STAR General File Format Requirements
Version 2.0

Table Of Contents

1. Document Information

1.1 Revision History

Revision Date	Version	Effective Date	Description
12/09/1996	1.00	04/01/1997	Initial Version
09/16/1998	1.1	10/01/1998	New Vendors/Codes added and minor changes to "Record Format" area. Reworked for clarity, added references to picture clauses
08/12/1999	1.2	9/01/1999	New Manufacturer and Vendor Codes, additions to data dictionary section and file naming sections based on STAR/DSP Data Element subcommittee analysis.
5/17/2000	1.3	8/01/2000	Added new short transaction Ids. Add EBCDIC as valid file transfer format, clarify file retention time, error trapping for duplicate file names. Added CR/LF formatting rules for text fields, added DIS short DSP code, added verbiage for real time data transfer (XML). Added removal of '.' from file name when necessary.
10/12/2000	1.94		Revised file naming. Clarified Picture Clause and data formatting. Modified retention time on DCS/DSP computer. Added note for directory/folder creation. Added Cycle date format. Added Comment use note.
12/12/00	1.95		Clarified use of quotations in file naming convention. Added additional examples of subdirectory naming convention. Added verbiage to File Format section to clarify record type.
01/02/01	1.96		Revised examples of subdirectory naming convention
01/24/01	1.97		Revised file naming and location. Revised example of subdirectory naming convention. Clarified the case of multiple transactions in one file. Revised XML basic guidelines to reference OAGI and to version the XML section beginning with 1.0.
02/01/01	1.98		Removed references to subdirectories. Increased file randomizer from 6 characters to 7 characters. Removed examples 2, 3, and 4. Edited example 1 a and b to reflect new file naming convention.
03/01/01	2.0		Removed "source" and "vendor" from the 2-character manufacturer code. Indicated that other directory delimiters may be used as required by the

			operating system. Replaced all 'AMDSG' references to 'STAR'.
05/15/01	2.0	01/01/02	Removed "DRAFT" from title page and headers. Made effective for implementation January 1st, 2002

2. File Naming and Location
Describes file naming convention and where they may be located

- The interface file may be located on any magnetic media, or as a data stream through a supported LAN or serial connection. If the file is located on a magnetic media it must be located in one or more directories found on the media's directory (Location and creation of directory/subdirectory needs to be negotiated between sender and receiver when implementation plan is created) in the following format:

"sss2ccc"	where quotes are for punctuation and not to be included in directory name and …
"sss"	identifies the source of the information being transferred. Can be either from Dealer System Provider "DSP" or Dealer Communications System "DCS".
"2"	is a constant "2" indicating information being transferred from a source system to a target system.
"ccc"	identifies the target of the information being transferred. Can be either to Dealer System Provider "DSP" or Dealer Communications System "DCS".
"mmttnnn.nnndddddddddddddd"	is an alphanumeric filename up to 24-characters in length. Where…
"mm"	is a 2-character *manufacturer* code found on the "*Short Manufacturer and Vendor Codes*" table.
"tt"	is a 2-character *transaction code* that indicates the content/nature of the file data. This code may be found in the "*Short Transaction Codes*" table.
"nnnnnnn"	is a zero-padded, 7-character code used to randomize what might otherwise be duplicate names when many files containing similar information are created. When created, filenames should be numbered from 0000000-9999999 as necessary to avoid duplicate naming. Each file created should use the smallest 7-character code available. When the upper limit of the code is reached, it should roll over to 0000000. Application error trapping may be required to resolve duplicate file names.
"." *	The optional file-naming delimiter required by some systems. The optional decimal point is considered a significant position in all applications.
"ddddddddddddd"	is an alphanumeric left justified target *dealer identifier,* up to 13-characters in length.

- Interface 1 : Interface filenames are of the following format: "mmttnnn.nnndddddddddddddd"   or "mmttnnnnnnndddddddddddddd" where...

"mm"	is a 2-character *target manufacturer* code found on the "*Short Manufacturer and Vendor Codes*" table.
"tt"	is a 2-character *transaction code* that indicates the content/nature of the file data. This code may be found in the "*Short Transaction Codes*" table.
"nnnnnnn"	is a zero-padded, 7-character code used to randomize what might otherwise be duplicate names when many files containing similar information are created. When created, filenames should be numbered from 0000000-9999999 as necessary to avoid duplicate naming. Each file created should use the smallest 7-character code available. When the upper limit of the code is reached, it should roll over to 0000000. Application error trapping may be required to resolve duplicate file names.
"."	The optional file-naming delimiter required by some systems. The optional decimal point occupies a location that is considered a significant position.
"ddddddddddddd"	The dealer code up to 13-characters in length.

* **Optional only in the case that the filename can be uniquely attributed to the sender/recipient by the file name alone or if the location of the file uniquely identifies the sender/recipient in some other fashion (i.e., a host machine that only contains files for a specific dealer).**

Example :　　　**Directory/Filename**

Example a:　　　**"DCS2DSP/OEPR123456757896" Would indicate that the 1,234,567th parts return file from the OEM was transferred from a Dealer Communications System to a Dealer Service Provider for Dealer 57896. In this example, the source manufacturer/vendor code, transaction type, and file sequence number are concatenated with the dealer number.**

Example b:　　　**"DSP2DCS/OEPR123.45657896" Would indicate that the 123,456th parts return file from Dealer 57896 was transferred from the DMS to the OEM (OE). In this example, the source manufacturer/vendor code, transaction type, and file sequence number are concatenated with the dealer number. In this example, the "." Indicates the optional file-naming delimiter required by some systems. The optional decimal point is considered a significant position.**

- **Once created, the interface files shall remain on the DCS/DSP computer for an agreed upon period of time between the sender and receiver of the file. Provisions for allowing selection and deletion of individual files before agreed upon retention period should be available to the user. This applies only to those files that are created and reside on the DCS/DSP computer's hard disk/network system.**

- **The interface file may be located on the Windows Clipboard as "CF_TEXT" format and must otherwise conform to the indicated file and record format. Only 1 interface file may reside on the Windows Clipboard at any given time. A file residing on the windows clipboard must be preceded with header information of the following format:　　"directory/filename/<CRLF>..."　　where....**
  - **"directory"　is either "DCS2DSP" or "DSP2DCS" as specified above.**
  - **"filename"　　　　is a filename as described above, with manufacturer code, transaction type, file sequence number, and dealer number.**
  - **"/"　　　　　　are literal slashes and are directory delimiters required where indicated. Other directory delimiters may be used as dictated by the operating system.**
  - **<CRLF>　　　　is a required Carriage-Return/Line-Feed sequence.**
  - **"..."　　　　　　is where the file contents starts and runs to its end.**

3.　Batch Data Transfer Format

3.1　File Format

- **In conventional terms, the file can be considered a "stream" or "flat" ASCII file.**
- **The interface file must use the ASCII character set. Only standard printable characters can be included.**
- **The file will contain a number of records of variable length each ending with a Carriage-Return/Line-Feed character sequence as described in the Record Format section below.**
- **All records and data fields in the file are optional with the exception of a few fields and the Ident Record. Without the Ident Record, the file content cannot be determined and there is no reason to send the file. For all other records, they are optional and are specific to the transaction being sent and may or may not be available for creation depending on the data in the file.**
- **The file may reside in the Windows Clipboard, subject to requirements mentioned in the File Naming and Location section above.**
- **Big-endian/Little-endian issues do not apply as data is sequential text.**

- **STAR file/field data is not presented under Unicode standards. Data is represented as a single byte-per-character.**
- **It may be necessary to transfer, manipulate and store data in a non-ASCII environment (such as EBCDIC). In such non-ASCII environments, the resident (Native) system/application and its support infrastructure is fully responsible for the following:**
    1. **All data transfers to and from the ASCII environment.**
    2. **All ASCII–to–Native data translations and vice-versa.**
    3. **The accuracy and integrity of the above-mentioned translations.**
    4. **Insuring the proper local/Native processing of STAR-conforming files as necessary.**
    5. **Insuring that any resulting/output ASCII files conform to STAR specifications.**
    Intermediate data and files residing in a non-ASCII environment are not described within the STAR batch standards documentation.

3.2   Record Format

- **Records are of variable length, each being terminated by the carriage-return/line-feed character sequence (CRLF).**
- **The CRLF is not required for last record of the file.**

3.2.1   Data Record Type

Contains data in format previously defined by record definition records.

- **The format of the data record type is:**

    "record.identifier: field.data.1, field.data.2, field.data.3, ... field.data.*n*"   where...

"record.identifier"	**is the identifier of the record. It must match a Record Identifier name defined for the transaction in process.**
":"	**delimits the "record.identifier" from "field.data.1" and is required.**
"field.data.1", *etc...*	**are positional data for the data fields (Data Elements) defined for the "record.identifier". Data fields must be separated by a comma. If a data field contains one or more blanks, commas or double-quotes, the data must be enclosed in double-quotes. A double-quote inside a quoted data item must be represented by 2 adjacent double-quotes. If a data item is not quoted, it will be considered to contain no leading, trailing, nor imbedded blanks. Any apparent leading or trailing blanks are assumed to be insignificant and not part of the data item.**

    NOTES:
    1. **All "field.data..." values are positional and relate directly to the Data Elements found on the related record definition. If any fields are empty or contain no value, then a follow-on comma appear as a place-holder for relevant fields that follow. If at any position in the record should all remaining fields be empty or contain no value, then the record may end at that point.**
    2. **Should all fields of an entire data record be empty or contain no value, the record may be omitted from the file.**

3.2.2   Comment Record Type

Contains comments not related to definitions nor data, and may occur anywhere in the file *after* the "xx.IDENT" record.

- **The format of the comment record type is:**

"!xxx..."	**where...**
"!"	**denotes the comment and is required.**
"xxx..."	**is any free-form comments and is optional.**

NOTE: Comment statements may occur in all specifications following the IDENT record in any sequence.

3.3   Multiple transactions in same file

Most transactions have a certain record denoted as a *header* record (PO.HEADER for Parts Orders, and DR.VEHICLE for Delivery Registration are 2 examples of header records). For such transactions, there is 1 *header* record per transaction. In cases where there may be more than 1 transaction present per file, the transactions will be the same transaction type, will be consistent with the filename, and most transactions accept for Pick List have certain records noted as *header* records. Besides containing information pertaining to the current transaction, the *header* record indicates the *beginning* of a new transaction. Any and all other records encountered are considered part of the *header*'s transaction, until a new *header* or the end-of-file is encountered. Records should appear in the order and groupings as indicated in each specification unless otherwise stated.

4.   Real Time Data Transfer

4.1   XML Basic Guidelines v1.0

- **All objects (data elements) that are optional in STAR transactions will be optional in their XML equivalent definitions.**
- **XML Schema names will be Camel Case abbreviated version of Data Dictionary Data Element Names based on the Common STAR Abbreviations Table.**
- **OAGI XML guidelines will be adopted for STAR XML standards. XML messaging naming guidelines will be maintained with a version control separate from file transfer specifications.**

5.   Data Dictionary

The intention of having a data element type specified in the STAR data dictionary is so that numeric data would be passed in a data element that is defined as numeric.   Specifying that a data element should be numeric does not imply any special validation rules for the system from which the information is created. When data is extracted and formatted for transfer, if the data element is defined as numeric it should have a numeric value.

5.1   Picture Clauses

The STAR Data Dictionary contains a number of picture clauses used to define the appearance of fields. Picture clauses are used to indicate the length/capacity and characteristics of a field. The characters used for the picture clause are as follows:

Character	Meaning of character in Picture Clause context
9	Denotes numeric data. Each occurrence of a "9" holds the place of a single numeric digit ("0"-"9"), and indicates another digit position in the field. Trailing zeros to the right of a decimal point may be omitted.
V	Indicates decimal position. This decimal point indicates how much significance is represented on either side of the decimal point. It is required to have a hard-coded decimal point present in all numeric data values that contain a decimal.
S	Indicates that the numeric field may be "signed". That is, may have a "+" or "-" sign preceding the first digit of the field value. This picture character *does not* represent a digit position and is always the first character of the picture clause. Within the

Character	Meaning of character in Picture Clause context
	context of STAR interface files, records and fields, all *unsigned* (no sign) numeric values will be assumed positive. Values that are to be represented as negative must be preceded by the negative ("-") sign.
X	Indicates a field and length that may contain any printable characters. One "X" is used to represent each character in length.
(*n*)	The *n* indicates the number of times the preceding character ("9" or "X") occurs. This is used to short-hand the number of occurrences of a picture character. The parentheses serve no other purpose than to bracket the *n*. For example, "9(5)" is equivalent to "99999", "S9(2)v9" is equivalent to "S99v9", "9(5)v9(5)" is equivalent to "99999v99999", X(10) is equivalent to "XXXXXXXXXX", and so on.

5.2   Data Examples

Some examples of field data values and how they would be read under some picture representations, under normal circumstances per above:

Data element data value as it appears in file	Picture clause used 99999 **5 digits, no decimal**	Picture clause used 999v99 **5 digits, 2 decimals**	Picture clause used S99999 **5 digits, signed** [1]	Picture clause used X(8) **8 characters**	Picture clause used X(3) **3 characters**
01234	1234	01234	1234	01234	01234
012.34	12	12.34	12	012.34	012.34
-1841	-1841	-1841	-1841	-1841	-1841
-184.1	-184	-184.10	-184	-184.1	-184.1
"AB CD"	*n/a - error*	*n/a - error*	*n/a - error*	AB CD	AB CD

NOTES: [1] The "signed" indicator ("S") indicates the sign is desired. However, a positive value may be represented without a sign.

5.3   Data Dictionary Report Field Explanations

Item	Description	Sample Values
Data Element Name	Name of data element.	Vehicle Identification Number
XML Schema Name	XML Data Element schema name	PartNo
Data Element Definition	Business use description of data element.	Federally defined 17 position vehicle identification number
Interface	DMS Application interface associated with data element	Delivery Reporting

Item	Description	Sample Values
Record	DMS Application interface sub-classification	Vehicle
Data Root Reference Number	Root Reference number used to assign attributes to Data Element	1310
Min	Minimum field length for data item if data is available for transfer.	1
Max	Maximum field length for data item. Must be equal to or greater than Minimum length.	18
Pic Clause	COBOL defined picture clause for data item.	X(18)
Data Element Number	Unique identification number assigned to data element	103680
Valid Value Rule	Valid Value Rule for data element. Rule definitions are contained in the Valid Value Rules Table.	Pass Through
Add/ Chg	Codes identifying type of change made to data element since last publication	NM DE
Sample Data	Sample of data element information	125.78

6.   Data Format

6.1   Character Format
- **As a general rule, character fields should be presented "trimmed" of leading and trailing blanks.**
- **Character fields need not be quoted unless they contain one or more of the following: a space (blank), comma and/or double-quote (each imbedded double-quote must be represented by 2 adjacent double-quotes "").**
- **If a field's character data exceeds the size indicated by the picture clause, the full value of the field should be presented. For example, if the field data for a picture clause of X(4) is "123456", then the field value should be "123456".**
- **If a field's character data length is less than the size indicated by the picture clause, the exact length of the field should be presented. For example, if the field data for a picture clause of X(4) is "23", then the field value should be "23".**

- If the attribute of a data field is X(255), it designates that the text data field does not have a maximum length. i.e.; NOMAX

6.2 Comment Field Format

- Because of potential translation problems with remote systems that use the EBCDIC character set, the carriage return – line feed (CR/LF) get stripped from text fields which contain them. This primarily affects comment fields defined in the STAR interface specifications. Since the CR/LF characters get stripped out of the comment fields that are sent in the STAR interface files, there is no way to preserve the original formatting of the comment field as it was originally entered by the dealership.

    In order to preserve the original formatting of the comment fields, the STAR recommended practice is to replace the carriage return (CR) character with the tilde (~) and the line feed (LF) character with the caret (^) where those characters are within a single field to be transmitted. Note that only inline CR/LF combinations should be translated, end-of-record CR/LF combinations should not be.

    For example: The comment field as entered at the dealership looks like (The CR/LF characters are not typed in by the dealer, they are only shown for reference.):

    "The customer complains of three problems: (CR/LF)
    1. Loud noise when accelerating. (CR/LF)
    2. The vehicle appears to be leaking oil from the rear main seal. (CR/LF)
    3. Difficulty shifting the car into drive from reverse. (CR/LF)"

    The comment field as transmitted in STAR format should look like:

    "*The customer complains of three problems:~ ^1. Loud noise when accelerating.~ ^ 2. The vehicle appears to be leaking oil from the rear main seal.~ ^ 3. Difficulty shifting the car into drive from reverse.~ ^*"

6.3 Date Format

- Full Dates should be formatted in century, year, month and day sequence without any separators. i.e., YYYYMMDD
- Financial Period Dates should be formatted in century, year and month sequence without any separators. i.e., YYYYMM
- Cycle Dates should be formatted in century, year, month and cycle number without any separators. i.e., YYYYMMC

6.4 Full Name Format

- Format for names that are not separated into separate data fields should be "First Last" with no commas between first and last name.

6.5 Numeric Format

- Numeric data fields in interface files should always have a decimal point present. If a numeric field exists without a decimal point, it will be considered a whole number.
- Positive numeric field data need not have the positive sign (+) present, even if indicated in the picture clause (such as S9v99). Numeric values will always be assumed positive unless preceded by a negative sign (-). For example, 5.50 is equivalent to +5.50.
- Numeric field data that is negative (-) rather than positive (+) should be represented signed, such as -25.50.

- If a data field's numeric data exceeds the size/precision indicated by the picture clause, the full value of the field should be presented. For example, if the field data for a picture clause of "999" is 1002, then the data field value should be 1002. Also, if the field data for a picture clause of 99.99 is 13.342, then the field value should be 13.342.
- Numeric fields should not be enclosed in double-quotes.

6.6   Social Security Number Format

- Social Security numbers should be formatted as nine numbers with no dashes. i.e.; 123456789
- Some DSPs may not be able to support the transmission of social security numbers based upon pending legislation.

6.7   Time Format

- Time should be formatted in hours, minutes and second sequence without any separators. i.e., HHMMSS

7. Short Manufacturer/DSP Codes

Short Manufacturer/DSP Codes Report available on STAR Web site.

8. Short Transaction Codes for File Naming

Short Transaction Codes Report available on STAR Web site.