Paper for Consideration by the S-100WG

ISO/IEC8211 Summary

Submitted by:	SevenCs GmbH – Holger Bothien
Executive Summary:	Describes the use of the ISO8211 Data Descriptive Record
Related document(s):	S-100 Part 10a

Introduction / Background

In S-57 a summary of ISO 8211 exists which helps the OEMs to understand how the encoding works in general and gives some implementation details.

S-100 Part 10a does not have such a summary. Especially, no information can be found about the general record structure and the content of the Data Descriptive Record (DDR).

Analysis/Discussion

Though, so far there are no requirements from OEMs, but we think, that this kind of documentation would be improve the usability of the standard. And it won't make it worse.

One interesting question here is which fields must be used within the DDR. We have seen most data sets (S-57 ENCs) describing all possible fields for the standard or for the data product. This leads to unnecessary large DDRs and the problem is even bigger in S-100. There are more different fields available. In many cases not all the fields will be used, especially in update files the actual number of different fields might be quite small.

The DDR describes the structure of the following records of the file and not the structure of any possible datasets for a given product. That's why it is sufficient to limit the DDR to what is really used within the file. The software overhead to achieve this is manageable. Nevertheless, this should be a minimum requirement only and old data will not become invalid.

Conclusions

A clause with appropriate sub clauses should be added after S-100 10a-3.6 Data Descriptive Fields.

Note that a separate proposal will renumber the clauses, but the logical place will be maintained.

The structure will be

- ISO/IEC 8211 file structure
  - General structure
  - The Leader
  - The Directory
  - The Field Area
  - An Example

The detailed text is proposed as follows:

3.7 ISO/IEC 8211 File Structure

3.7.1 General Structure

This clause does not replace ISO/IEC 8211, nor does it give a comprehensive overview of ISO/IEC 8211. The reader is referred to ISO/IEC 8211:1994 for a complete description and explanation.

The clause will give a short overview of ISO/IEC 8211 by explaining those parts of the encapsulation structure which are of relevance to S-100.

ISO/IEC 8211 files are organized by Logical Records (LR), the first record is the Data Descriptive Record (DDR) and all subsequent records are Data Records (DR).

The DDR contains information on the hierarchy and structure of the remaining part of the file. It does not define the semantic of the data.

Each Logical Record (both DDR and DR) contains three basic elements

- Leader
- Directory
- Field Area

3.7.2 The Leader

The leader of a logical record contains the parameters necessary to read records and to disaggregate the directory into its entries. In addition, the DDR leader contains a few data descriptive parameters applicable to the entire file. It has a fixed length of 24 bytes.
The first five bytes in any leader will contain the length of the complete record in bytes encoded as a decimal number in ASCII representation. (e.g., a record of 242 Bytes will a record length entry of "00242". If a record has a size of 100000 bytes and larger than the value must be set to "00000". In this case the software must be able to calculate the record size from the information in the directory.

3.7.2.1 The DDR Leader

The structure of the DDR leader

RP	Len	Entry name	Content
0	5	Record length	number of bytes in record
5	1	Interchange level	"3"
6	1	Leader identifier	"L"
7	1	In line code extension indicator	"E"
8	1	Version number	"1"
9	1	Application indicator	SPACE
10	2	Field control length	"09"
12	5	Base address of field area	Start address of field area (number of bytes in leader and directory)
17	3	Extended character set indicator	" ! " (SPACE,!,SPACE)
20	4	Entry map	(see below)

Entry map of DDR leader

RP	Sub-entry name	Len	Content
20	Size of field length field	1	Variable "1"-"9" (defined by encoder)
21	Size of field position field	1	Variable "1"-"9" (defined by encoder)
22	Reserved	1	"0"
23	Size of field tag field	1	"4"

3.7.2.2 The DR Leader

The structure of the DR leader

RP	Len	Entry name	Content
0	5	Record length	number of bytes in record
5	1	Interchange level	SPACE
6	1	Leader identifier	"D"
7	1	In line code extension indicator	SPACE
8	1	Version number	SPACE
9	1	Application indicator	SPACE
10	2	Field control length	2 SPACEs
12	5	Base address of field area	Start address of field area (number of bytes in leader and directory)
17	3	Extended character set indicator	3 SPACEs
20	4	Entry map	(see below)

Entry map of DR leader

RP	Sub-entry name	Len	Content
20	Size of field length field	1	Variable "1"-"9" (defined by encoder)
21	Size of field position field	1	Variable "1"-"9" (defined by encoder)
22	Reserved	1	"0"
23	Size of field tag field	1	"4"

3.7.3 The Directory

The directory of a logical record contains the parameters necessary to identify and locate each field in the field area. The directory consists of repeated directory entries containing the
- field tag
- field length
- field position

The directory ends with the field terminator (1/14). The field positions are relative to the beginning of the field area. The position of the first field following the directory is 0. The number of bytes used for the three elements (the field entry) is defined by the entry map in the leader of the logical record.

3.7.4 The Field Area

The field area is different for the DDR and DR. In the first record only, the DDR, the field area contains data descriptive fields. Each data descriptive field contains information necessary to decode the user data in the field area of the DR('s). The fact that the data description is contained in the interchange file makes it possible to exchange data without an external description, though the semantic of the elements is not known. The S-100 standard and the product specifications that uses an ISO/IEC 8211 data encoding does contain an external data description used for the exchange of the data. However, the data descriptive fields can only be omitted from the DDR if they are not used in the current file not because of the existence of an external data description. The data descriptive fields of the DDR form an integral part of an ISO/IEC 8211 conforming file.

The limitation to the used fields is a minimum requirement and other fields may be defined by Data Descriptive Fields in the DDR. However, this adds unnecessary data to the data set and should be avoided.

The field area of the DR contains the actual data to be transferred

3.7.4.1 The Field Area of the DDR
a. Field control field

The first field of the DDR is the field control field. The field tag for the field control field is "0000". The field control field contains a list of field tag pairs. The list defines the hierarchy of all the fields described in the DDR. The list contains pairs of partent/child tags and together with the preorder traversal sequence of the field descriptions in the DDR describes a generic tree structure for the exchange file. The pairs may be placed in the list in any sequence and must be contiguous. The following figure gives an example of a tree. The set of field tag pairs is HE, EA, EB, HF, HG, GC and GD.
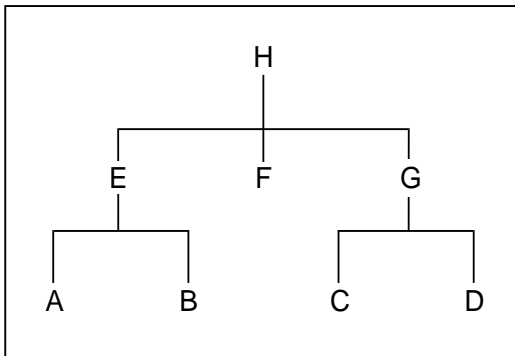
The structure of the field control field is as following

Field controls	External file title (Optional for S-100)	UT	List of field tag pairs	FT

The field controls of the field control field are : "0000;&" + 3 SPACEs.

b. Data descriptive fields

The successive fields of the field area contain the data descriptive fields. The data descriptive fields are encoded in the DDR in a preorder traversal sequence. The preorder traversal sequence of the tree shown above is HEABFGCD.

The structure of a data descriptive field is as following

Field controls	Field name	UT	Array descriptor	UT	Format controls	FT

The field controls describe the level and data type of the data fields defined by the data descriptive fields. The structure of the field controls is shown in the following table.

RP	Len	Entry name	Content
0	1	Data structure code	"1" – linear structure "2" – multi-dimensional structure "3" – concatenated structure
1	1	Data type code	"1" – implicit point (integer) "2" – implicit point (float) "6" – mixed data types

2	2	Auxiliary controls	"00"
4	2	Printable graphics	";&"
6	3	Truncated escape sequence	" " (3 SPACEs) – ASCII Encoding ISO 646 "%/G" – UTF8 Encoding implementation level 1

The field name contains the long description of the data fields as defined in the tree structures given in this part of S-100. The Array descriptor and format controls define the inner field structure for the associated data fields. Refer to ISO/IEC 8211 for a complete description.

3.7.4.2 Field area of the DR's

The data fields in the DR's must be encoded in the preorder traversal sequence as defined in the DDR. The structure of the data fields is defined by the data descriptive fields in the DDR.

3.7.5 An Example

The following rather simple example shows an S-100 conformal ISO/IEC 8211 conformal dataset file.

It contains a single feature type (*BuoySafeWater*) with the following attributes set:

Attribute Code	Value	Remark
buoyShape	"4"	Pillar
Colour[1]	"3"	Red
Colour[2]	"1"	White
colourPattern	"3"	Diagonal Stripes
featureName[1].language	"eng"	English
featureName[1].name	"Example buoy"	
featureName[2].language	"deu"	German
featureName[2].name	"Beispiel Tonne"	

The Feature Object Id is

Sub-Field	Integer value	Hexadecimal representation
Producing agency	31868	7C7C
Feature identification number	12345678	00BC614E
Feature identification subdivision	42	002A

The position is

	Geo position	Integer value	Hexadecimal representation
Latitude	42.42° N	424200000	1948C740
Longitude	12.1234° W	-121234000	F8C61DB0

The example contains non-printable characters and binary codes. They are replaced with the denotation as defined in the following table:

Character	Code	Denotation	Remark
Space	Hex 20	□	
UT (Unit Terminator)	Hex 1F	▲	
FT (Field Terminator)	Hex 1E	▼	

Binary code b11	Hex xx	[xx]	
Binary code b12	Hex xxyy	[xxyy]	Due to the LSBF encoding this is equivalent to [yy][xx]
Binary code b14	Hex wwxxyyzz	[wwxxyyzz]	Due to the LSBF encoding this is equivalent to [zz][yy][xx][ww]
Binary code b24	Hex wwxxyyzz	[wwxxyyzz]	For negative numbers the two complement is encoded LSBF
Binary code b48	x.y	[x.y]	Encoded as defined by the IEEE 754 double precision encoding LSBF

DDR Leader

011803LE1□0900155□!□3304

DDR Directory

0000090000DSID132090DSSI118222ATCS044340FTCS047384CSID084431CRSH095515PRID0
71610C2IT055681FRID084736FOID064820ATTR058884SPAS083942▼

DDR Field Area

0000;&□□□S100Example.000▲DSIDDSSIDSIDATCSDSIDFTCSCSIDCRSHPRIDC2ITFRIDFOIDFR
IDATTRFRIDSPAS▼

3600;&%/GData□Set□Identification▲RCNM!RCID!ENSP!ENED!PRSP!PRED!PROF!DSNM!DS
TL!DSRD!DSLG!DSAB!DSED*DSTC▲(b11,b14,7A,A(8),3A,{b11})▼

1600;&□□□Data□Set□Structure□Information▲DCOX!DCOY!DCOZ!CMFX!CMFY!CMFZ!NOIR!
NOPN!NOMN!NOCN!NOXN!NOSN!NOFR▲(3b48,10b14)▼

2600;&□□□Attribute□Codes▲*ATCD!ANCD▲(A,b12)▼

2600;&□□□Feature□Type□Codes▲*FTCD!FTNC▲(A,b12)▼

1100;&□□□Coordinate□Reference□System□Record□Identifier▲RCNM!RCID!NCRC▲(b11,
b14,b11)▼

1600;&%/GCoordinate□Reference□System□Header▲CRIX!CRST!CSTY!CRNM!CRSI!CRSS!S
CRI▲(3b11,2A,b11,A)▼

1100;&□□□Point□Record□Identifier▲RCNM!RCID!RVER!RUIN▲(b11,b14,b12,b11)▼

1100;&□□□2-D□Integer□Coordinate□Tuple▲YCOO!XCOO▲(2b24)▼

1100;&□□□Feature□Type□Record□Identifier▲RCNM!RCID!NFTC!RVER!RUIN▲(b11,b14,2
b12,b11)▼

1100;&□□□Feature□Object□Identifier▲AGEN!FIDN!FIDS▲(b12,b14,b12)▼

2600;&%/GAttribute▲*NATC!ATIX!PAIX!ATIN!ATVL▲(3b12,b11,A)▼

2100;&□□□Spatial□Association▲*RRNM!RRID!ORNT!SMIN!SMAX!SAUI▲(b11,b14,b11,2b
14,b11)▼

DR 1 Leader (Data Set General Information record)

00321□D□□□□□00065□□□3304

DR 1 Directory

DSID104000DSSI065104ATCS070169FTCS017239▼

DR 1 Field Area

[0A][00000001]S-100□Part□10a▲5.0▲INT.IHO.S-101.1.1▲1.1▲1▲S100Example.000▲S-100□Encoding□example▲20221019EN▲▲1▲[0E][12]▼

[0.0][0.0][0.0][00989680][00989680][00000064][00000000][00000001][00000000][00000000][00000000][00000000][00000001]▼

buoyShape▲[0001]colour▲[0002]colourPattern▲[0003]featureName▲[0004]language▲[0005]name▲[0006]▼

BuoySafeWater▲[0001]▼

DR 2 Leader (Data Set Coordinate Reference System record)

00064□D□□□□□00039□□□2104

DR 2 Directory

CSID070CRSH187▼

DR 2 Field Area

[0F][00000001][01]▼

[01][01][01]WGS□84▲4326▲[02]▲▼

DR 3 Leader (Point Record)

00055□D□□□□□00037□□□1104

DR 3 Directory

PRID90C2IT99▼

DR 3 Field Area

[6E][00000001][0001][01]▼

[1948C740][F8C61DB0]▼

DR 4 Leader (Feature Type Record)

00218□D□□□□□00065□□□3304

DR 4 Directory

FRID011000FOID009011ATTR117020SPAS016137▼

DR 4 Field Area

[64][00000001][0001][0001][01]▼

[7C7C][00BC614E][002A]▼

[0001][0001][0000][01]4▲
[0002][0001][0000][01]3▲
[0002][0002][0000][01]1▲
[0003][0001][0000][01]3▲
[0004][0001][0000][01]▲
[0005][0001][0005][01]eng▲

```
[0006][0001][0005][01]Example□buoy▲
[0004][0002][0000][01]▲
[0005][0001][0008][01]deu▲
[0006][0001][0008][01]Beispiel□Tonne▲▼

[6E][00000001][FF][FFFFFFFF][00000000][01]▼
```

Recommendations

To increase the quality of the specification and to make implementation easier we recommend to endorse the proposed changes to S-100 Part 10a.

Action Required by the S-100WG

The S-100WG is invited to:

a. Note this paper
b. Discuss this paper
c. Endorse the proposals