Questions and Answers on Document Construction with XML

Questions and Answers on Document Construction with XML

Table of Contents

Purpose of this Q and A type document

The purpose of this list is to check and enhance the *understanding* of XML use in the areas of publishing and industry schemas.

Of course knowledge about specific XML technologies is needed as well, so the questions go from more complex ones which offer opportunities for design discussions to simpler ones which cover specific technologies.

I would like to thank Martin Maier, Oliver Margraf for contributing to this list.

Chapter 1. Markup Languages in General

XML and HTML

HTML, XHTML and XML are all so called markup languages. This raises some questions about general concepts of markup languages as well as what makes them different.

1.

   What makes markup languages different e.g. from the concept used by WYSIWYG editors like word, wordperfect, Quark Express etc?

   tbd

2.

   Was macht markup Sprachen anders im Vergleich zu WYSIWIG Editoren wie Word, Wordperfekt, Quark Express etc?

   tbd

3.

   What is the advantage of a document written in XML compared to proprietary formats like Word, Quark Express etc.

   tbd

4.

   Welchen Vorteil bringt ein XML Dokument im Vergleich zu propietären Formaten wie Word, Quark Express etc?

   tbd

5.

   What are the main differences between XML and HTML?

   ### Tip

   Could you e.g. express a database record well in HTML? Think about what kind of tags are available and what their main purpose is.

   tbd

6.

   Nennen Sie die wesentlichen Unterschiede zwischen XML und HTML.

   ### Tipp

   Könnten Sie einen Datanbank Rekord gut in HTML ausdrücken? Was für Tags sind verfügbar und welchem Zweck dienen sie in erster Linie?.

   tbd

7.

   XML is said to be extensible and HTML not. Explain

Tip

Think about who decides which tags are allowed?

tbd

8.

In the above question, is the "extensibility" of XML something technical or something political/social?

Tip

Can you extend *every* xml schema?

tbd

9.

What is the advantage of using UNICODE compared to ASCII in XML documents?

Tip

Think about WHERE xml will be used?

tbd

10.

Was ist der Vorteil von XML Dokumenten in Unicode anstatt ASCII?

Tipp

Denken Sie daran wo XML überall benutzt werden kann?

tbd

Chapter 2. Xtended Markup Language XML

XML parts

Here we take a look at the main parts of an XML file.

1.

What are the main ingredients of an xml instance that you will most likely need or encounter in documents?

XML declaration, element tags, attributes, comments, processing instructions, name space declarations and entities

2.

Does an XML instance always need a DTD?

foo

3.

Braucht eine XML Instanz immer eine DTD?

tbd

4.

Explain the meaning of * + ? in einer DTD?

"*" means zero, one or many occurrences of an element in an XML instance are legal. "+" means that an element needs to be there at least once but possibly more often and "?" means that an element can be there or not.

5.

Erklären Sie die Bedeutung von * + ? in einer DTD?.

"*" bedeutet keinmal, einmal oder beliebig oft kann ein Element in der XML Instanz vorkommen. "+" bedeutet mindestens einmal aber auch mehrmals und "?" bedeutet einmal oder keinmal ist legal.

6.

Explain the meaning of #REQUIRED, #IMPLIED, #FIXED when used in the definition of attributes?

Tip

What would you like to enforce about the use of attributes? What if the value of an attribute would always be the same?

tbd

7.

Erklären Sie die Bedeutung von #REQUIRED, #IMPLIED, #FIXED bei der Definition eines Attributes?

Tipp

Was möchten Sie bei Attributen eventuell erzwingen? Was ist wenn der Wert eines Attributes immer gleich ist?

tbd

8.

A DTD defines a catalog element. Add an attribute "created" to this element. <!ELEMENT catalog (#PCDATA)>

tbd

9.

Eine DTD definiert ein katalog Element. Fügen Sie ein Attribute "erstellt" hinzu. <!ELEMENT katalog (#PCDATA)>

tbd

10.

What is the purpose of attributes in XML. Do you have to use them? Give some guiding rules for the use of attributes.

Tip

Think about extensibility. How does this differ between attributes and elements? Think about your application and how it will react.

tbd

11.

What happens in the following fragment of an instance when it is parsed? <!ENTITY bmw "Bayrische Motorenwerke" > <text>And the worlds best motorbikes come from &bmw; in Munich </text>

tbd

12.

Was passiert wenn folgendes Fragment einer Instanz durch den Parser geht? <!ENTITY bmw "Bayrische Motorenwerke" > <text> Und die besten Motorräder der Welt kommen von &bmw; in Müchen </text>

tbd

13.

What kind of definitions are used in DTDs?

Tip

Think about the xml parts that show up in instances. How are they defined?

tbd

14.

In the above question, is the "extensibility" of XML something technical or something political/social?

Tip

Can you extend *every* xml schema?

tbd

Namespaces in XML

This section covers namespace use in XML.

1.
   The example below shows an XML fragment. What does it mean? xml namespace decl and use goes here

   tbd

2.
   What are namespaces mainly used for in XML?

   ## Tip

   XML lets you define your own tags. Think about the consequences if everybody does define their own tags.

   Eindeutigkeit Kontext. Namensraum wird verwendet um die XSL-Tags von anderen Tags zu unterscheiden (?)

1.
   Geben Sie für jedes Element und Attribut des abgebildeten XML-Dokuments den zutreffenden Namensraum an. .

   tbd

Chapter 3. XML examples

XML instances

Here we take a look at some specific features of XML instances.

1.

    The example below shows a xml segment. Can you explain what you see and what it might be?

    Creating relations in XML documents

2.

    The example shows two different means of creating relatins within XML. One uses special XML features while the other one is made by the authors themselves. Explain both mechanisms and tell which one is application specific?

    ### Tip

    a segment of JDF linking is shown.

    tbd

3.

    In the above question, is the "extensibility" of XML something technical or something political/social?

    ### Tip

    Can you extend *every* xml schema?

    tbd

4.

    Given the following XML instance, write a proper DTD for it: <CD> <artist>....................</artist> <title>.................................</title> <price>...............................</price> </CD>

    tbd

5.

    Schreiben Sie eine DTD für die folgende XML Instanz: <CD> <Interpret>....................</Interpret> <Title>.................................</Title> <Preis>...............................</Preis> </CD>

    tbd

Chapter 4. Modelling XML information

Modelling XML information - using DTDs or XML Schema or something else like Relax NG - is perhaps the most time consuming and important task. Only topped by the effort to create an industry-wide schema like JDF.

General questions about modelling

This section covers questions about when to use what modelling construct - DTD, schema etc. -

1.

Welche Bedeutung hat eine DTD? Was wird darin definiert?

Tipp

Denken Sie an die Vorteile in Bezug auf die Korrektheit eines Dokuments. Korrekt gegenüber was?

tbd

2.

You and your friends are all heavy into lounge music from Paris and you decide to share information about CDs that you own or want to have. With ratings etc. Whhat would be the advantage of creating a DTD if you want to share this info?

Tip

Think about what happened when we created instances of our catalog without a dtd first. We had some model, yes, but....

tbd

3.

Some XML instances are called well-formed and others are called valid? What is the difference?

Tip

Think about what you need for a document to be valid. Certain elements? Or just good looking tags?

tbd

4.

Was ist ein gültiges und was ein wohlgeformtes XML-Dokument?

Tipp

Denken Sie daran was Sie brauchen damit ein Dokument valide sein kann. Gewisse Elemente oder nur Tags die gut aussehen?

Wohlgeformt: Tags stimmen überein. Auf jedes öffnende Tag gibt es auch ein schließendes Tag. es gibt eine Wurzel. Gültig dann wenn es wohlgeformt ist, Dokument verweist auf eine DTD Dokument hält die DTD auch ein

5.

Exactly what allows a Document Type Definition to be checked? The existence of certain elements? The order of elements? The proper text in elements?

Tip

Think about creating a valid instance of xml that fits to a DTD. What can you type in without getting an error and where does the parser or editor tell you that it is wrong?

Order of elements, number of elements and types of attributes can be checked. There is no way to check whether an elements text is OK except if it has been declared as empty and contains text.

6.

Look at the example XML instance and DTD below and say if the instance is a proper instance with respect to the DTD or if there are mistakes in the instance. Which mistakes would you fix where?

Tip

The mistake can be in instance (e.g. if the DTD defines an element that the instance does not have) or it can be in the DTD because it does not correctly fit to our model that we need.

Example of instance with missing element and dtd with wrong cardinality

tbd

7.

Was ist der Unterschied zwischen einem DTD und XML-Schema ?

beide beschreiben die Grammatik eines XML Dokus, aber: XML-Schema ist in XML Syntax formuliert (???) DTD hat nur beschränkte Datentypen und Inhalte XML-Format ermöglicht die Verwendung der XML-Techniken (XLink, XSL) sowie der bestehenden XML-Applikationen für XML-Schema

8.

What is the alternative to a Document Type Definition (DTD) and when would you use it?

Tip

DTDs where invented first and used a lot by authors and publishers. With XML more and more programmatic use has become popular e.g. to represent database records or programming language types in XML.

tbd

9.

Was kann man alternativ zu einer DTD verwenden und wann tut man das in erster Linie?

Tipp

DTDs wurden zuerst erfunden und vor allem von Autoren viel genutzt. Durch XML ist auch die Nutzhung in Programmen stark vermehrt worden, z.B. um Datenbank Einträge oder Typen aus Programmiersprachen abzubilden in XML.

tbd

10.

You have some very data oriented information that needs to get modelled. What construct: DTD or XML schema would you pick for the job and why?

Tip

Think about what makes data *correct* and how you would express and check correctness.

tbd

11.

You need to write a book or article and you want it transformed into html and pdf with little effort from your side. What modell for your text comes to mind?

If you decide to use docbook e.g., you will get a model that will certainly cover your needs with respect to publishing constructs (from book and authorinfo over references, mediaobjects, indexes and table of contents etc. And on top of it there exist free rendering stylesheets to produce various output formats