



SDMX und DDI

August 2010

Data Integration, Tabulation and Dissemination
Government | Commercial | Research



Was ist SDMX/DDI?

- Das “Statistical Data Meta eXchange” (SDMX) Format ist ein XML-basierter Standard für den Austausch von (aggregierten) statistischen Daten mitsamt ihren Metadaten
- Der Standard wird gefördert von einem Konsortium bestehend aus *UNO, IWF, OECD, der Weltbank, EZB und BIS*.
- Der SDMX-Standard macht den Statistik-Austausch zwischen Nutzern und Lieferanten sicherer und einfacher
- Der DDI-Standard beschreibt Individualdaten/Mikrodaten in gleicher Weise wie SDMX aggregierte Daten.
- Die Schlüssel-Eigenschaften von SDMX und DDI sind die Werkzeug-Unabhängigkeit und die Einbettung von Metadaten mitsamt Daten
- Die DDI- und SDMX-Standards sind veröffentlicht und sind nicht auf Statistik begrenzt
- Der Einsatz des SDMX/DDI-Standards bringt das eGovernment-Ideal näher

Wann kann SDMX/DDI sinnvoll eingesetzt werden?



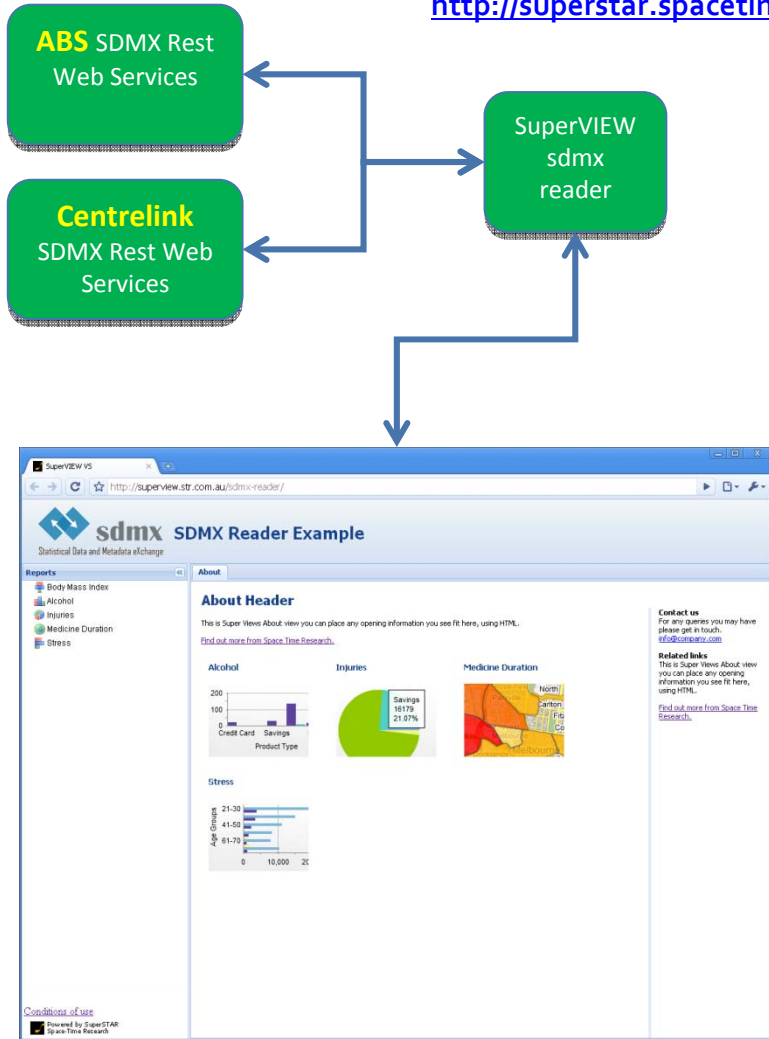
- Statistikämter und andere Ämter können mit Vorteil SDMX/DDI einsetzen, um (statistische) Daten sicherer und leichter
 - *einzusammeln,*
 - *an anderen Ämter weiterzugeben und*
 - *zu veröffentlichen,* sei es für
 - a) Firmen und Personen, die automatisiert (statistische) Daten über das Internet abfragen. Die Metadaten warnen zum Beispiel bei Anomalien.
 - b) Einzelpersonen, die mit interaktiven Internet-Abfragen Daten und Metadaten eruieren können. Je nach Werkzeug können die Ergebnisse in anderen Formaten heruntergeladen werden
- Alle neuen Projekte, die (statistische) Daten sammeln und/oder veröffentlichen, können mit Vorteil die SDMX/DDI-Standards einsetzen

- SuperSTAR:
 - Mikrodaten in DDI-Format einlesen und als .SXV4-Format speichern
 - .SXV4-Daten in SDMX-Format auslesen
 - .SXV4-Daten in SDMX-Format u.a. exportieren
 - SDMX-Daten mit SuperVIEW/SDMX ansehen
- Open Source:
 - Mikrodaten in DDI-Format umformattieren

Der SuperVIEW-"SDMX-Reader"



<http://superstar.spacetime-research.com/sdmx-reader/>



This section displays five detailed screenshots of the application's data visualizations:

- Stress:** A horizontal bar chart showing stress levels across different categories. A table below lists data for categories like "Low", "Medium", "High", and "Very High".
- Body Mass Index:** A horizontal bar chart showing BMI distribution by age group and sex. A table below provides summary statistics for "Male" and "Female".
- Alcohol:** A grouped bar chart showing alcohol consumption by state/territory. A table below lists consumption levels for "None", "Moderate", and "Excessive".
- Medicine Duration:** A map of Australia showing the duration of medicine use by region. A table below lists "Number of medicine" for various regions.
- Injuries:** Two pie charts showing the distribution of injuries by sex (Male and Female). A table below lists "Number of injuries" for different categories.

Mikrodaten aus DDI-definierter Quelle in SuperSTAR einlesen



```

61 <!-- LogicalProduct id="TM_DATASET_1169711_LP" -->
62 <!-- User ID type="Information Warehouse" dataset="1169711" -->
63 <!-- Label: Health Survey (National) - Year Ending June 2008 -->
64 <!-- Description: Health Survey (National) - Year Ending June 2008 -->
65
66 <!-- Data Relationship id="TM_DATASET_1169711_DR" -->
67 <!-- User ID type="Information Warehouse" dataset="1169711" -->
68 <!-- Label: Health Survey (National) - Year Ending June 2008 -->
69 <!-- Description: The data is stored/described in/by 2 different record structures that have a
70 defined hierarchical relationship
71 The levels are :
72 1 Person level
73 <!-- Description -->
74 <!-- LogicalRecord id="TM_DATASET_1169711_3_LR" hasLocation="false" variableQuantity="23" -->
75 <!-- User ID type="Information Warehouse" dataset="1169711" recordType="3" -->
76 <!-- Label: Person Level -->
77 <!-- Description: Person Level -->
78 <!-- Case Identification -->
79 <!-- Variable Specification Reference -->
80 <!-- Variable Reference -->
81 <!-- Scheme -->
82 <!-- ID: TM_DATASET_1169711_LR_KEYS_VS -->
83 <!-- Scheme -->
84 <!-- ID: CLS_1169711_3_LR_PERSON_ID -->
85 <!-- Variable Reference -->
86 <!-- Case Identification Reference -->
87 <!-- Case Identification -->
88 <!-- Variables In Record all Variables In LogicalProduct="false" -->
89 <!-- Variable Scheme Reference -->
90 <!-- All variables in this variable scheme are used in this record -->
91 <!-- ID: TM_DATASET_1169711_LR_3_VS -->
92 <!-- Variable Scheme Reference -->
    
```



person.csv - Microsoft Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	1	0									9	2304	0	7
2	2	2									9	1302	5	7
3	3	0									9	1101	3	7
4	4	0					6					1101	1	7
5	5	2					6				9	9232	1	0
6	6	0									9	1101	0	7
7	7	0	1	2								7103	5	5
8	8	0	1				6	7			9	1101	0	7
9	9	0										2201	5	7
10	10	0										9225	1	7
11	11	0							7		9	1101	5	7
12	12	0				4	5					1101	0	7
13	13	0						5	6			1101	4	7
14	14	0										1101	5	7
15	15	0				4		7	8			1101	5	0
16	16	0				4		7				2102	0	7
17	17	0										1101	5	7
18	18	0			3							1101	5	5
19	19	0						6	8			1101	1	7
20	20	2				2					9	5204	5	0
21	21	0										1101	3	7
22	22	0					4		6	7	9	1101	5	7
23	23	0	1				4				9	1101	0	2
24	24	0						4	6	7	9	1101	0	7



Untitled - SuperCHANNEL

File View Source Target Tools Help

New Open Save Source Attributes Logs Type Map Preferences Target Grouping Build All Update All Resume All Build Update Revise Resume Stop

Source View

Source Attributes

Foreign Key Column

Label: All types of action taken to manage diabetes/HSL

Name: CLS_ITEM_1628837_VAR_GROUP

Index: 2

Default:

Size: 256

Type: VARCHAR

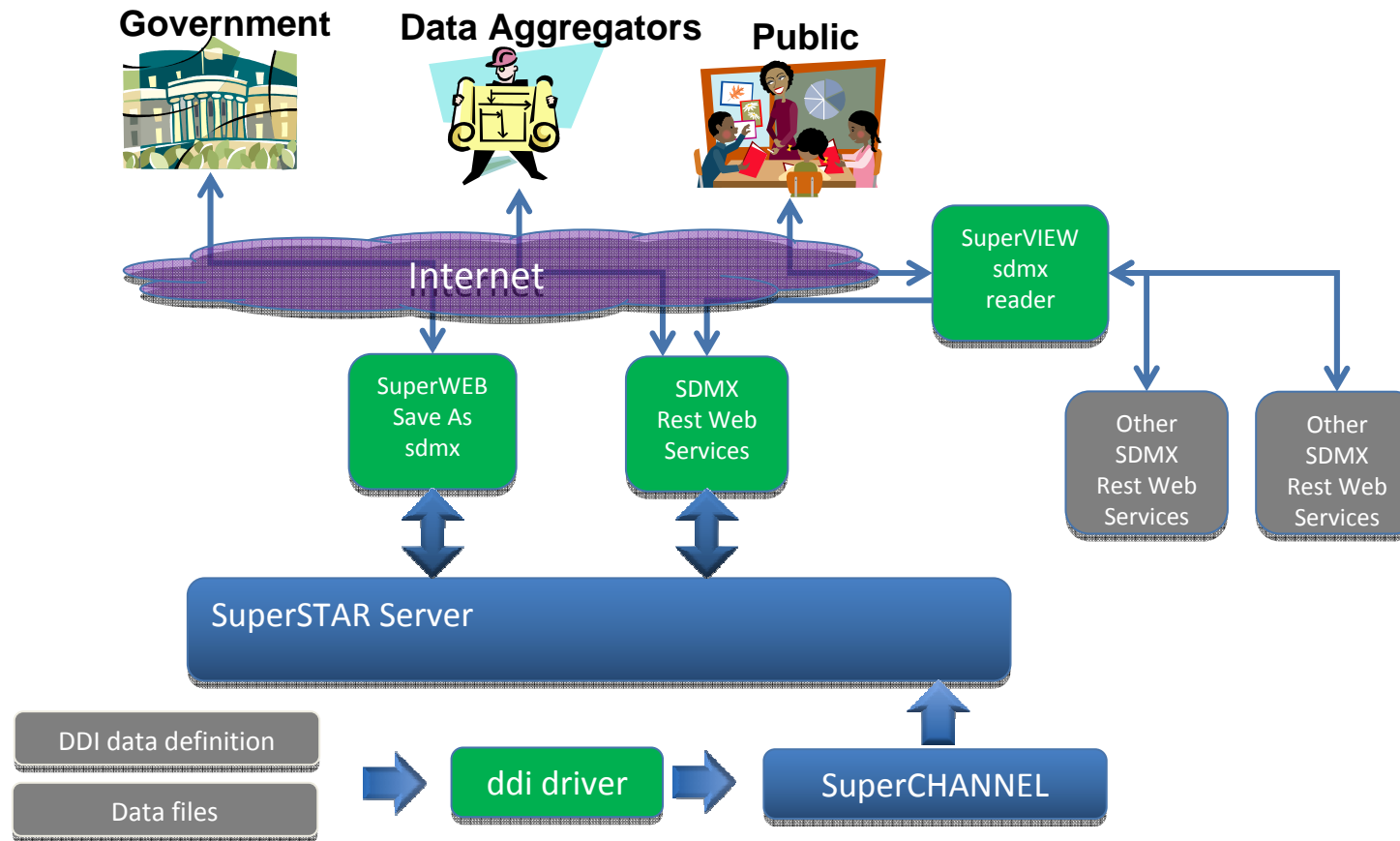
Usage: Classification

Target View

Test sample from NHS

- Fact Tables
 - All types of action taken to manage diabetes/HSL
 - Random person number
 - NHS_ACTDIABET_NHS_ACTDIABET_AS_CODED_CODS : code
 - Condition level (for use with Condition level items only)
 - Days in last week consumed alcohol
 - Person level
 - Time since last consulted an eye specialist/optometrist about any eye condition
 - Country of birth - classification suits SuperCross (3)
 - Registered Marital Status (3)
 - Type of educational institution currently attending
 - Highest year of school completed
 - Random person number
 - Type of other health professional consulted in the last 12 months
 - Random condition number
 - Type of other health professional consulted in the last 12 months
 - Classifications
 - Actions taken to manage diabetes/HSL
 - Frequency of visits to health professional
 - Highest year of school completed
 - Marital status
 - Standard Australian Classification of Countries Revision 2.03 (Operational Standard), lev
 - Standard Australian Classification of Countries Revision 2.03 (Operational Standard), lev
 - Standard Australian Classification of Countries Revision 2.03 (Operational Standard), lev
 - Standard Australian Classification of Countries Revision 2.03 (Operational Standard), lev
 - Time since last consulted an eye specialist/optometrist about any eye condition
 - Type of educational institution currently attending
 - Type of other health professions consulted
 - Which days in the last week alcohol was consumed on
 - Plain Tables

Superstar-Architektur-Ergänzungen für SDMX/DDI



Die grün-gefärbten Teile zeigen die SDMX/DDI-Ergänzungen in der SuperSTAR suite.

SuperSTAR kann

- mit dem **“SDMX-Webserver”** die SDMX-Daten im Internet bereitstellen
- mit der **“save-as-sdmx”**-SuperWEB-Funktion SDMX-Dateien erstellen
- mit dem SuperVIEW **“SDMX-Reader”** SDMX-Daten aus Webservices lesen

- mit dem **SuperCHANNEL “DDI-driver”** die DDI-definierten Daten in die Datenbank laden

SDMX-Daten von Internet-Diensten automatisiert abholen

Query URL

```
Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://localhost:8080/sdmxservices/Data/GenderByMaritalStatus/USER-user1
http://localhost:8080/USER-user1
http://www.sdmx.org/docs/2_0/SDMXGenericData.xsd http://www.SDMX.org/resources/SDMXXML/schemas/v2_0/message http://www.sdmx.org/docs/2_0/SDMXMessage.xsd"
- <message:Header>
  <message:ID>people_user1_2010-07-28T150851</message:ID>
  <message:Test>>false</message:Test>
  <message:Prepared>2010-07-28T15:08:51.648+10:00</message:Prepared>
  - <message:Sender id="AGENCY_ID">
    <message:Name xml:lang="en">AGENCY_NAME</message:Name>
    </message:Sender>
  </message:Header>
- <message:DataSet>
  <KeyFamilyRef>GenderByMaritalStatus</KeyFamilyRef>
  - <Series>
    - <SeriesKey>
      <Value concept="Gender" value="M"/>
      <Value concept="Maritalstatus" value="SING"/>
    </SeriesKey>
    - <Obs>
      <Time>9999</Time>
      <ObsValue value="8.0"/>
    </Obs>
  </Series>
  - <Series>
    - <SeriesKey>
```

Ergebnis

Abfrage-Ergebnisse von SuperWEB als SDMX-Dateien



Download Table: Excel 2007 (.xlsx)(max 16,384 columns x 65,000 rows and < 100,000 cells) Go

- Excel 2007 (.xlsx)(max 16,384 columns x 65,000 rows and < 100,000 cells)
- Excel 2003 (.xls)(max 255 columns x 65,000 rows)
- Comma Separated Value (.csv)
- CSV String Value (.csv)
- SDMX Structure Definition (.xml)**
- SDMX Archive (.zip)

data

structure

```
<?xml version="1.0" encoding="UTF-8"?><message:GenericData xmlns="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/generic" xmlns:message="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/message" xmlns:common="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/common" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/generic http://www.sdmx.org/docs/2_0/SDMXGenericData.xsd http://www.SDMX.org/resources/SDMXML/schemas/v2_0/message http://www.sdmx.org/docs/2_0/SDMXMessage.xsd"><message:Header><message:ID>NHS1user12010-07-28T092253</message:ID><message:Test>false</message:Test><message:Prepared>2010-07-28T09:22:53.968+10:00</message:Prepared><message:Sender id="AGENCY_ID"><message:Name xml:lang="en">AGENCY_NAME</message:Sender></message:Header><message:DataSet><KeyFamilyRef>NHS1user12010-07-28T092253</KeyFamilyRef><Series><SeriesKey><Value concept="Household@20@level_1642944" value="NHS5f@BMBMICT@5f@1642944@20@value@20@set_01"></Value><Value concept="Household@20@level_1601858" value="LFS@5f@SEX@5f@1601858@20@value@20@set_1"></Value><Value concept="MEASURES_DIMENSION" value="Person@20@level_1178259"></Value></SeriesKey><Attributes><Value concept="UNIT_MEASURE" value="number"></Value><Value concept="UNIT_MULT" value="0"></Value><Value concept="AGGREGATION_LEVEL" value="Person@20@level"></Value><Value concept="MEASURE_TYPE" value="count"></Value><Value concept="TIME_FORMAT" value="yyyy-MM-dd"></Value></Attributes><Obs><Time>9999</Time><ObsValue value="24699.0"></ObsValue></message:DataSet></message:GenericData>
```

```
<?xml version="1.0" encoding="UTF-8"?><message:Structure xmlns="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/structure" xmlns:message="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/message" xmlns:common="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/common" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/message http://www.sdmx.org/docs/2_0/SDMXMessage.xsd http://www.SDMX.org/resources/SDMXML/schemas/v2_0/structure http://www.sdmx.org/docs/2_0/SDMXStructure.xsd"><message:Header><message:ID>NHS1user12010-07-28T092253</message:ID><message:Test>false</message:Test><message:Prepared>2010-07-28T09:22:53.968+10:00</message:Prepared><message:Sender id="AGENCY_ID"><message:Name xml:lang="en">AGENCY_NAME</message:Sender></message:Header><message:CodeLists><CodeList agencyID="AGENCY_ID" id="NHS5f@BMBMICT@5f@1642944@20@value@20@set"><Name xml:lang="en">NHS_BMBMICT_1642944 value set</Name><Code value="NHS5f@BMBMICT@5f@1642944@20@value@20@set_01"><Description xml:lang="en">Grade 3 thinness (BMI<16.00)</Description></Code><Code value="NHS5f@BMBMICT@5f@1642944@20@value@20@set_02"><Description xml:lang="en">Grade 2 thinness (BMI16.00-16.99)</Description></Code><Code value="NHS5f@BMBMICT@5f@1642944@20@value@20@set_03"><Description xml:lang="en">Grade 1 thinness (BMI 17-18.49)</Description></Code><Code value="NHS5f@BMBMICT@5f@1642944@20@value@20@set_04"><Description xml:lang="en">Normal range</Code></CodeList></message:Structure>
```