Interface File Specfication

MVI SOLVE-IT Einstellungstest

Autoren	**Markus Bayerlein**
	Thomas Linke

Version	**1.1**
Datum	**16.02.2015**
Status	**In Arbeit**

Table of Contents

1 Introduction

On a daily basis, data files reflecting new or modified data are being created on an external server. These files are downloaded in order to import them into our in-house database.

The document explains the interface, which files types exist, what their structure is, when they will be available and how you can download them. The reader is supposed to have IT engineering skills.

The data files described in this document are tabulator separated flat files (see "→File format").

2 Available File Types

The data files can be categorized into common and company specific files. The latter can only be downloaded by the company whose data the files contain. Users need an appropriate user profile to download any of these files.

There are files containing data that has been released or changed on a certain due date (daily files), files containing initial data (from the beginning of delivering System until a certain due date and files containing data of a certain time period (start date, end date). Some of them will be created automatically every day, some of them have to be requested individually.

(you can click on the data type to get to the file specification)		Daily	Initial	
ID	**Data Type**			
Common data files:				
Substances	"→ Basic Substances"	✓	✓	
Companies	"→ Company / Contact"	✓	✓	

✓ This file will automatically be generated every day.

3 File Specification

3.1 Basis-BNF (Backus Naur Form)

The metasyntax BNF is the basis for all further definitions.

Data Type	Description
DIGIT	<"0".."9">
NUMBER	*(DIGIT)
FLOAT	NUMBER "," NUMBER
Nx	NUMBER with x digits
Nx,y	FLOAT with (x minus y) digits in front of the decimal point and y digits behind (precision)
Ax	Character string with max x characters
B	Boolean, 0 (False) or 1 (True)
B(-1)	Boolean, 0 (False), 1 (True), -1 (Unspecified, for example if a yes/no question has not been answered)
DT	Date (TT.MM.JJJJ, e.g. 24.12.1999)
TI	Time (HH:MM:SS, e.g. 23:30:05)
RECSEP	Record separator: <US-ASCII LF, line feed (10)> (Java: "\n")
SEP	Field separator: <US-ASCII TAB, tab> (Java: "\t")

NA users please note that the decimal "point" in decimal values is represented by a comma due to the European development.

Numeric values in the download files are always displayed with the maximum of precision even if the value equals zero (e.g. "0,000000" for a field with format N12,6).

Some numeric fields may contain a "-1" or "-1,0000" as an unspecified value, meaning for example that an optional weight field has not been filled out by the creator.

Empty text fields contain no character, even no spaces.

3.2 File format

The data files are so-called "flat files", built of ANSI characters.

Each line contains several fields. The fields are separated by tabulators (SEP), the lines are separated by line feeds (RECSEP).

The first field per line is always an identifier that links to the line specification with the definition how many fields follow and what data type they are.

3.3 Syntax Description

This chapter describes the structure and the format of the data files.

There is one definition table for each possible record (line in the flat file). The field order in the table will be reflected in the flat files. All fields are separated by a tabulator (SEP). The last field in a record is always terminated by a line feed (RECSEP). Because of optional fields in some records, the number of fields may vary from company to company.

As the separators are always the same, they are not mentioned in the definition tables.

3.3.1 File Header

The file headers are always the first line in the data files. They may vary depending on the time frame the file has been generated for (daily, initial or time period).

In all cases, field number two (DataFileIdentifier) will be replaced by the identifier of the data file, for example "SUBSTANCE". In that case the placeholder "A?" would be replaced by A9, reflecting the 9 characters of the word "SUBSTANCE".

INITIALFILEHEADER		
No.	Name	Type
1	"#" (Header record identifier)	A1
2	DataFileIdentifier	A?
3	Specification version (currently 2,60)	N7,2
4	Date of file generation	DT
5	Time of file generation	TI
6	Start date: "01.01.2000" *	DT
7	End date	DT

* The start date "01.01.2000" was the Date when the delivering system went productive.

3.3.2 Basic Substances

This file with the ID "Substances" contains all basic substances created or changed on due date (daily file).

BNF Structure:

SUBSTANCESTRUCT	=	SUBSTANCEHEADER SUBSTANCEBODY
SUBSTANCEHEADER	=	(DAILYFILEHEADER | INITIALFILEHEADER) with DataFileIdentifier = "SUBSTANCE" (Type: A9), see chapter "→File Header"
SUBSTANCEBODY	=	*SUBSTANCEENTRY
SUBSTANCEENTRY	=	SUBSTANCE *SYNONYM

SUBSTANCE		
No.	Name	Type
1	"R" (Record Identifier)	A1
2	Node ID of the basic substance	N18
3	CAS Code (if available)	A20
4	EU Index Code (If Available)	A20
5	EINECS / ELINCS Code (If Available)	A20
6	GADSL-classification "duty-to-declare"	B
7	GADSL -classification "IsUnwanted"	B(-1)
8	GADSL -classification "IsProhibited"	B
9	IsReach	B
10	IsDeleted	B
11	IsHidden	B
Sort order: Node ID		

In 2001 the field "IsUnwanted" has been removed from the ILRS (former VDA) classifications. For that reason newer substances have an "IsUnwanted" value of "-1" (unspecified).

Either “IsDeleted” or “IsHidden” can possess value “1”, never both at the same time (but both can be “0” of course).

SYNONYM		
No.	Name	Type
1	"RN" (Record Identifier)	A2
2	Node ID of the basic substance	N18
3	Synonym ID (starting at 0)	N4
4	ISO Language	A2
5	Synonym (Name)	A250
Sort order: ISO Language, Synonym ID		

Each SUBSTANCE contains 1 mandatory substance name (SynonymId 0) and up to 3 optional substance synonyms (SynonymId 1, 2 and 3) per ISO Language. Currently the languages "EN" (English) and "DE" (German) are available.

3.3.3 Company / Contact

This file with the ID "Companies" contains all companies and contacts created or changed on due date (daily file).
It consists of complete COMPCONTACTENTRY blocks. That means if for example a contact person has changed on the due date, you will find a COMPCONTACTENTRY with the contact person's company, all its contact persons and its org-units in the data file.

The initial file contains all companies and contacts ever created from the beginning of the delivering system on until the date of file generation, independent of the end date in the header.

BNF Structure:

COMPCONTACTSTRUCT	=	COMPCONTACTHEADER COMPCONTACTBODY
COMPCONTACTHEADER	=	(DAILYFILEHEADER | INITIALFILEHEADER) with DataFileIdentifier = "COMPCONTACT" (Type: A11), see chapter "→File Header"
COMPCONTACTBODY	=	*COMPCONTACTENTRY
COMPCONTACTENTRY	=	COMPANY *CONTACT *COMPUNITS

COMPANY		
No.	Name	Type
1	"FI" (Record Identifier)	A2
2	Company ID	N9
3	Company Name	A50
4	Street	A50
5	Postbox	A20
6	ISO Country Code	A2
7	Postal Code	A10
8	City	A50
9	IsDeleted	B
10	IsOEM	B
11	Duns - Number	A11
Sort order: Company ID		

CONTACT		
No.	Name	Type
1	"FK" (Record Identifier)	A2
2	Company ID	N9
3	Contact ID in the company's contact list (starting at 1)	N4
4	IsDeleted	B
5	Contact Last name	A50
6	Contact First name	A50
7	Contact Phone number	A50
8	Contact Fax number	A50
9	Contact e-mail address	A80
10	Contact Department name	A50
11	Contact Mailbox	A20
12	IsIMDSContact	B
13	IsREACHContact	B
Sort order: Company ID, Contact ID		

COMPUNITS		
No.	Name	Type
1	"FE" (Record Identifier)	A2
2	Company ID (Root)	N9
3	Org-Unit ID	N9
4	Org-Unit Name	A50
5	Org-Unit Street	A50
6	Org-Unit Postbox	A20
7	Org-Unit ISO Country Code	A2
8	Org-Unit Postal Code	A10
9	Org-Unit City	A50
10	IsDeleted	B
11	Duns - Number	A11
Sort order: Company ID (Root), Org-Unit ID		