

Postal services — Hybrid mail — Document type definitions for customer to operator: a common set of default tags

ICS 03.240; 35.240.99

National foreword

This Draft for Development is the English language version of ENV 14014:2001.

This publication is not to be regarded as a British Standard.

It is being issued in the Draft for Development series of publications and is of a provisional nature because it is considered that further experience is required in its application before it is converted into a British Standard. It should be applied on this provisional basis, so that information and experience of its practical application may be obtained.

Comments arising from the use of this Draft for Development are requested so that UK experience can be reported to the European organization responsible for its conversion into a European Standard. A review of this publication will be initiated 2 years after its publication by the European organization so that a decision can be taken on its status at the end of its three-year life. The commencement of the review period will be notified by an announcement in *Update Standards*.

According to the replies received by the end of the review period, the responsible BSI Committee will decide whether to support the conversion into a European Standard, to extend the life of the prestandard or to withdraw it. Comments should be sent in writing to the Secretary of BSI Technical Committee SVS/4, Postal Services, at 389 Chiswick High Road, London W4 4AL, giving the document reference and clause number and proposing, where possible, an appropriate revision of the text.

A list of organizations represented on this committee can be obtained on request to its secretary.

Cross-references

The British Standards which implement international or European publications referred to in this document may be found in the BSI Standards Catalogue under the section entitled “International Standards Correspondence Index”, or by using the “Find” facility of the BSI Standards Electronic Catalogue.

This Draft for Development, having been prepared under the direction of the Management Systems Sector Policy and Strategy Committee, was published under the authority of the Standards Policy and Strategy Committee on 30 August 2001

Summary of pages

This document comprises a front cover, an inside front cover, the ENV title page, pages 2 to 67, and a back cover.

The BSI copyright date displayed in this document indicates when the document was last issued.

Amendments issued since publication

Amd. No.	Date	Comments

© BSI 30 August 2001

ISBN 0 580 38413 6

English version

Postal services - Hybrid mail - Document type definitions for customer to operator: a common set of default tags

This European Prestandard (ENV) was approved by CEN on 7 April 2001 as a prospective standard for provisional application.

The period of validity of this ENV is limited initially to three years. After two years the members of CEN will be requested to submit their comments, particularly on the question whether the ENV can be converted into a European Standard.

CEN members are required to announce the existence of this ENV in the same way as for an EN and to make the ENV available promptly at national level in an appropriate form. It is permissible to keep conflicting national standards in force (in parallel to the ENV) until the final decision about the possible conversion of the ENV into an EN is reached.

CEN members are the national standards bodies of Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom.



EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

Management Centre: rue de Stassart, 36 B-1050 Brussels

Contents

	page
1 INTRODUCTION	4
2 SCOPE	5
3 NORMATIVE REFERENCES	7
4 TERMS, DEFINITIONS AND ACRONYMS	8
4.1 Terms and definitions	8
4.2 Acronyms	8
5 META-SYNTAX	13
6 DEFINITION OF THE HML DTD	15
6.1 General rule for HML documents	17
6.2 General elements	18
6.3 Bundle elements	27
6.4 Letter elements	28
6.5 Page elements	36
ANNEX A (NORMATIVE): LIST OF ELEMENTS	38
ANNEX B (INFORMATIVE): DTD DESCRIPTION	41
ANNEX C (INFORMATIVE): EXAMPLES OF HML DOCUMENTS	48
BIBLIOGRAPHY	67

Foreword

This European Prestandard has been prepared by Technical Committee CEN/TC 331 "Postal services", the secretariat of which is held by NEN.

This ENV is related to Hybrid Mail and contains the syntax rules for a data stream for the submission of printing data to a Hybrid mail operator. It is a Document Type Definition (DTD) describing the data stream.

Annex A is normative. Annexes B and C are informative.

This standard includes a Bibliography.

According to the CEN/CENELEC Internal Regulations, the national standards organizations of the following countries are bound to announce this European Prestandard: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom.

1 Introduction

Hybrid Mail is the technology whereby input in one communication medium is converted for delivery on another communication medium according to the sender's instructions and/or the recipient's capabilities. The typical application of Hybrid Mail is to provide a Hybrid Mail operator with printing data as well as processing and delivery instructions, and request the operator to secure the print, enveloping and delivery of the physical letters.

The transfer of data to a Hybrid Mail operator requires that the printing data be linked to a number of data items related to the management, production, finishing, etc. of the data to be printed. Such data items secure that all relevant information is accompanying the printing data. Also it will enable the Hybrid Mail operator to automate his processes.

There is a need for a standardised yet flexible way to present the data to the Hybrid Mail operator. This will enable customers and Hybrid Mail operators to have a seamless exchange of information. It will allow makers of applications for document creation (letters, marketing mailing, etc.) and output management from other applications (accounting systems, production management, etc.), to adhere to the same data presentment and to offer the seamless data interchange.

2 Scope

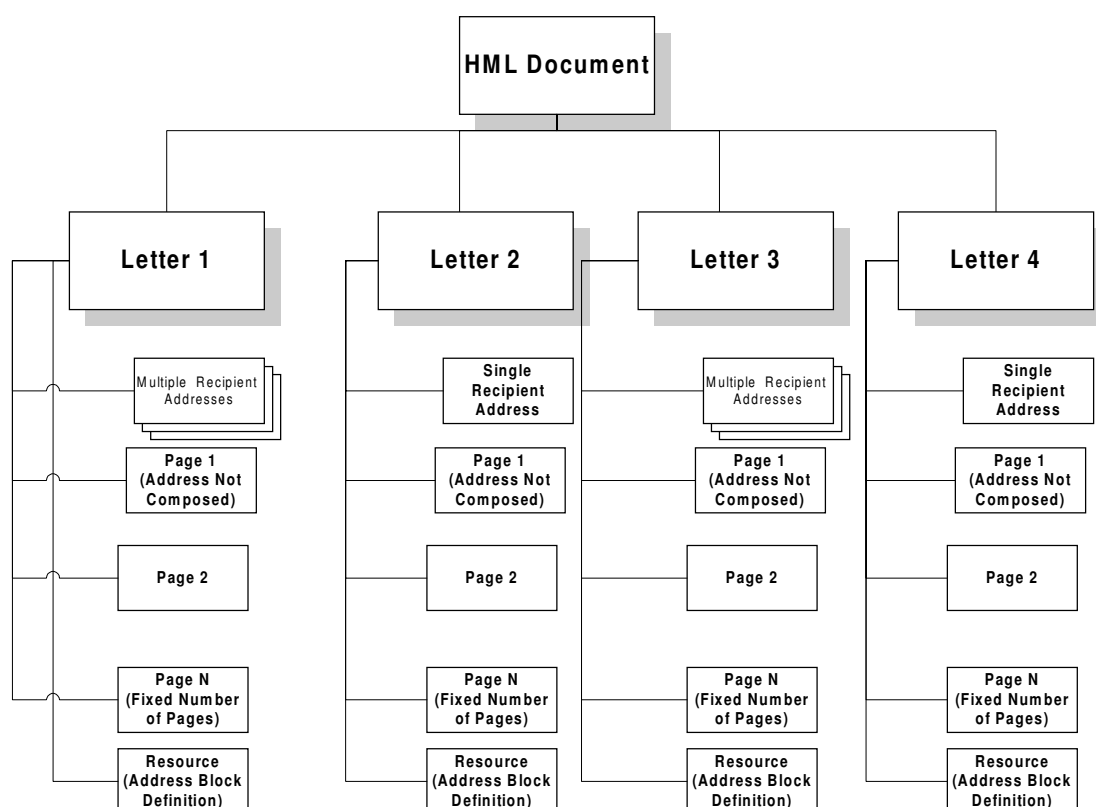
The purpose of this European Prestandard is to define the syntax rules for a data stream for the submission of printing data to a Hybrid mail operator. The Prestandard defines a Document Type Definition (DTD) describing the data stream.

The description is based upon the XML (eXtended Markup Language) definition of rules and semantics for defining a DTD. The purpose of this is to offer a generalised syntax description that can provide seamless integration with a number of existing applications generating data that is liable to be forwarded to a Hybrid Mail operator.

The use of a DTD will ensure that the documents confirm to the standard defined and that the output has the correct syntax. Software manufacturers can use a DTD to program applications that will produce “correct” outputs.

This European Prestandard defines the syntax for creating a data stream that will eventually be converted into a deliverable. The overall object (a batch) can be divided into one or more objects that again can be divided into objects. The hierarchy includes bundles that can hold letters that can hold pages. Each object has a number of characteristics attached to it.

This diagram shows the structure of a HML document: each letter has all the information (except recipient address) fully composed on the page.



Each letter can have one or many recipient addresses. The number of pages for each letter is known (but may be different for each letter). The only external resources are:

- the resource (e.g., CSS, XSL) to define how the recipient address is composed on the page
- the fonts.

Before this file can be printed one action must be performed: each letter must be

replicated as many times as it has recipient addresses and each recipient address must be composed using the external resource.

This European Prestandard does not define the specific services offered by local Hybrid Mail operators.

This European Prestandard does not define the communication method used. It does only define the format of Hybrid Mail as such.

3 Normative references

This European Prestandard incorporates by dated or undated reference, provisions from other publications. These normative references are cited at the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this European Prestandard only when incorporated in it by amendment or revision. For undated references the latest edition of the publication referred to apply (including amendments).

ISO 269	Correspondence envelopes - Designation and sizes
ISO 3166-2	Codes for the representation of names of countries and their subdivisions - Part 1: Country Subdivision codes
ISO 8879:1986	Information processing - Text and office systems - Standard Generalized Markup Language (SGML)
	Technical corrigendum 1:1996
	Amendment 1:1988
ISO/IEC 10646-1	Information technology - Universal Multiple-Octet Coded Character Set (UCS) - Part 1: Architecture and Basic Multilingual Plane
	Technical Corrigendum 1:1996
	Technical Corrigendum 2:1998
	Amendment 1:1996 Transformation Format for 16 planes of group 00 (UTF-16)
	Amendment 2:1996 UCS Transformation Format 8 (UTF-8)
	Amendment 3:1996
	Amendment 4:1996
	Amendment 5:1998 Hangul syllables
	Amendment 6:1997 Tibetan
	Amendment 7:1997 33 additional characters
	Amendment 8:1997
	Amendment 9:1997 Identifiers for characters
	Amendment 10:1998 Ethiopic
	Amendment 11:1998 Unified Canadian Aboriginal Syllables
	Amendment 12:1998 Cherokee
	Amendment 13:1998 CJK unified ideographs with supplementary sources
	Amendment 16:1998 Braille patterns
	Amendment 19:1998 Runic
	Amendment 20:1998 Ogham

4 Terms, definitions and acronyms

4.1 Terms and definitions

For the purposes of this European Prestandard, the following terms and definitions apply.

4.1.1

batch

data structure that contains **bundles** as well as administrative and other data common to all **bundles**

NOTE One HML document will contain one batch. A batch may contain one or several bundles.

4.1.2

bundle

data structure that contains **letters** which are processed as a group as well as administrative and other data common to these **letters**

NOTE A bundle may contain one or more letters.

4.1.3

letter

data structure that contains the data to be rendered as one integral piece of information which is to be delivered to one recipient in physical or electronic format

4.1.4

page

data structure that contains the data to be rendered within one physical or logical space whose dimensions correspond to the relevant presentation method and medium

4.1.5

resource

data structure that contains static data or formatting information for **letters** and/or **pages**

NOTE The static data or information contained in a resource may be relevant to only one letter or page or may be shared between several letters or pages.

4.2 Acronyms

The following acronyms are used in the document:

4.2.1

AFP

Advance Function Presentation – PDL defined by IBM

4.2.2

ASCII

American Standard for Character Information Interchange – character encoding

4.2.3

BDW

Block Descriptor Word

4.2.4

CMYK

Cyan Magenta Yellow Black – colour space definition

4.2.5

CRC

Cyclical Redundancy Check

4.2.6

DES

Data Encryption Standard – encryption standard

4.2.7

DTD

Document Type Definition – part of the XML and SGML standard

4.2.8

ePDS:

ePOST/DS as defined by IDP A/S

4.2.9

FNT

Font object – Xerox – file extension

4.2.10

FRM

Form object – Xerox – file extension

4.2.11

GIF

Graphical Interchange Format – image format

4.2.12

HML

Hybrid Mail Language

4.2.13

HTML

Hyperlink Text Markup Language

4.2.14

IMG

Image object (image format) – Xerox – file extension

4.2.15

IOCA

Image Object Content Architecture – IBM image format

4.2.16

JPEG

Joint Photo Group – image format

4.2.17

LGO

Logo object – Xerox – file extension

4.2.18

LZ77

Lempel-Ziv 1977 – compression technique

4.2.19

LZ78

Lempel-Ziv 1978 – compression technique

4.2.20

LZW

Lempel-Ziv-Welch – compression technique

4.2.21

PCL

Print Control Language – PDL defined by HP

4.2.22

PDF

Portable Document Format – PDL defined by Adobe

4.2.23

PDL

Print Description Language

4.2.24**PGP**

Pretty Good Privacy – encryption standard

4.2.25**PI**

Processing Instruction – part of the XML standard

4.2.26**PNG**

Portable Network Graphics – public domain graphics format

4.2.27**PS1-3**

PostScript level 1 to 3 – PDL defined by Adobe

4.2.28**PSEG**

Page Segment – IBM image format

4.2.29**RDF**

Resource Description Frame

4.2.30**RDW**

Record Descriptor Word

4.2.31**RGB**

Red Green Blue – colour space definition

4.2.32**SGML**

Standard Generalized Markup Language

4.2.33**TIFF**

Tagged Image File Format – image format

4.2.34**URL**

Universal Resource Locator

4.2.35

W3C

World Wide Web Consortium – see <http://www.w3.org/>

4.2.36

XML

eXtensible Markup Language

4.2.37

XSL

eXtensible Stylesheet Language

5 Meta-syntax

This section introduces a syntactic notation, later used in the present document. The notation is adopted to define the syntactic rules of the target language: in this sense, the notation is a meta-syntax for the syntax of the target language.

HML is based on XML version 1.0 as described in REC-xml-19980210 [XML-1998]. This is a subset of Standard Generalized Mark up Language (SGML) as defined in ISO 8879:1986(E).

For the sake of generality, in the following the term **target language** will be used for specifying the language to be defined and to be later used for writing documents.

In the present document the target language is **HML**.

Syntactic rules of the target language are defined by means of syntactic clauses, classified as either **element declarations**, **attribute list declarations** or **comments**. In the following, the first two of these syntactic clauses will be described in detail. Here, only their abstract characteristics are introduced.

Element declarations define **elements**, which are logical parts of the documents.

Elements may contain other elements, to be considered as sub-parts of the first ones. To describe this relationship among elements, element declarations can define elements in terms of other elements.

The whole document itself is considered as an element, and is described by an element declaration: this element is the unique root, which all the other elements start from.

On the other side, elements not further subdivided in parts are simply streams of characters allowed in the documents. They are defined as well by element declarations.

In this way, element declarations express the inner structure of documents: this structure can be easily reconstructed by going through the chain of element declarations, starting from the one that declares the root element.

An element declaration is defined by the following syntactical construction:

<!ELEMENT *element_name* *contentspec* >

Where *element_name* is the name of the element defined and *contentspec* is the list of elements which constitute the set of elements which defines the named element.

Attribute list declarations define characteristics of elements.

In describing the notation, some rules will be followed.

In the syntactic clauses:

- Syntactic items independent of the particular target language (i.e. keywords, symbols and so on) are written in regular font (that is, without using bold or italic forms)
- Identifiers used as placeholders for other things to be later made actual in the syntax of the target language (as for example syntactic items dependent of the particular target language) are written in italic font.

In the examples:

- Syntactic items independent of the particular target language (i.e. keywords, symbols and so on) are written in regular font (that is, without using bold or italic forms)
- Syntactic items dependent of the particular target language (as for example constant names of the target language) are written in italic font.

An attribute declaration is defined by the following syntactical construction:

<!ATTLIST *element_name attribute_name attribute_type default_decl* >

Where the *element_name* identifies which element the attributes belong to, *attribute_name* is the name of the attribute, *attribute_type* is the type of the attribute and *default_decl* informs whether the attribute has a default value that is used if the attribute is not present.

For a more detailed description of the syntax of XML please see [XML-1998]

6 Definition of the HML DTD

In this chapter, syntactic rules of HML are given in XML. They completely define the concrete syntax of HML. The structure of the HML DTD is illustrated in Figure 1 and Figure 2:



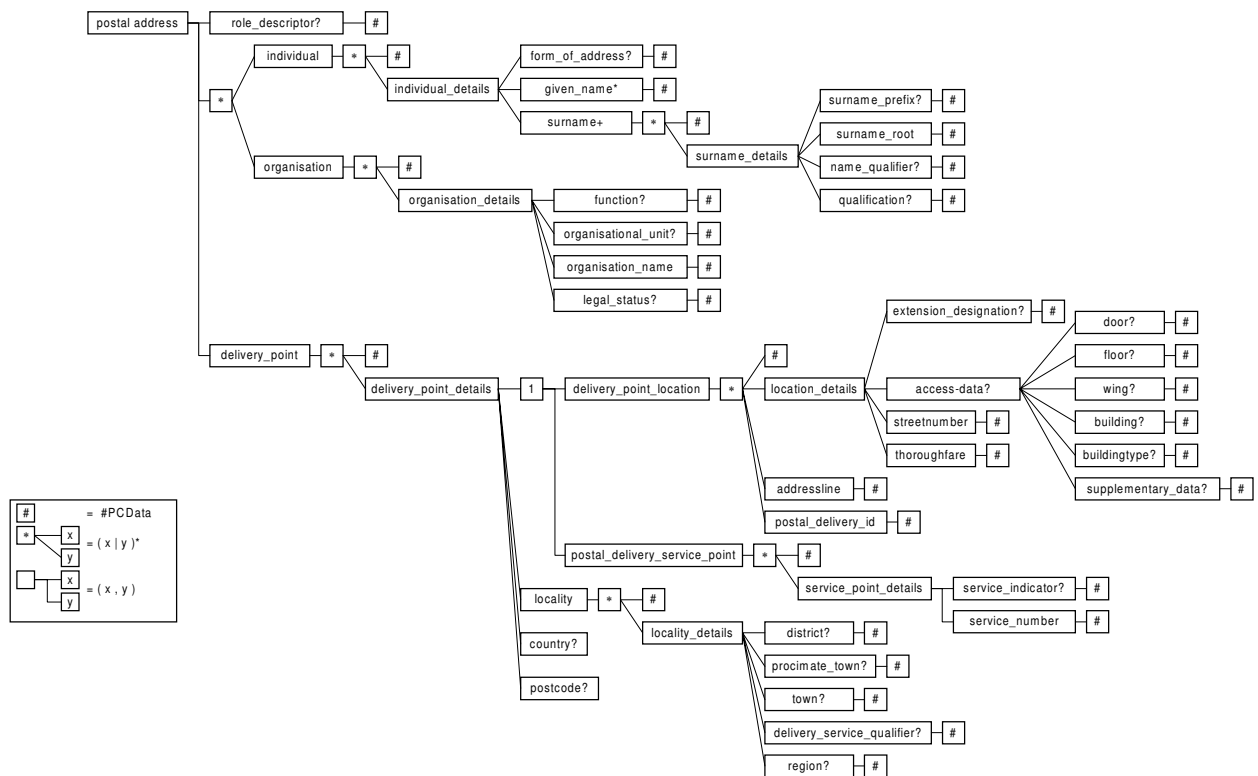


Figure 2 - Illustration of the second part of the HML DD

6.1 General rule for HML documents

HML is based on eXtensible Markup Language (XML) version 1.0 as described in REC-xml-19980210 [XML-1998]. This is a subset of Standard Generalized Markup Language (SGML) as defined in ISO 8879:1986(E).

As HML is specified as a Document Type Definition (DTD) instance of XML, the general rules for XML compliant DTD's apply to HML. However this chapter specifies how a compliant HML document shall be processed when using the following components from the XML standard which are not unique to HML:

- Comments
- Processing Instructions
- CDATA Sections
- Name Spaces

6.1.1 Comments

Comment ::= '<!--' ((Char - '-') | ('-' (Char - '-')))* '-->'

A conforming HML system shall make the comments in a conforming HML document available to external processors on request.

6.1.2 Processing Instructions

PI ::= '<?' PITarget (S (Char* - (Char* '?>' Char*)))? '?>'

PITarget ::= Name - (('X' | 'x') ('M' | 'm') ('L' | 'l'))

A conforming HML system shall ensure that the processing instructions are presented to the relevant processor if this is available. This can either be pre- or postprocessor. Otherwise the processing instructions shall be ignored.

6.1.3 CDATA Sections

```

CDSECT ::= CDStart CData CEnd

CDStart ::= '<![CDATA['

CData ::= (Char* - (Char* ']]>' Char*))

CEnd ::= ']]>'

```

A conforming HML system shall ensure that the data inside a CDATA section is passed through unmodified to the relevant processor.

6.1.4 Name space

The use of name spaces in HML shall comply with the W3C recommendations – see [REC-xml-names-19990114]. The processing of names from other name spaces than the HML name space shall be accepted and either be handled by the HML system or alternatively by a pre- or postprocessor.

6.1.5 Hybrid Mail Language (HML) Extensions:

In the future new equipment must be controllable in Hybrid Mail Operations. To include new features to make use of this equipment and to fulfil the Hybrid Mail Language it is allowed to include new elements.

The way of doing this is make use of the defined way as mentioned in Chapter 2.8 of the [XML-1998]: the document type declaration can point to an external subset containing mark-up declarations.

6.2 General elements

The encoding for any text strings in this Prestandard is the ISO/IEC 10646-1. This will have to be specified in the XML header section of the HML document. Text strings with alternative encoding have to be enclosed in a CDATA section.

There are three types of processing that can be applied within a compliant HML document:

- Pre-processing
- Processing
- Post-processing

Pre-processing is anything that is to be processed before the document is handled by an HML system. Processing is anything that should be handled by a HML system. Post-processing is anything that shall be handled by the post-processing equipment. <crypt>, <compress> and <authenticate> are pre-processing tags.

6.2.1 hml

<!ELEMENT hml (crypt | authenticate | compress | (preprocessing*, common, bundle+))>

The root level tag defines the document.

<!ATTLIST hml

version

CDATA

#REQUIRED >

6.2.1.1 version

In order for a Hybrid Mail processor to process documents that adhere to the HML standard while ensuring that the HML standard can be updated at a later stage as Hybrid Mail processing systems develop any compliant Hybrid Mail processor has to understand the version attribute. The current legal values to the version attribute are:

- 1.0

6.2.2 crypt

<!ELEMENT crypt (#PCDATA) >

The crypt section is to inform any pre-processor of the crypting algorithm used to crypt the HML data with. This applies to the CDATA section. The decrypted contents of the CDATA section will contain the remaining part of the HML document. After the pre-processing has decrypted the CDATA section the HML document is re-phased.

<!ATTLIST crypt

algorithm	CDATA	#REQUIRED
sender	CDATA	#REQUIRED
encryptionkey_id	CDATA	#IMPLIED
encryptionkey_length	CDATA	#IMPLIED>

6.2.2.1 algorithm

The algorithm attribute tells which crypt algorithm has been used to encrypt the data contained in the CDATA section. By default a conforming HML system shall recognise the following values in the algorithm attribute:

- DES
- 3DES
- IDEA
- RSA
- PGP

6.2.2.2 sender

The sender attribute tells who sent the encrypted HML document in order to enable the pre-processor to pick up which decryption key to use.

6.2.2.3 encryptionkey_id

The encryptionkey_id tells which encryption key is used for the encryption of the HML document.

6.2.2.4 encryptionkey_length

The encryptionkey_length tells what the length is of the encryption key. See Appendix C.1. for an example.

NOTE The exchange of encryption key pairs is no part of this specification. This will have to be done separate from the exchange of HML documents.

6.2.3 compress

<!ELEMENT compress (#PCDATA) >

The compress section is to inform any pre-processor of the compress algorithm used to compress the HML data with. This applies to the CDATA section. The decompressed contents of the CDATA section will contain the remaining part of the HML document. After the pre-processing has decompressed the CDATA section the