concepts qui identifient les données doivent être dimensionnés. Les concepts qui fournissent des informations supplémentaires sur les données doivent devenir des attributs.
- Si un concept est une dimension, il est possible d'avoir des séries temporelles qui ne diffèrent que par la valeur de ce concept.
 - Par exemple. si l'unité de mesure est une dimension, il est possible d'avoir des séries chronologiques distinctes pour « T » et « T/HA » ou, plus controversé, « KG » et « T »

Dimension ou Attribut ? (2)

Cambodia		
Fixed and Mobile telephone subscriptions	2013	20.6 million
Fixed and Mobile telephone subscriptions	2012	19.7 million
Fixed and Mobile telephone subscriptions	2013	140.9 per 100 pop.

Unité de mesure comme dimension... (dimensions soulignées)

<u>Zone de ref.</u>	<u>Indicateur</u>	<u>Période</u>	<u>Unité de mesure</u>	<u>Multiplicateur d'unité</u>	<u>Valeur d'obs.</u>
Cambodia	Fixed and Mobile telephone subscriptions	2013	Number	Millions	20.6
Cambodia	Fixed and Mobile telephone subscriptions	2012	Number	Millions	19.7
Cambodia	Fixed and Mobile telephone subscriptions	2013	Per 100 pop.	Units	140.9

Dimension ou Attribut ? (3)

Unité de mesure comme attribut...

Violation!

<u>Zone de ref.</u>	<u>Indicateur</u>	<u>Periode</u>	Unité de mesure	Multipliateur d'unité	Valeur d'obs.
Cambodia	Fixed and Mobile telephone subscriptions	2013	Number	Millions	20.6
Cambodia	Fixed and Mobile telephone subscriptions	2012	Number	Millions	19.7
Cambodia	Fixed and Mobile telephone subscriptions	2013	Per 100 pop.	Units	140.9

- L'ensemble de données ci-dessus n'est pas valide : observation en double
- Les deux valeurs ci-dessus ne diffèrent que par leurs attributs

Dimension ou Attribut ? (4)

Unité de mesure en tant que dimension...



<u>Zone de ref.</u>	<u>Indicateur</u>	<u>Période de temps</u>	<u>Unité de mesure</u>	Multiplieur d'unité	Valeur d'obs.
Cambodia	Fixed and Mobile telephone subscriptions	2013	Number	Millions	20.6
Cambodia	Fixed and Mobile telephone subscriptions	2012	Number	Millions	19.7
Cambodia	Fixed and Mobile telephone subscriptions per 100 population	2013	Per 100 pop.	Units	140.9

- Maintenant, il n'y a pas de violation car chaque ligne a une clé unique
- Le concept d'unité est toujours utile

Attachement d'attribut

- Dans SDMX 2.0, les attributs peuvent être attachés au niveau de l'observation, de la série chronologique, du groupe ou de l'ensemble de données.
- Dans SDMX 2.1, les attributs peuvent être attachés à l'observation, à la ou aux dimensions, au groupe ou à l'ensemble de données.
 - Lorsqu'un attribut est attaché à toutes les dimensions sauf le temps, il est effectivement attaché aux séries temporelles
- Pour des raisons pratiques, les attributs sont souvent attachés à l'observation ou à la série chronologique.
- De plus, les attributs peuvent être désignés comme **obligatoires** ou **conditionnels** (facultatifs). Les attributs obligatoires doivent être présents au niveau de leur pièce jointe pour que l'ensemble de données soit valide, tandis que les attributs conditionnels peuvent être ignorés.
 - Les dimensions, en revanche, doivent toujours être fournies.

Concepts inter-domaines

- Le groupe de travail statistique SDMX (SWG) développe et publie des concepts inter-domaines
- Il s'agit d'identifiants de concept recommandés qui sont partagés entre les domaines thématiques statistiques et qui peuvent être réutilisés dans de nombreuses DSD. La liste complète des concepts inter-domaines est disponible sur le site Web SDMX sous Directives : https://sdmx.org/?page_id=3215
- Le schéma de concept inter-domaines est également publié sur le SDMX Global Registry : <https://registry.sdmx.org>

Concepts inter-domaines : exemples

- Certains des concepts inter-domaines largement utilisés incluent :
 - Indicateur statistique : **INDICATEUR**
 - Zone de référence : **REF_AREA**
 - Sexe : **SEXE**
 - Âge : **ÂGE**
 - Unité de mesure : **UNIT_MEASURE**
 - Multiplicateur d'unité : **UNIT_MULT**
 - Période : **TIME_PERIOD**
 - Valeur d'observation : **OBS_VALUE**

Modèle de données jusqu'à présent...

Concept	ID	Role	Attachement
Indicateur	INDICATOR	Dimension	
Zone de reference	REF_AREA	Dimension	
Période	TIME_PERIOD	Dimension	
Multiplicateur d'unité	UNIT_MULT	Attribute	séries chronologiques
Valeur d'observation	OBS_VALUE	Mesure Primaire	

Exercice 1 : Identifier les concepts

- Identifier les concepts dans le tableau
- Marquez chaque concept comme:
 - Dimension
 - Mesure primaire (c'est-à-dire valeur d'observation)
 - Attribut

Representation

- DSD définit une plage de valeurs valides pour chaque concept.
- Lorsque les données sont transférées, chacun de ses concepts descripteurs doit avoir des valeurs valides.
- Un concept peut être
 - codé
 - Non codé avec format
 - Texte libre non codé

Code

- "Un ensemble de lettres, de chiffres ou de symboles indépendant de la langue qui représente un concept dont la signification est décrite dans une langue naturelle."
- Une séquence de caractères qui peut être associée à des descriptions dans n'importe quel nombre de langues.
 - Les descriptions peuvent être mises à jour sans perturber les mappages ou d'autres composants de l'échange de données.

Liste de codes

- "Une liste prédéfinie à partir de laquelle certains concepts codés statistiques prennent leurs valeurs"
- Une liste de codes est un ensemble de codes conservés comme une unité.
- Une liste de codes énumère toutes les valeurs possibles pour un concept ou un ensemble de concepts
 - Liste des codes sexuels
 - Liste des codes de pays
 - Liste des codes indicateurs, etc.

Liste de codes: exemples

CL_SERIES	
Code	Description
SI_POV_DAY1	Population below international poverty line (1.1.1)
SI_POV_EMP1	Employed population below international poverty line (1.1.1)
SI_POV_NAHC	Population below national poverty line (1.2.1)
SI_COV_BENFTS	Population covered by at least one social protection floor/system (1.3.1)
SI_COV_CHLD	Children covered by social protection (1.3.1)
SI_COV_DISAB	Population with severe disabilities collecting disability social protection benefits (1.3.1)
SI_COV_LMKT	Population covered by labour market programs (1.3.1)
SI_COV_MATNL	Mothers receiving maternity benefits and benefits for newborns (1.3.1)
SI_COV_PENSN	Population above retirement age receiving a pension (1.3.1)

CL_AREA	
Code	Description
1	World
2	Africa (M49)
4	Afghanistan
5	South America (M49)
8	Albania
9	Oceania (M49)
10	Antarctica
11	Western Africa (M49)
12	Algeria

CL_EDUCATION_LEV		
Code	Description (EN)	Description (FR)
_T	Total or no breakdown by education level	Total ou aucune ventilation par niveau de s
ISCED11_0	Early childhood education	Education de la petite enfance
ISCED11_01	Early childhood educational development	Développement éducatif de la petite enfance
ISCED11_02	Pre-primary education	Enseignement préprimaire
ISCED11_1	Primary education	Enseignement primaire
ISCED11_10	Primary education	Enseignement primaire

Concepts SDMX et listes de codes

- Les listes de codes fournissent une représentation des concepts, en termes de codes.
- Les codes sont indépendants de la langue et peuvent inclure des descriptions dans plusieurs langues.
- Les listes de codes doivent être harmonisées entre tous les fournisseurs de données qui seront impliqués dans l'échange.

Concepts non codés

- Peut être en texte libre : tout texte valide peut être utilisé comme valeur pour le concept.
 - note de bas de page
- Peuvent avoir leur format spécifié
 - Code postal : 5 chiffres
 - Dernière mise à jour : date/heure

Représentation des concepts en SDMX

- **Les dimensions** doivent être codées ou avoir leur format spécifié.
 - Le texte libre n'est pas autorisé.
- **Les attributs** peuvent être codés ou non codés ; format peut éventuellement être spécifié.
- Le rôle et la représentation des concepts sont définis dans le DSD.
- Pour chaque concept d'un schéma conceptuel, une représentation par défaut peut ou non être fournie. Cette représentation par défaut peut ou non être remplacée dans le DSD.
- En d'autres termes, Concept Scheme suggère la représentation d'un concept, que le DSD peut accepter ou remplacer.

Modèle de données jusqu'à présent...

Concept	ID	Role	Attachement	Representation
Indicateur	INDICATOR	Dimension		CL_INDICATOR
Zone de reference	REF_AREA	Dimension		CL_AREA
Période	TIME_PERIOD	Dimension		Date/time (YYYY)
Multiplicateur d'unité	UNIT_MULT	Attribute	séries chronologiques	CL_UNIT_MULT
Valeur d'observation	OBS_VALUE	Mesure Primaire		Nombre à virgule flottante

Data model so far: Code Lists

CL_INDICATOR	
Code	Nom
POP	Total Mid-Year Population

CL_UNIT_MULT	
Code	Nom
0	Units
1	Tens
2	Hundreds
3	Thousands
6	Millions
9	Billions

CL_AREA (partielle)	
Code	Nom
AF	Afghanistan
AL	Albania
AQ	Antarctica
DZ	Algeria
AS	American Samoa
AD	Andorra

Listes de codes inter-domaines

- Comme avec les concepts inter-domaines, le groupe de travail statistique SDMX (SWG) développe et publie des listes de codes inter-domaines.
- Lorsqu'ils sont disponibles, ceux-ci sont basés sur des classifications statistiques existantes et contiennent les codes de ces classifications. Sinon, les codes sont développés par le SWG.
- Ces codes doivent être utilisés chaque fois que possible dans l'échange ou la diffusion SDMX. Les listes de codes sont souvent complétées par des codes spécifiques au pays ou à l'organisation. Par exemple, la liste mondiale des indicateurs régionaux de référence est souvent complétée par des indicateurs régionaux de référence infranationaux pour la diffusion des données nationales.
- Comme d'autres artefacts mondiaux, les listes de codes inter-domaines SDMX sont publiées sur le registre mondial SDMX : <https://registry.sdmx.org>

Listes de codes inter-domaines: exemples

- Zone de reference: **REF_AREA** → **CL_AREA**
- Sexe: **SEX** → **CL_SEX**
- Multiplicateur d'unité: **UNIT_MULT** → **CL_UNIT_MULT**

Codes génériques inter-domaines

Valeur de code recommandée	Description de code recommandé
_L	Extension locale (peut être utilisée comme préfixe)
_N	Non réponse
_O	Autre
_S	Sous-total
_T	Totale
_U	Pas de données/inconnu
_X	Non attribué/non spécifié
_Z	N'est pas applicable

- Il est recommandé d'utiliser ces codes dans toutes les listes de codes, le cas échéant.

Exercice 2 : Representation

- En travaillant avec votre modèle, déterminez la représentation de chaque concept
 - Codé, formaté, texte libre
- Développer des listes de codes et des formats pour vos concepts
 - Choisissez n'importe quelle approche pour vos codes et utilisez-la de manière cohérente

Définition de la structure de données : résumé



Concept	Concept ID	Role	Attachement	Representation	Code list ID
Indicateur	INDICATOR	Dimension		CL_INDICATOR	CL_INDICATOR
Zone de reference	REF_AREA	Dimension		CL_AREA	CL_AREA
Période	TIME_PERIOD	Dimension		Date/time (YYYY)	
Multiplicateur d'unité	UNIT_MULT	Attribute	séries chronologiques	CL_UNIT_MULT	CL_UNIT_MULT
Valeur d'observation	OBS_VALUE	Mesure Primaire		Nombre à virgule flottante	

Importance du modèle de données

- Le modèle de données, représenté par DSD, définit quelles données peuvent être encodées et transmises.
- Les défauts d'un DSD peuvent avoir un impact négatif important sur l'échange de données
 - Concepts manquantes
 - Rôle incorrect des concepts
 - Modèle non optimisé

Jeu de données

- Collecte organisée de données définies par une définition de structure de données (DSD)*
- Un jeu de données est structuré conformément à un DSD
- Sert comme conteneur pour les séries chronologiques ou les séries transversales dans les messages de données SDMX.

*Source: **SDMX Glossary**

Série chronologique

- Un ensemble d'observations d'une variable particulière, prises à différents moments dans le temps.
- Les observations qui appartiennent à la même série temporelle diffèrent dans leur dimension temporelle.
 - Toutes les autres valeurs de cote sont identiques.
 - Les attributs au niveau de l'observation peuvent différer entre les observations d'une même série chronologique.

Série chronologique: Demonstration

1.1 Proportion of population below \$1 (PPP) per day													
Series	1990	1992	1994	1996	1998	1999	2000	2002	2006	2007	2008	2009	2011
Rwanda													
MDG Population below \$1 (PPP) per day, percentage Last updated: 02 Jul 2012							74.6 ^{1,3}		72.1 ^{1,3}				63.2 ^{1,3}
State of Palestine													
MDG Population below \$1 (PPP) per day, percentage Last updated: 02 Jul 2012										0.4 ^{1,2,3}		0.0 ^{1,2,3}	
Thailand													
MDG Population below \$1 (PPP) per day, percentage Last updated: 02 Jul 2012	11.6 ^{1,3}	8.6 ^{1,3}	4.1 ^{1,3}	2.5 ^{1,3}	2.1 ^{1,3}	3.2 ^{1,3}	3.0 ^{1,3}	1.6 ^{1,3}	1.0 ^{1,3}		0.4 ^{1,3}	0.4 ^{1,3}	
1.2 Poverty gap ratio													
Series	1990	1992	1994	1996	1998	1999	2000	2002	2006	2007	2008	2009	2011
Rwanda													
MDG Poverty gap ratio at \$1 a day (PPP), percentage Last updated: 02 Jul 2012							36.9 ^{1,3}		34.8 ^{1,3}				26.6 ^{1,3}
State of Palestine													
MDG Poverty gap ratio at \$1 a day (PPP), percentage Last updated: 02 Jul 2012										0.1 ^{1,2,3}		0.0 ^{1,2,3}	
Thailand													
MDG Poverty gap ratio at \$1 a day (PPP), percentage Last updated: 02 Jul 2012	2.4 ^{1,3}	1.6 ^{1,3}	0.7 ^{1,3}	0.4 ^{1,3}	0.3 ^{1,3}	0.5 ^{1,3}	0.5 ^{1,3}	0.3 ^{1,3}	0.2 ^{1,3}		0.0 ^{1,3}	0.1 ^{1,3}	
Footnotes													
1 Based on nominal per capita consumption averages and distributions estimated from household survey data.													
2 Based on Purchasing Power Parity (PPP) dollars imputed using regression.													
3 Source: http://research.worldbank.org/PovcalNet/index.htm													

Données non chronologiques (données transversales¹)

- Une ou plusieurs dimensions non temporelles sont choisies le long desquelles un ensemble d'observations est construit.
 - Par exemple. pour une enquête ou un recensement, le temps est généralement fixe et une autre dimension peut être choisie pour être rapportée au niveau de l'observation
- Utilisé moins fréquemment que la représentation des séries chronologiques

¹ Le terme "données transversales" a été supprimé dans SDMX 2.1

Vue de série chronologique vs vue transversale

2.1 Net enrolment ratio in primary education

	2009	2010	2011
Morocco			
Total net enrolment ratio in primary education, both sexes		94.1	96.2
Total net enrolment ratio in primary education, boys		95	96.8
Total net enrolment ratio in primary education, girls		93.3	95.6
State of Palestine			
Total net enrolment ratio in primary education, both sexes	88.2	89.2	
Total net enrolment ratio in primary education, boys	88.2	89.8	
Total net enrolment ratio in primary education, girls	88.2	88.5	
Uganda			
Total net enrolment ratio in primary education, both sexes	94.2	91	
Total net enrolment ratio in primary education, boys	93.1	89.7	
Total net enrolment ratio in primary education, girls	95.3	92.3	

- La dimension Sexe a été choisie comme mesure transversale.

- Notez que l'heure est toujours applicable.

2.1 Net enrolment ratio in primary education 2010

	Total	Boys	Girls
Morocco	94.1	95	93.3
State of Palestine	89.2	89.8	88.5
Uganda	91	89.7	92.3

Clés dans SDMX

- La **clé de série** identifie de manière unique une série
 - Dans le cas des séries chronologiques, se compose de toutes les dimensions sauf le **temps**
- La clé de groupe identifie de manière unique un groupe de séries chronologiques
 - Se compose d'un sous-ensemble de la clé de série

Métadonnées structurelles et de référence

- Métadonnées structurelles : identificateurs et descripteurs, par ex.
 - Définition de la structure de données
 - Schéma conceptuel
 - Code
- Métadonnées de référence : Décrit le contenu et la qualité des données, par ex.
 - Définition de l'indicateur
 - Commentaires et limites



What we
have
covered
so far

Métadonnées de référence dans SDMX

- Peut être stocké ou échangé séparément de l'objet qu'il décrit, mais être lié à celui-ci
- Peut être indexé et recherché
- Rapporté selon une structure définie

Définition de la structure des métadonnées (MSD)

- MSD définit :
 - Le type d'objet auquel les métadonnées de référence peuvent être associées
 - Par exemple. DSD, flux de données.
 - Les composants composant l'identifiant d'objet de l'objet cible
 - Par exemple. le projet de SDG MSD permet d'attacher aux métadonnées une clé partielle.
 - Concepts utilisés pour exprimer les métadonnées (« attributs des métadonnées »).
 - Par exemple. Définition des indicateurs, gestion de la qualité

Métadonnées de référence dans SDMX 3.0

- La prise en charge des métadonnées de référence a été repensée dans SDMX 3.0.
 - Pour simplifier la mise en œuvre, la découverte et l'utilisation
- Il est possible de déclarer des concepts de métadonnées de référence directement dans le DSD et de transmettre des métadonnées de référence dans le cadre d'un ensemble de données ou dans un message séparé.
- Les définitions de structure de métadonnées, les ensembles de métadonnées et les flux de métadonnées seront toujours pris en charge pour les métadonnées de niveau supérieur qui ne sont pas attachées à des ensembles de données spécifiques.
- On s'attend à ce que la simplification des métadonnées de référence conduise à leur mise en œuvre dans des outils communs ainsi qu'à une utilisation beaucoup plus large.

Flux de données et flux de métadonnées

- Le flux de données peut être vu comme une "vue" sur une définition de structure de données
 - Peut être limité à un sous-ensemble de codes dans n'importe quelle dimension
 - Peut être catégorisé, c'est-à-dire peut avoir des catégories attachées
 - Dans sa forme la plus simple, définit toute donnée valide selon un DSD
- De même, le métadonnées flux définit une vue sur une définition de structure de métadonnées.
- Alors que le DSD définit la structure d'un jeu de données, un flux de données est un canal de transmission qui peut être utilisé pour le reporting ou la diffusion.
- Lorsqu'un ensemble de données est interrogé via une API, un flux de données est spécifié. Le flux de données peut couvrir l'intégralité de l'ensemble de données ou seulement certaines parties de l'ensemble de données.

Contraintes de contenu

- Les contraintes peuvent être utilisées pour définir quels codes ou combinaisons de codes sont autorisés (ou interdits)
- Les contraintes peuvent définir des règles de validation plus granulaires qu'une simple validation de codes
- Sont souvent attachés au Flux de données mais peuvent également être attachés à DSD, Accord de mise à disposition, etc.

Types de contraintes de contenu

- Les **Contraintes de Cube de Région** définissent les codes valides (ou non valides) comme un sous-ensemble de ceux définis dans les listes de codes d'un DSD. Par exemple, pour le pays dans le flux de données Global , la seule valeur valide pour la dimension REPORTING_TYPE est N ("National").
- Les **contraintes de série** définissent des combinaisons valides (ou non valides) de codes définis dans les listes de codes d'un DSD.
 - SERIES=**SH_STA_STNT** (Proportion d'enfants présentant un retard de croissance modéré ou sévère)
 - AGE=**Y0T4** (Moins de cinq ans)
 - COMPOSITE_BREAKDOWN=**_T,MS_MIGRANT,MS_NOMIGRANT,MS_EUMIGRANT** ou **MS_NONEUMIGRANT**
 - PRODUIT=**_T**
 - ACTIVITE=**_T**

Catégorie, schéma de catégories et catégorisation

- La catégorie est un moyen de classer les données pour les rapports ou la diffusion
 - Les domaines thématiques sont généralement mis en œuvre sous forme de catégories, telles que « Statistiques démographiques », « Statistiques économiques »
- Le schéma de catégorie regroupe les catégories dans une unité maintenable.
- La catégorisation lie une catégorie à l'objet auquel elle s'applique.

Messages SDMX

- Toute information relative à SDMX est échangée sous la forme de documents appelés messages. Un message SDMX peut être envoyé dans un certain nombre de formats standard, notamment XML, JSON, CSV.
- Il existe plusieurs types de messages SDMX, chacun ayant un but particulier, par ex.
 - Le message de **structure** est utilisé pour transmettre des informations structurelles telles que DSD, MSD, Concept Scheme, etc.
 - **GenericData, StructureSpecificData** et d'autres messages sont utilisés pour envoyer des données.
- Les messages SDMX au format XML sont appelés messages SDMX-ML.

SDMX Artefact

- Bien qu'il existe de nombreux types d'artefacts, dans presque toutes les situations, "SDMX Artefact" fait référence à un composant maintenable et versionnable de métadonnées structurelles.
 - Schéma conceptuel
 - DSD
 - Liste de codes
- Notez que par ex. un code individuel n'est généralement pas considéré comme un artefact, car il ne peut exister et être transmis que dans le cadre d'une liste de codes.

Identification et gestion des versions d'artefacts SDMX

- L'identification de SDMX Artefact se compose de 3 champs : ID ; Agence d'entretien ; Version

Type	Nom	ID	Agence d'entretien	Version
DSD	NA Main Aggregates	NA_MAIN	ESTAT	1.11
DSD	NA Main Aggregates	NA_MAIN	ESTAT	1.10
Concept Scheme	SDG Concept Scheme	SDG_CONCEPTS	IAEG-SDGs	1.0
Code List	Reference area code list	CL_AREA	SDMX	2.0

Merci!

Dany Ghafari, SDG and Environment Statistics Unit, UNEP

United Nations Avenue, Gigiri
PO Box 30552 – 00100 GPO Nairobi, Kenya

www.unep.org