
Example set of DFDL 1.0 properties

Status of This Document

This working draft document provides an example to the OGF community of the Data Format Description Language (DFDL) standard. Distribution is unlimited

Copyright Notice

Copyright © Open Grid Forum (2012). Some Rights Reserved. Distribution is unlimited.

Abstract

This document provides a set of DFDL properties that can be used as defaults for creating DFDL schema.

Table of contents

1. Introduction.....	2
2. Example DFDL defineFormat	3
3. Authors and Contributors	6
4. Intellectual Property Statement	6
5. Disclaimer	6
6. Full Copyright Notice	6
7. References	6

1. Introduction

Data Format Description Language 1.0 [DFDL] is a language with a rich set of properties for modeling text and binary data formats. For portability of DFDL schemas, the language is designed to have very few built-in default values for these properties. This means that a DFDL schema can not be silent about any property that is applicable to any of the objects in the schema. To assist first-time DFDL schema authors, this document provides a `defineFormat` annotation that can be used to provide defaults for just about all DFDL properties.

The `defineFormat` annotation is defined within a DFDL schema which itself defines no schema objects. The recommended usage is for the DFDL author to refer to this DFDL schema from his own DFDL schema using `xs:import [XSDLV1]`. If any property values need to be changed, it is recommended that the DFDL author either:

- overrides the property values in his own DFDL schema
- makes a copy of this DFDL schema in a different namespace and overrides the property values in the copy.

The values for the properties are suitable for modeling text data, in particular they can be used without modification when modeling standard Comma Separated Value (CSV) text data [RFC4180].

The `defineFormat` annotation provides defaults for all properties where it makes sense to do so. However there are some properties where a default is not appropriate, and for these the `defineFormat` is silent. An example is the *length* property, which if used would typically vary on a per element basis.

DFDL escape scheme property values are defined in an accompanying `defineEscapeScheme` annotation which is referenced from the `defineFormat` annotation.

Section 2 defines the DFDL schema containing the `defineFormat` and `defineEscapeScheme` annotations.

2. Example DFDL defineFormat

In the schema below, the file name is assumed to be DFDLGeneralFormatDefinition.xsd.

When copying the text into a DFDL schema file, the copyright statement must be retained.

```
<?xml version="1.0" encoding="UTF-8"?>
<!--
*****
* (C) Copyright IBM Corp 2011
* All Rights Reserved
*
* This DFDL schema provides a set of DFDL property values to assist
* in the creation of DFDL schema that model general text or binary
* data, or a mixture of text and binary data.
*
* DFDL Version 1.0
* Schema Version 1.0
*****
-->
<!--
*****
* The primary purpose of this schema is to provide sensible values
* for the vast majority of DFDL properties as an illustration.
* However the properties are a consistent set that model a variable
* length text format consisting of a single record of CSV data,
* (for example, aaa,bbb,ccc,ddd,eee).
*
* The recommended usage is to refer to this schema from other DFDL
* schema using a xs:import statement. If any property values need
* to be changed, it is possible to update this schema directly,
* but it is recommended either to override them locally in the
* schema that includes this one, or make a copy of this schema in
* a different namespace.
*
* For example, to model fixed length, binary data then override
* "representation" to "binary", override "lengthKind" to "explicit"
* and set "length".
*
* Example usage:
*
* 1) To apply the properties in this schema as defaults for all
* objects in another DFDL schema:
*
* <xs:schema xmlns:tns="http://www.ogf.org/dfdl/dfdl-1.0/examples/" ... >
*
*   <xs:import namespace="http://www.ogf.org/dfdl/dfdl-1.0/examples/"
*     schemaLocation="DFDLGeneralFormatDefinition.xsd"/>
*
*   <xs:annotation>
*     <xs:appinfo source="http://www.ogf.org/dfdl/">
*       <dfdl:format ref="tns:GeneralFormat"
*         ... override property values if desired ...
*       </dfdl:format>
*     </xs:appinfo>
*   </xs:annotation>
*
*   ...
*
* </xs:schema>
*
* 2) To apply the properties in this schema to specific schema objects
* and override selected property values:
*
* <xs:schema xmlns:tns="http://www.ogf.org/dfdl/dfdl-1.0/examples/" ... >
```

```
*
*   <xs:import namespace="http://www.ogf.org/dfdl/dfdl-1.0/examples/"
*           schemaLocation="DFDLGeneralFormatDefinition.xsd"/>
*
*   <xs:element name="root" dfdl:ref="tns:GeneralFormat"
*           dfdl:lengthKind="explicit" dfdl:length="80">
*       <xs:complexType>
*           <xs:sequence dfdl:ref="tns:GeneralFormat" dfdl:separator="">
*               ...
*           </xs:sequence>
*       </xs:complexType>
*   </xs:element>
*
* </xs:schema>
*****
-->

<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:dfdl="http://www.ogf.org/dfdl/dfdl-1.0/"
    targetNamespace="http://www.ogf.org/dfdl/dfdl-1.0/examples/"
xmlns:tns="http://www.ogf.org/dfdl/dfdl-1.0/examples/" >

    <xs:annotation>
        <xs:appinfo source="http://www.ogf.org/dfdl/">

            <!--
            A general purpose defineFormat annotation.
            -->

            <dfdl:defineFormat name="GeneralFormat">
                <dfdl:format
                    encoding="UTF-8" utf16Width="fixed" byteOrder="bigEndian"
                    ignoreCase="no" outputNewLine="%CR;%LF;"

                    alignment="1" alignmentUnits="bytes" fillByte="0"
                    leadingSkip="0" trailingSkip="0"

                    lengthKind="delimited" lengthUnits="characters"
                    prefixIncludesPrefixLength="no"

                    representation="text" textPadKind="none" textTrimKind="none"
                    escapeSchemeRef="tns:GeneralBlockEscapeScheme"

                    textBidi="no" textBidiTextOrdering="implicit"
                    textBidiSymmetric="yes" textBidiTextShaped="no"
                    textBidiNumeralShapes="nominal" textBidiOrientation="RTL"

                    textStringJustification="left" textStringPadCharacter="%SP;"
                    truncateSpecifiedLengthString="no" textOutputMinLength="0"

                    textNumberJustification="right" textNumberPadCharacter="0"
                    decimalSigned="yes" textNumberCheckPolicy="lax"
                    textNumberRep="standard" textStandardBase="10"
                    textNumberRounding="pattern" textNumberRoundingMode="roundUp"
                    textNumberRoundingIncrement="0.0"
                    textStandardDecimalSeparator="." textStandardGroupingSeparator=", "
                    textStandardExponentCharacter="E" textStandardZeroRep=""
                    textStandardInfinityRep="Inf" textStandardNaNRep="NaN"
                    textNumberPattern="#0" textZonedSignStyle="asciiStandard"

                    textBooleanJustification="left" textBooleanPadCharacter="%SP;"
                    textBooleanTrueRep="true" textBooleanFalseRep="false"

                    textCalendarJustification="left" textCalendarPadCharacter="%SP;"
                    calendarPatternKind="implicit" calendarPattern="yyyy-MM-dd'T'HH:mm:ss"
                    calendarCheckPolicy="lax" calendarTimeZone="UTC"
                    calendarObserveDST="yes" calendarFirstDayOfWeek="Monday"
                    calendarDaysInFirstWeek="4" calendarCenturyStart="53"
```

```
calendarLanguage="en-US"

occursCountKind="parsed"

sequenceKind="ordered" separator="," separatorPolicy="required"
separatorPosition="infix" initiatedContent="no" floating="no"

choiceLengthKind="implicit"

initiator="" terminator="" documentFinalTerminatorCanBeMissing="no"
emptyValueDelimiterPolicy="none"

nilKind="literalValue" useNilForDefault="no" nilValue="NIL"
nilValueDelimiterPolicy="none"

binaryNumberRep="binary" binaryPackedSignCodes="C D F C"
binaryDecimalVirtualPoint="0" binaryNumberCheckPolicy="lax"
binaryFloatRep="ieee"

binaryCalendarRep="bcd" binaryCalendarEpoch="1970-01-01T00:00:00+00:00"

binaryBooleanTrueRep="1" binaryBooleanFalseRep="0">
</dfdl:format>
</dfdl:defineFormat>

<!--
An example defineEscapeScheme annotation.
Assumes 'block' style escaping.
-->

<dfdl:defineEscapeScheme name="GeneralBlockEscapeScheme">
  <dfdl:escapeScheme
    escapeKind="escapeBlock"
    escapeBlockStart="" escapeBlockEnd=""
    escapeEscapeCharacter="" generateEscapeBlock="whenNeeded"
    extraEscapedCharacters="&#x0D; &#x0A;">
  </dfdl:escapeScheme>
</dfdl:defineEscapeScheme>

</xs:appinfo>
</xs:annotation>
</xs:schema>
```

3. Authors and Contributors

Stephen M. Hanson,
IBM Software Group,
Hursley,
Winchester, UK
smh@uk.ibm.com

We greatly acknowledge the contributions made to this document by the following people.

- Suman Kalia, IBM Software Group, Markham, Ontario, Canada

4. Intellectual Property Statement

The OGF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the OGF Secretariat.

The OGF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights, which may cover technology that may be required to practice this recommendation. Please address the information to the OGF Executive Director.

5. Disclaimer

This document and the information contained herein is provided on an "As Is" basis and the OGF disclaims all warranties, express or implied, including but not limited to any warranty that the use of the information herein will not infringe any rights or any implied warranties of merchantability or fitness for a particular purpose.

6. Full Copyright Notice

Copyright © Open Grid Forum (2012). Some Rights Reserved. Distribution is unlimited.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included as references to the derived portions on all such copies and derivative works. The published OGF document from which such works are derived, however, may not be modified in any way, such as by removing the copyright notice or references to the OGF or other organizations, except as needed for the purpose of developing new or updated OGF documents in conformance with the procedures defined in the OGF Document Process, or as required to translate it into languages other than English. OGF, with the approval of its board, may remove this restriction for inclusion of OGF document content for the purpose of producing standards in cooperation with other international standards bodies.

The limited permissions granted above are perpetual and will not be revoked by the OGF or its successors or assignees.

7. References

- [DFDL] DFDL 1.0, <http://www.ogf.org/documents/GFD.174.pdf>
- [XSDLV1] XML Schema Part 1: Structures, <http://www.w3.org/TR/xmlschema-1/>
- [RFC 4180] Common Format and MIME Type for CSV Files, <http://www.ietf.org/rfc/rfc4180.txt>