INTERNATIONAL STANDARD

**ISO/IEC 15444-6**

First edition
2003-10-15

**AMENDMENT 1**
2007-08-15

# Information technology — JPEG 2000 image coding system —

## Part 6: Compound image file format

## AMENDMENT 1: Hidden text metadata

*Technologies de l'information — Système de codage d'image JPEG 2000 —*

*Partie 6: Format de fichier d'image de composant*

*AMENDEMENT 1: Métadonnées de texte caché*

Reference number
ISO/IEC 15444-6:2003/Amd.1:2007(E)

© ISO/IEC 2007

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS

Not for Resale

**PDF disclaimer**

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.



## COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2007

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland



Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

© ISO/IEC 2007 – All rights reserved

# Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Amendment 1 to ISO/IEC 15444-6:2003 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

--`,,```,,,,````-`-`,,`,,`,`,,`---

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

# Information technology — JPEG 2000 image coding system —

## Part 6:
## Compound image file format

## AMENDMENT 1: Hidden text metadata

*Add the following normative references to 2.2:*

IETF RFC 1950, *ZLIB Compressed Data Format Specification version 3.3*, May 1996

IETF RFC 1951, *DEFLATE Compressed Data Format Specification version 1.3*, May 1996

IETF RFC 2045, *Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies*

IETF RFC 2396, *Uniform Resource Identifiers (URI): Generic Syntax,* August 1998

W3C, *Cascading Style Sheets, level 1 (CSS1) Specification*, http://www.w3.org/pub/WWW/TR/REC-CSS1

W3C, *Cascading Style Sheets, level 2 (CSS2) Specification*, http://www.w3.org/TR/REC-CSS2

W3C, *HTML 4.01 Specification,* http://www.w3.org/TR/html401

W3C, *XHTML 1.0 Extensible HyperText Markup Language, Second Edition*, http://www.w3.org/TR/xhtml1

W3C, *XML Schema Part 0: Primer, Second Edition*, http://www.w3.org/TR/xmlschema-0

W3C, *XML Schema Part 1: Structures, Second Edition*, http://www.w3.org/TR/xmlschema-1

W3C, *XML Schema Part 2: Datatypes, Second Edition*, http://www.w3.org/TR/xmlschema-2

*Add the following terms and definitions to Clause 3:*

**3.23**
**hidden text**
symbolic representation for the characters and words found in an image

**3.24**
**annotation**
particular region of a page in a JPM document that has associated a URL reference, a note or a highlight

**3.25**
**hidden text XML**
XML data which describe hidden text and annotations for a single page in a JPM file and which conform to the schema in Annex H

**3.26**
**compressed hidden text XML**
hidden text XML data compressed using the mechanisms defined in F.2

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

**3.27**
**hidden text UUID box**
UUID box containing compressed hidden text XML

**3.28**
**hidden text XML Schema**
XML Schema for hidden text XML, as defined in H.1

*Add the following abbreviations to Clause 4:*

**HTX** Hidden Text XML

*Add the following subclause after 5.2.8:*

## 5.3 Hidden Text Metadata

Hidden text metadata is data representing the text, text elements and text flow associated with an image. In the context of this standard, hidden text is associated with a particular region of a page in a JPM document.

Common uses for hidden text include text searching and highlighting, cut-and-paste, and text-to-speech processing. Hidden text describes the flow of the text on a page as well as the text elements.

JPM allows a rich, multiple content-type representation of a document. Each region of a page may be encoded with a compression technique best suited to its characteristics. In regions containing text, high fidelity reproduction of the source image is retained by not replacing the text regions with a character-based rendition through OCR, but rather by using advanced coding methods such as JBIG2. Even OCR results with a 99 percent accuracy contain substantial numbers of errors per page which require expensive human labour to correct. The searchable nature of a character-based rendition can be obtained instead by associating hidden "dirty OCR" results with the corresponding text image. This standard defines a format for hidden text metadata.

A key issue with hidden text is capturing the ambiguities seen by the OCR engine in a way that allows properly-constructed search engines to find whether and where a given word might be present in a text image. Properly captured, this information provides nearly as much searching precision as an approach using human-corrected "clean OCR" data, but at much lower cost. Search results are most useful where there are fewer false positives to weed through. Intelligent search engines can take account of such data as confidence and alternate characters or alternate words to appropriately alter the ranking of search hits on less certain characters.

In many cases, true ambiguity exists in the image and it would confuse a human observer as well. In these cases, saving confidence values for characters and their alternatives or describing several alternative parsings of a string of characters into words can amount to saving the state of the OCR process to allow the problem to be revisited in a later stage, perhaps by a different engine or by access to first a general dictionary and then a set of more specialized dictionaries.

As a last step, when a person is presented with the search results, they can dismiss a given search hit by comparison to the actual image data for a character or word. For this purpose (and to allow later-stage OCR processes to resume analysis on the image), bounding box rectangles can be defined for all the elements of the hidden text such as characters, words, lines, paragraphs and regions. By indicating a container relationship among these items, intelligent navigation and text selection can occur at character, word, line, paragraph boundaries. A reading order through these rectangles can be defined for what was in the image just a random placement of unrelated glyphs.

While it is primarily designed for use by machines such as search engines, the hidden text can also serve as a crude (if "dirty") or adequate (if "clean") alternate representation for an image region to allow it to display on character-based devices (such as mobile phones) or small-area graphics devices (such as PDAs).


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Annotations are added to the document typically with a WYSIWYG editor to indicate URL references, notes, and to highlight key sections of the document text. Each annotation is associated with a particular region of a page in a JPM document.

XML is used for hidden text and annotations because it is a format widely used to store structured information, and can be machine processed.

*Renumber the original 5.3 as 5.4.*

*Add the following rows at the correct alphabetical location in Table A.1 of A.4:*

**Table A.1 — Boxes defined or referenced within this International Standard**

| Box name | Type | Superbox | Comments (Informative) |
|---|---|---|---|
| Hidden Text Metadata | 'htxb' (0x68747862) | Yes | This optional box contains hidden text and annotations. |
| HTX Reference Box | 'phtx' (0x70687478) | No | This optional box can be used to point to Hidden Text Metadata box contents at top file level. |

*Add the following subclauses after B.6.4:*

### B.6.5 Hidden Text Metadata box (superbox)

Box type: 'htxb' (0x68747862)

Container: Page box or File

Mandatory: No

Quantity: At most one if the container is the Page box, any number if the container is the file

Location: Anywhere in the Page box after the Page Header box if the container is the Page box, or anywhere after the File Type box if the container is the file

The Hidden Text Metadata box ('htxb') serves as a container for hidden text data. It is a superbox that may contain an optional Label box and must contain one of two box types. It may either contain one XML box containing hidden text metadata, or it may contain one UUID box containing hidden text metadata as specified in F.2.

The type of a Hidden Text Metadata box shall be htxb' (0x68747862). The contents of a Hidden Text Metadata box shall be as in Figure B.25:

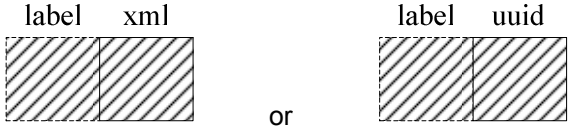



**Figure B.25 — Organization of the contents of a Hidden Text Metadata box**

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### B.6.6 HTX Reference box

Box type: 'phtx' (0x70687478)

Container: Page box

Mandatory: No

Quantity: At most one

Location: Anywhere in the Page box after the Page Header box

If the hidden text for a page is contained in a Hidden Text Metadata box within the corresponding Page box, this box must not appear. If the hidden text for a page is contained in a series of one or more Hidden Text Metadata boxes at the file level, one HTX reference box has to be included in the corresponding Page box.

The type of a HTX Reference box shall be 'phtx' (0x70687478). The contents of a HTX Reference box shall be as in Figure B.26:

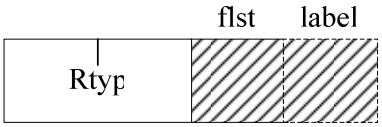



**Figure B.26 — Organization of the contents of a HTX Reference box**

**Rtyp**: Referenced box type. This field specifies the actual type (as would be found in the TBox field in an actual box header) of the box referenced by this HTX Reference box. However, a reader shall not attempt to locate a physically stored box header for the box represented by this HTX Reference box, as it is legal to use a HTX Reference box to create a new box that is not contiguously contained in other locations within this or other files, and thus the box header will not exist.

**flst**: Fragment List box. This box specifies the actual locations of the fragments of the referenced HTX element. When those fragments are concatenated, in order, as specified by the Fragment List box definition, the resulting byte-stream shall be the contents of the referenced HTX element, which contains hidden text data, and shall not include the box header fields. The format of the Fragment List box is specified in B.5.1.1. If Rtyp is 'uuid' and the UUID signals deflate compression as defined in F.2, the number of fragments of the Fragment List box must be one.

**label**: Label box. This optional box may contain a Label box which specifies a label or name for the hidden text of the corresponding page. The structure of a Label box is specified in B.6.3.

**Table B.31 — HTX Reference box contents data structure values**

| Parameter | Size (bits) | Value |
|---|---|---|
| Rtyp | 32 | See Table B.32 |
| flst | Variable | Variable |
| label | Variable | Variable |

**Table B.32 — Legal Rtyp values**

| Value | Meaning |
|---|---|
| xml\40 | The referenced HTX data shall be contained in an XML box as described in Annex F. The XML box is defined in I.7.1 of ITU-T Rec T.800 (2002) \| ISO/IEC 15444-1:2004. |
| uuid | The referenced HTX data shall be contained in a UUID box as described in Annex F. The UUID box is defined in I.7.2 of ITU-T Rec T.800 (2002) \| ISO/IEC 15444-1:2004. |
| | All other values reserved |


Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
© ISO/IEC 2007 – All rights reserved
Not for Resale

*Renumber the original B.6.5 as B.6.7.*

*Add the following annexes after Annex E:*

# Annex F
(normative)

# Hidden Text and Annotations Storage

## F.1 Storage of HTX in JPM

A hidden text XML element is restricted to represent text for a single page. It is stored in a Hidden Text Metadata box as defined in B.6.5. The Hidden Text Metadata box either appears within the corresponding Page box or is placed at the top level of the file. If placed on top level, an HTX Reference box as defined in B.6.6 must be placed in the corresponding Page box to point to the Hidden Text Metadata boxes that composes the hidden text of the page.

When a Hidden Text Metadata box is small in size, it is reasonable to place it directly in Page box. In keeping with the usual JPM approach, large objects are generally placed at the top file level. In this case, the much smaller HTX Reference box is placed in the page box and points to the actual data. Also in this case a single HTX Reference box can point to multiple file level Hidden Text Metadata boxes. This can be used to compose the HTX for many pages from combinations of fixed page content (such as page headers and footers) and variable page content unique to each page.

XML data representing hidden text and annotations is defined using XML 1.0, and conforms to the schemas in Annex H. It shall be referred to as ***Hidden Text XML or HTX***.

HTX shall be stored in a Hidden Text Metadata box as defined in B.6.5.

The storage of uncompressed HTX may increase file size considerably. In order to minimise the increase in file size, HTX may be compressed using the mechanisms defined in F.2.

## F.2 Compression of HTX

HTX may be compressed using the zlib format defined in IETF RFC 1950 with DEFLATE compression defined in IETF RFC 1951.

UUID boxes shall be used for the storage of compressed HTX in the JPM file format.

Compressed HTX shall be stored in a UUID box, as defined in I.7.3 of ISO/IEC 15444-1:2004, with the following contents:

**ID** This field shall contain the following 16 hexadecimal bytes:

```
c2 f3 66 a4 27 ec 40 c4 a0 9a 7e 65 2f 36 eb 59
```

**DATA** This field will contain hidden text XML compressed to the DEFLATE format, as specified in F.1.

A UUID box with the above content shall be referred to as a ***hidden text UUID box***.

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

The following URL may be used in a UUID Data Entry box, as defined in I.7.3.2 of ISO/IEC 15444-1:2004, to describe the format of the data contained in hidden text UUID boxes:

http://www.jpeg.org/hiddentext/htx.html


Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
© ISO/IEC 2007 – All rights reserved
Not for Resale

# Annex G
(normative)

# Hidden Text and Annotations Types and Elements

## G.1 Overview

This section describes each of the HTX types and elements, and how they are to be used and interpreted.

Annex H formally describes the schemas that the hidden text XML must conform to. Here the text is a description of each of the elements, what they are for, how they relate to each other, how often they can occur, how they are to be interpreted.

Hidden text can be encoded using subelements at different levels of detail as described in this section. This can be used to structure the hidden text and give it a text flow in regions, paragraphs, lines, words, etc. Whenever this kind of structured information is not available, the hidden text can be directly put into the appropriate elements, omitting specific positioning of lines inside paragraphs, words inside lines, etc. The following picture gives an overview of the various elements that can be used to store the hidden text of a page:

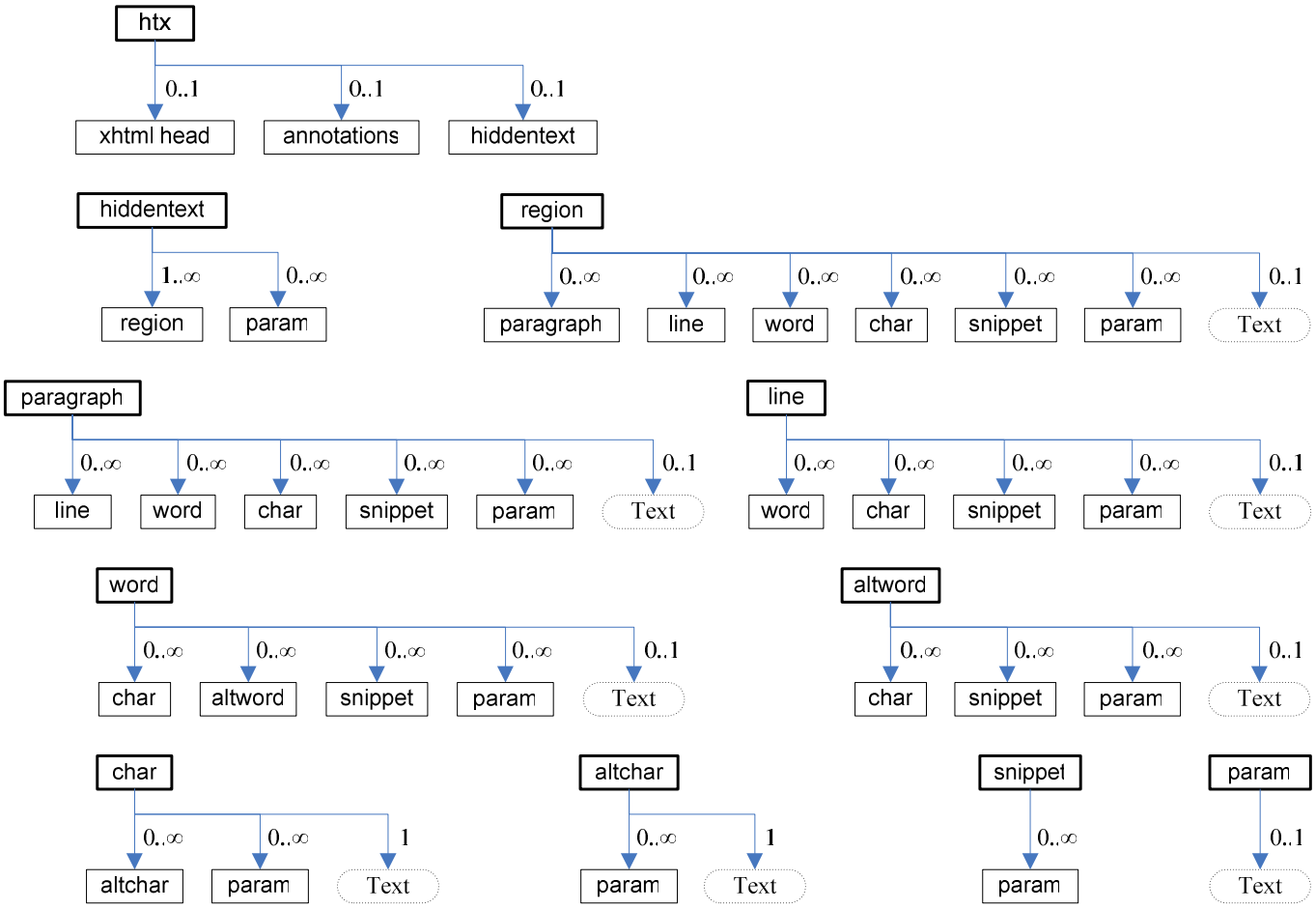



**Figure G.1 — Structure of HTX**

© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

The hidden text XML schema (see H.1) uses some types and elements defined in the XHTML 1.0 XML Schema. See the XHTML 1.0 reference for full details of these types and elements.

The following additional types and elements are defined:

## G.2 Types

### G.2.1 Shape

The **Shape** type is used to describe the shape of a region in the document and is defined by the following XML schema declaration:

```
<xs:simpleType name="Shape">
   <xs:annotation>
         <xs:documentation>Enumeration of shapes.</xs:documentation>
  </xs:annotation>
    <xs:restriction base="xs:token">
        <xs:enumeration value="rect"/>
        <xs:enumeration value="poly"/>
    </xs:restriction>
</xs:simpleType>
```

### G.2.2 Coordinates

The **Coords** type is used to store a comma separated sequence of non-negative integer values. This type is similar to the XHTML 1.0 **Coords** type but excludes negative and percentage values. The attribute specifies the position and shape of the area. The number and order of values depends on the value of the shape attribute. Possible combinations:

- **rect**: left-x, top-y, right-x, bottom-y.
- **poly**: x1, y1, x2, y2, ..., xN, yN.
  If the first and last x and y coordinate pairs are not the same, user agents must infer an additional coordinate pair to close the polygon.

The **Coords** element is defined by the following XML schema declaration:

```
<xs:simpleType name="Coords">
   <xs:annotation>
      <xs:documentation>
          Comma separated list of integer values.
       </xs:documentation>
   </xs:annotation>
   <xs:restriction base="xs:string">
      <xs:pattern value="\d+,\s*\d+(,\s*\d+,\s*\d+)+"/>
   </xs:restriction>
</xs:simpleType>
```


Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
© ISO/IEC 2007 – All rights reserved
Not for Resale

### G.2.3 Percentage

A simple type **Percentage** is defined to store a string that holds a percent value indicating the confidence of a hidden text word or character match. **Percentage** is defined as follows:

```
<xs:simpleType name="Percentage">
   <xs:annotation>
      <xs:documentation>Percentage value.</xs:documentation>
   </xs:annotation>
   <xs:restriction base="xs:string">
      <xs:pattern value="\d+(\.\d+)?%?"/>
   </xs:restriction>
</xs:simpleType>
```

### G.2.4 Angle

A simple type **Angle** is defined to store a string that indicates an angle for use in hidden text. The **Angle** type is defined as follows:

```
<xs:simpleType name="Angle">
    <xs:annotation>
        <xs:documentation>
            nn for radian measure or nn° for degree
        </xs:documentation>
    </xs:annotation>
    <xs:restriction base="xs:string">
        <xs:pattern value="[\-\+]?\d+(\.\d+)?°?"/>
    </xs:restriction>
</xs:simpleType>
```

### G.2.5 Resolution

A simple type **Resolution** is defined to store a string that indicates a resolution for use with coordinates in hidden text and annotations. The **Resolution** type is defined as follows:

```
<xs:simpleType name="Resolution">
    <xs:annotation>
        <xs:documentation>
            Resolution value in dots per inch (dpi). A single number stands
for
            horizontal and vertical resolution having the same values.
            Two numbers can be used to define different resolutions for
horizontal
            (first number) and vertical (second number).
        </xs:documentation>
    </xs:annotation>
    <xs:restriction base="xs:string">
        <xs:pattern value="\d+(\.\d+)?(,\*s\d+(\.\d+)?)?"/>
    </xs:restriction>
</xs:simpleType>
```

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

## G.3 Common Attributes

### G.3.1 Core Attributes

**Coreattrs**, a set of core attributes that are common to most elements, is defined as follows:

```
<xs:attributeGroup name="coreattrs">
    <xs:annotation>
        <xs:documentation>
            core attributes common to most elements
            id          document-wide unique id
            class       space separated list of classes
            lang        language code (backwards compatible)
            xml:lang    language code (as per XML 1.0 spec)
            dir         direction for weak/neutral text
            iref        URI of the image corresponding to the region
        </xs:documentation>
    </xs:annotation>
    <xs:attribute name="id"    type="xs:ID"       />
    <xs:attribute name="class" type="xs:NMTOKENS" />
    <xs:attribute name="iref"  type="xs:anyURI" />
    <xs:attributeGroup ref="xhtml:i18n"/>
</xs:attributeGroup>
```

The following attributes are members of the **Coreattrs** group:

- **lang** (optional)

  An optional attribute of type **LanguageCode** to indicate the default language for text in the hidden text XML. Refer to the XHTML 1.0 specification for further details.

- **xml:lang** (optional)

  An optional attribute of type **xml:lang** to indicate the default language for text in the hidden text XML. Refer to the XHTML 1.0 specification for further details.

- **dir** (optional)

  An optional attribute containing the string **rtl** or **ltr**, indicating the default direction for text in the hidden text XML. Refer to the XHTML 1.0 specification for further details.

- **id** (optional)

  An optional attribute of type xs:ID. Contains an id that is unique in the scope of this document. This attribute can be used for referencing a certain element (e.g. in a style sheet). See XML Schema specification for further details.

- **class** (optional)

  This attribute can contain a space separated list of classes. Useful for convenient style sheet usage.

- **iref** (optional)

  URI which points to an image file corresponding to the region.
  (ex.1 iref="http://jpeg.org/image.jp2", ex.2 iref="jpip://jpeg.org/image.jp2?fsize=32,32&rsiz=32,32")


Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
© ISO/IEC 2007 – All rights reserved
Not for Resale

### G.3.2 Position Attributes

**Posattrs**, a set of position attributes that are common to most visual elements, is defined as follows:

```
<xs:attributeGroup name="posattrs">
    <xs:annotation>
        <xs:documentation>
            positioning attributes common to most elements
            shape       shape of an element
            coords      coordinates of an element
            angle       angle of text direction
                        0 is horizontal to the right, positive values
                        mean counter-clockwise rotation
            baseline    angle of the characters in a line of
                        text
        </xs:documentation>
    </xs:annotation>
  <xs:attribute name="shape"    type="Shape" default="rect"   />
  <xs:attribute name="coords"   type="Coords"                 />
  <xs:attribute name="angle"    type="Angle" default="0"      />
     <xs:attribute name="baseline" type="Angle" default="0"      />
</xs:attributeGroup>
```

The following attributes are members of the **Posattrs** group:

- **shape** (optional)

  An optional attribute of type **Shape** containing the shape of the region bounding the element. Possible values are ‘rect’ for a rectangle and ‘poly’ for a polygon.The default value for this attribute is **rect.** If this attribute is missing then the bounding shape for this element is the bounding shape of the parent element (which is the whole page in case of hiddentext).

- **coords** (optional)

  The logical coordinates of the shape bounding the hidden text for this page. The unit is pixels. A resolution can be defined as an attribute on the htx element. If this attribute is missing then the bounding shape for this element is the bounding shape of the parent element (which is the whole page in case of hiddentext). How the value of cords is to interpreted depends on the shape attribute. The coord values unit is pixel, no percentage or any length unit like inch or centimetre.

  The origin (coordinates ‘0, 0’) is the upper left corner of the page.

- **angle** (optional)

  An attribute of type Angle that indicates the angle of orientation of the element, relative to the direction of the element's parent.

  Can either be in degree (value followed by a ° sign) or radian measure (value without unit). A value of 0 means same direction as the element's parent, positive values mean rotating counter-clockwise relative to that direction. Default value is ‘0’.

- **baseline** (optional)

  An attribute of type Angle that indicates the relative orientation of the sub elements and direct content contained in the element with respect to the direction given by the angle attribute.

  Can either be in degree (value followed by a ° sign) or radian measure (value without unit). A value of 0 means same direction as the element, positive values mean rotating counter-clockwise relative to that direction. Default value is ‘0’.

The values of shape and cords attribute should be interpreted as described in HTML 4.01 subclause 13.6.1 section “AREA attribute definitions”.

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

## G.4 Elements

### G.4.1 HTX

The **htx** element, the global container and root elements for hidden text and annotations, is declared as follows:

```
<xs:element name="htx">
   <xs:annotation>
      <xs:documentation>
           Global container for hidden text and annotations. Contains
           language attributes, an optional xhtml head and a
           mandatory body.
      </xs:documentation>
   </xs:annotation>
   <xs:complexType>
      <xs:sequence>
         <xs:element ref="xhtml:head"    minOccurs="0"/>
         <xs:element ref="annotations"   minOccurs="0"/>
         <xs:element ref="hiddentext"    minOccurs="0"/>
      </xs:sequence>
      <xs:attribute name="res"      type="xs:Resolution" />
      <xs:attribute name="width"    type="xs:integer"    />
      <xs:attribute name="height"   type="xs:integer"    />
      <xs:attributeGroup ref="coreattrs"/>
   </xs:complexType>
</xs:element>
```

This root element contains an optional **xhtml:head** element, a annotations element and a **hiddentext** element.

**Attributes**:

Core attributes apply.

- **res** (optional)

  An optional attribute of type Resolution indicating the resolution for any coordinates in dots per inch (dpi). A single number stands for horizontal and vertical resolution having the same values. Two numbers can be used to define different resolutions for horizontal (first number) and vertical (second number).

- **width** (optional)

  The width of the page in pixels.

- **height** (optional)

  The height of the page in pixels.

**Elements**

- **xhtml:head** (at most one)

  An optional element containing general header data for the hidden text XML elements, including any required Cascading Style Sheet data. Refer to the XHTML 1.0 specification for further details.

- **annotations** (at most one)

  An optional annotations element may be used to attach notes and to describe clickable and highlighted regions on a page.

- **hiddentext** (at most one)

  An optional hiddentext containing the hidden text XML data for a page.

**Direct content**

- **none**


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.2 Parameter

The **param** element is declared as follows:

```
<xs:element name="param">
   <xs:annotation>
      <xs:documentation>
           User defined properties for a hidden text and annotations
           object.
       </xs:documentation>
   </xs:annotation>
   <xs:complexType mixed="true">
      <xs:attribute name="name"  type="xs:string" use="required"/>
   </xs:complexType>
</xs:element>
```

A **param** element of a HTX contains user defined properties for the associated element, e.g. to specify the OCR engine used to generate the hidden text.

**Attributes**:

- **name** (mandatory)

  A mandatory string attribute containing the name of this parameter.

**Elements**

- **none**

**Direct content**

- **value**

  An optional string attribute containing the value for this parameter.

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.3 Hidden Text

The **hiddentext** element is used to represent the hidden text for a page and is declared as follows:

```
<xs:element name="hiddentext">
   <xs:annotation>
      <xs:documentation>
           A hiddentext element contains hidden text for a page.
       </xs:documentation>
   </xs:annotation>
   <xs:complexType>
      <xs:sequence>
         <xs:element ref="param"    minOccurs="0" maxOccurs="unbounded"/>
         <xs:element ref="region"                 maxOccurs="unbounded"/>
      </xs:sequence>
      <xs:attributeGroup ref="coreattrs"   />
      <xs:attributeGroup ref="posattrs"    />
   </xs:complexType>
</xs:element>
```

A **hiddentext** element of a HTX contains a sequence of **region** elements.

**Attributes**:

Core and position attributes apply.

**Elements**

- **param** (zero or more)

  Optional sequence of parameter elements may be used to specify user defined properties for hidden text.

- **region** (one or more)

  Sequence of elements containing hidden text for regions in the page.

**Direct content**

- None.


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.4 Region

The **region** element is declared as follows:

```
<xs:element name="region">
   <xs:annotation>
      <xs:documentation>
           A region element contains hidden text for a page region.
       </xs:documentation>
   </xs:annotation>
   <xs:complexType mixed="true">
      <xs:sequence maxOccurs="unbounded">
         <xs:element ref="param"     minOccurs="0" maxOccurs="unbounded" />
         <xs:element ref="paragraph" minOccurs="0" maxOccurs="unbounded" />
         <xs:element ref="line"      minOccurs="0" maxOccurs="unbounded" />
         <xs:element ref="word"      minOccurs="0" maxOccurs="unbounded" />
         <xs:element ref="char"      minOccurs="0" maxOccurs="unbounded" />
         <xs:element ref="snippet"   minOccurs="0" maxOccurs="unbounded" />
      </xs:sequence>
      <xs:attributeGroup ref="coreattrs" />
      <xs:attributeGroup ref="posattrs"  />
   </xs:complexType>
</xs:element>
```

A **region** element of a HTX contains an optional sequence of **paragraph, line, word, char** and **snippet** elements and optional hidden text for a region.

**Attributes**:

Core and position attributes apply.

**Elements**

- **param** (zero or more)

  Optional sequence of parameter elements may be used to specify user defined properties for the region.

- **paragraph** (zero or more)

  Optional sequence of elements containing hidden text for paragraphs in the region.

- **line** (zero or more)

  Optional sequence of elements containing hidden text for lines in the region

- **word** (zero or more)

  Optional sequence of elements containing hidden text for words in the region

- **char** (zero or more)

  Optional sequence of elements containing hidden text for characters in the region

- **snippet** (zero or more)

  Optional sequence of elements each identifying an unrecognised portion of the raster.

**Direct content**

- Region level hidden text

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.5 Paragraph

The **paragraph** element is declared as follows:

```
<xs:element name="paragraph">
   <xs:annotation>
      <xs:documentation>
           A paragraph element contains hidden text for a paragraph.
       </xs:documentation>
   </xs:annotation>
   <xs:complexType mixed="true">
      <xs:sequence maxOccurs="unbounded">
         <xs:element ref="param"    minOccurs="0" maxOccurs="unbounded"  />
         <xs:element ref="line"     minOccurs="0" maxOccurs="unbounded"  />
         <xs:element ref="word"     minOccurs="0" maxOccurs="unbounded"  />
         <xs:element ref="char"     minOccurs="0" maxOccurs="unbounded"  />
         <xs:element ref="snippet"  minOccurs="0" maxOccurs="unbounded"  />
      </xs:sequence>
      <xs:attributeGroup ref="coreattrs" />
      <xs:attributeGroup ref="posattrs"  />
   </xs:complexType>
</xs:element>
```

A **paragraph** element of a HTX contains an optional sequence of **line, word, char** and **snippet** elements and optional hidden text for the paragraph.

**Attributes**:

Core and position attributes apply.

**Elements**

- **param** (zero or more)

  Optional sequence of parameter elements may be used to specify user defined properties for the paragraph.

- **line** (zero or more)

  Optional sequence of elements containing hidden text for lines in the paragraph.

- **word** (zero or more)

  Optional sequence of elements containing hidden text for words in the paragraph

- **char** (zero or more)

  Optional sequence of elements containing hidden text for characters in the paragraph

- **snippet** (zero or more)

  Optional sequence of elements each identifying an unrecognised portion of the raster.

**Direct content**

- Paragraph level hidden text

--`,,```,,,,````-`-`,,`,,`,`,,`---



© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.6 Line

The **line** element is declared as follows:

```
<xs:element name="line">
   <xs:annotation>
      <xs:documentation>
           A line element contains hidden text for a line.
       </xs:documentation>
   </xs:annotation>
   <xs:complexType mixed="true">
      <xs:sequence maxOccurs="unbounded">
         <xs:element ref="param"   minOccurs="0" maxOccurs="unbounded"  />
         <xs:element ref="word"    minOccurs="0" maxOccurs="unbounded"  />
         <xs:element ref="char"    minOccurs="0" maxOccurs="unbounded"  />
         <xs:element ref="snippet" minOccurs="0" maxOccurs="unbounded"  />
      </xs:sequence>
      <xs:attributeGroup ref="coreattrs" />
      <xs:attributeGroup ref="posattrs"  />
   </xs:complexType>
</xs:element>
```

A **line** element of a HTX contains an optional sequence of **word, char** and **snippet** elements and optional hidden text for a line.

**Attributes**:

Core and position attributes apply.

**Elements**

- **param** (zero or more)

  Optional sequence of parameter elements may be used to specify user defined properties for the line.

- **word** (zero or more)

  Optional sequence of elements containing hidden text for words in the line.

- **char** (zero or more)

  Optional sequence of elements containing hidden text for characters in the line

- **snippet** (zero or more)

  Optional sequence of elements each identifying an unrecognised portion of the raster.

**Direct content**

- Line level hidden text

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.7 Word

The **word** element is declared as follows:

```
<xs:element name="word">
   <xs:annotation>
      <xs:documentation>
           A word element contains hidden text for a word.
       </xs:documentation>
   </xs:annotation>
   <xs:complexType mixed="true">
      <xs:sequence maxOccurs="unbounded">
         <xs:element ref="param"   minOccurs="0" maxOccurs="unbounded"   />
         <xs:element ref="char"    minOccurs="0" maxOccurs="unbounded"   />
         <xs:element ref="snippet" minOccurs="0" maxOccurs="unbounded"   />
         <xs:element ref="altword" minOccurs="0" maxOccurs="unbounded"   />
      </xs:sequence>
      <xs:attribute name="conf"  type="Percentage"   />
      <xs:attributeGroup ref="coreattrs"             />
      <xs:attributeGroup ref="posattrs"              />
   </xs:complexType>
</xs:element>
```

A **word** element of a HTX contains an optional sequence of **char, altword** and **snippet** elements and optional hidden text for a word.

**Attributes**:

Core and position attributes apply.

- **conf** (optional)

  An optional attribute indicating the confidence of the accuracy of the given word as a percentage in the range of 0% to 100%.

**Elements**

- **param** (zero or more)

  Optional sequence of parameter elements may be used to specify user defined properties for the word.

- **char** (zero or more)

  Optional sequence of elements containing hidden text for characters in the word.

- **snippet** (zero or more)

  Optional sequence of elements each identifying an unrecognised portion of the raster.

- **altword** (zero or more)

  Optional sequence of elements containing hidden text for alternate words. It defines a possible alternative interpretation of the word in the raster image when uncertainties exist.

**Direct content**

- Word level hidden text


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.8 Alternative Word

The **altword** element contains alternative hidden text for a word and is declared as follows:

```
<xs:element name="altword">
   <xs:annotation>
      <xs:documentation>
           A altword element contains hidden text for a alternative word.
       </xs:documentation>
   </xs:annotation>
   <xs:complexType mixed="true">
      <xs:sequence maxOccurs="unbounded">
         <xs:element ref="param"    minOccurs="0" maxOccurs="unbounded"  />
         <xs:element ref="char"     minOccurs="0" maxOccurs="unbounded"  />
         <xs:element ref="snippet"  minOccurs="0" maxOccurs="unbounded"  />
      </xs:sequence>
      <xs:attribute name="conf"  type="Percentage"   />
      <xs:attributeGroup ref="coreattrs"   />
      <xs:attributeGroup ref="posattrs"    />
   </xs:complexType>
</xs:element>
```

An **altword** element of a HTX contains an optional sequence of **char** and **snippet** elements, and optional hidden text for an alternative word.

**Attributes**:

Core and position attributes apply.

- **conf** (optional)

  An optional attribute indicating the confidence of the accuracy of the given alternate word as a percentage in the range of 0% to 100%.

**Elements**

- **param** (zero or more)

  Optional sequence of parameter elements may be used to specify user defined properties for the alternate word.

- **char** (zero or more)

  Optional sequence of elements containing hidden text for characters in the alternate word.

- **snippet** (zero or more)

  Optional sequence of elements each identifying an unrecognised portion of the raster.

**Direct content**

- Word level hidden text specifying some alternative text.

© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.9 Character

The **char** element contains a hidden text character and is declared as follows:

```
    <xs:element name="char">
    <xs:annotation>
       <xs:documentation>
            A char element contains a hidden text character.
        </xs:documentation>
    </xs:annotation>
    <xs:complexType mixed="true">
       <xs:sequence maxOccurs="unbounded">
          <xs:element ref="param"    minOccurs="0" maxOccurs="unbounded"  />
          <xs:element ref="altchar"  minOccurs="0" maxOccurs="unbounded"  />
       </xs:sequence>
       <xs:attribute name="conf"  type="Percentage"   />
       <xs:attributeGroup ref="coreattrs"   />
       <xs:attributeGroup ref="posattrs"    />
    </xs:complexType>
</xs:element>
```

A **char** element of a HTX contains an optional sequence of **altchar** elements and an optional hidden text character.

**Attributes**:

Core and position attributes apply.

- **conf** (optional)

  An optional attribute indicating the confidence of the accuracy of the given character as a percentage in the range of 0% to 100%.

**Elements**

- **param** (zero or more)

  Optional sequence of parameter elements may be used to specify user defined properties for the character.

- **altchar** (zero or more)

  Optional sequence of elements containing hidden text for alternate characters in the alternate word. This element may only occur directly after a char or altchar element. It defines a possible alternative interpretation of the character in the raster image when uncertainties exist.

**Direct content**

- Character level hidden text


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.10 Alternative Character

The **altchar** element is declared as follows:

```
<xs:element name="altchar">
   <xs:annotation>
      <xs:documentation>
           A altchar element contains a hidden text alternative character.
       </xs:documentation>
   </xs:annotation>
   <xs:complexType mixed="true">
      <xs:sequence>
         <xs:element ref="param"    minOccurs="0" maxOccurs="unbounded"  />
      </xs:sequence>
      <xs:attribute name="conf"  type="Percentage"    />
      <xs:attributeGroup ref="coreattrs"   />
      <xs:attributeGroup ref="posattrs"    />
   </xs:complexType>
</xs:element>
```

A **altchar** element of a HTX contains an optional sequence of **snippet** elements and optional hidden text for an alternative character.

**Attributes**:

Core and position attributes apply.

- **conf** (optional)

  An optional attribute indicating the confidence of the accuracy of the given alternate character as a percentage in the range of 0% to 100%.

**Elements**

- **param** (zero or more)

  Optional sequence of parameter elements may be used to specify user defined properties for the alternate character.

**Direct content**

- Character level hidden text specifying an alternative character.

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.11 Snippet

The **snippet** element is declared as follows:

```
<xs:element name="snippet">
   <xs:annotation>
      <xs:documentation>
         A snippet element identifies an unrecognised portion of the image.
      </xs:documentation>
   </xs:annotation>
   <xs:complexType mixed="true">
      <xs:sequence minOccurs="0" maxOccurs="unbounded" >
         <xs:element ref="param" />
      </xs:sequence>
      <xs:attributeGroup ref="coreattrs"   />
      <xs:attributeGroup ref="posattrs"    />
   </xs:complexType>
</xs:element>
```

A **snippet** element of a HTX identifies an unrecognised portion of the image.

**Attributes**:

Core and position attributes apply.

**Elements**

- **param** (zero or more)

  Optional sequence of parameter elements may be used to specify user defined properties for the snippet.

**Direct content**

- None.

--`,,```,,,,````-`-`,,`,,`,`,,`---



© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.12 Annotations

The **annotations** element is declared as follows:

```
<xs:element name="annotations">
   <xs:annotation>
      <xs:documentation>
           The annotations element represents annotated, clickable and
           highlighted areas on a page.
      </xs:documentation>
   </xs:annotation>
   <xs:complexType>
      <xs:sequence maxOccurs="unbounded">
         <xs:element ref="area"/>
      </xs:sequence>
      <xs:attributeGroup ref="coreattrs" />
   </xs:complexType>
</xs:element>
```

An annotation element of a HTX contains a sequence of area elements defining clickable and/or highlighted areas as well as notes for a page.

**Attributes**:

Core attributes apply.

**Elements**

- **area** (at least one)

  Sequence of elements defining clickable and highlighted areas on a page.

**Direct content**

- **none**

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

### G.4.13 Area

The **area** element is declared as follows:

```
<xs:element name="area">
   <xs:annotation>
      <xs:documentation>
            An area element represents a clickable and/or highlighted area on a
            page.
      </xs:documentation>
   </xs:annotation>
   <xs:complexType>
      <xs:attribute name="href"  type="xhtml:URI"/>
      <xs:attribute name="alt"   type="xhtml:Text"/>
      <xs:attribute name="target"   type="xhtml:URI" default="_self"/>
      <xs:attribute name="shape"    type="Shape" default="rect"   />
      <xs:attribute name="coords" type="Coords"      />
      <xs:attributeGroup ref="coreattrs" />
   </xs:complexType>
</xs:element>
```

An **area** element of a HTX defines a clickable and/or highlighted area on a page. To define a highlighted, non-clickable area the **href** attribute is to be omitted.

**Attributes**:

Core and position attributes apply.

- **href** (optional)

  URI of the hyperlink to use when the area is clicked.

- **alt** (optional)

  An optional text descriptor for the hyperlink, possibly for use as a tooltip. Must not be used without the **href** attribute defined.

- **target** (optional)

  The frame that the document the hyperlink points to should be displayed in (if the viewing program supports any kind of framing). Default is "_self". See HTML 4.01 definition for further details. Must not be used without the **href** attribute defined.

**Elements**

- **none**

**Direct content**

- Notes for this area.

--`,,```,,,,````-`-`,,`,,`,`,,`---



© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

# Annex H
(normative)

# Hidden Text and Annotations Schema

## H.1 XML Schema

The following URL references the XML Schema for HTX:

**http://www.jpeg.org/hiddentext/htx.xsd**

This XML Schema shall be referred to as the ***Hidden Text XML Schema***.

### H.1.1 Version and Encoding

The HTX Schema uses XML 1.0 and UTF-8 encoding:

```
<?xml version="1.0" encoding="utf-8" ?>
```

### H.1.2 Schema

The schema wrapper for HTX is declared as follows:

```
<xs:schema   id                  = "htx"
             xmlns               = "http://www.jpeg.org/hiddentext/htx"
             xmlns:xs            = "http://www.w3.org/2001/XMLSchema"
             xmlns:xhtml         = "http://www.w3.org/1999/xhtml"
             targetNamespace     = "http://www.jpeg.org/hiddentext/htx"
             elementFormDefault  = "qualified">
```

The schema uses several types and elements from the XHTML 1.0 XML Schema, and imports this schema as follows:

```
<xs:import namespace      = "http://www.w3.org/1999/xhtml"
           schemaLocation  = "http://www.w3.org/2002/08/xhtml/xhtml1-strict.xsd"
```

The following tag is used to close the HTX schema:

```
</xs:schema>
```

© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

# Annex I
(informative)

# Hidden Text and Annotations Examples

## I.1 Example 1

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet href="#style" type="text/css"?>
<htx  xmlns               = "http://www.jpeg.org/hiddentext/htx"
      xmlns:xhtml         = "http://www.w3.org/1999/xhtml"
      xmlns:xsi           = "http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation  = "http://www.jpeg.org/hiddentext/htx htx.xsd"
      lang                = "en"
      dir                 = "ltr">

   <!-- ********************************************** -->
   <!-- Hidden text and annotations optional xhtml header  -->
      <!-- ********************************************** -->

    <xhtml:head lang="de" dir="ltr">
      <xhtml:title>Test Document Title</xhtml:title>
      <xhtml:style type="text/css" media="Screen"
            title="Style Sheet Name One" lang="de" dir="ltr" id="style">
            style                                  { display: none; }
            body, paragraph, line                  { display: block; }
            word, altword, char, altchar, snippet { display: inline; }
            body                                   { font-family: Arial,sans-
serif; }
            altword, altchar                       { font-style: italic; color:
grey; }
            [class~="red"]                         { background-color: red; }
            [class~="blue"]                        { color: blue; }
            [class~="large"]                       { font-size: 150%; }
      </xhtml:style>
    </xhtml:head>

    <!-- ********************************************** -->
    <!-- Optional annotations for page                  -->
    <!-- ********************************************** -->

   <annotations>
     <area  shape="rect" coords="0, 10, 50, 200" alt="link somewhere"
            href ="http://some.where.com/there" class="red" />
     <area  shape="poly" coords="23, 45, 0, 100, 90, 100"
             alt="link to somewhere alse" href ="http://somewhere.else.org/" />
   </annotations>

      <!-- ********************************************** -->
      <!-- Object for hidden text in the page             -->
      <!-- ********************************************** -->

   <hiddentext shape="rect" coords="0, 0, 100, 200">

      <!-- Optional, user defined page level parameters -->
      <param name="param1">abcd</param>
```


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
      <!-- First, mandatory region -->

      <region shape="rect" coords="0, 0, 100, 100">
         Here is some region level hidden text.

         <!-- Optional region parameters -->
         <param name="param2">efgh</param>

         <paragraph shape="poly"
                    coords="0, 0, 100, 0, 100, 50, 30, 50, 30, 60, 0, 60">
            Here is some paragraph level hidden text.

            <!-- Optional paragraph parameters -->
            <param name="param3">ijkl</param>
            <param name="param4">mnop</param>

            <line shape="rect" coords="0, 0, 100, 20" class="large" >
               <word conf="70%">
                  Word
                  <char    conf="90%">
                      c
                      <altchar conf="10%">o</altchar>
                  </char>
                  <char   conf="80%">
                      h
                     <altchar conf="12%">n</altchar>
                     <altchar conf= "8%">k</altchar>
                  </char>
                  <altword conf="15%">
                     Vordok
                  </altword>
                  <altword conf= "5%">
                     Wordoh
                  </altword>
               </word>

               <word>
                  Next
               </word>

               Some line level hidden text.
            </line>

         </paragraph>

         <line class="large blue">
            And some line level hidden text.
         </line>

      </region>

      <!-- Second, optional region -->
      <region>
          <!-- ... -->
      </region>

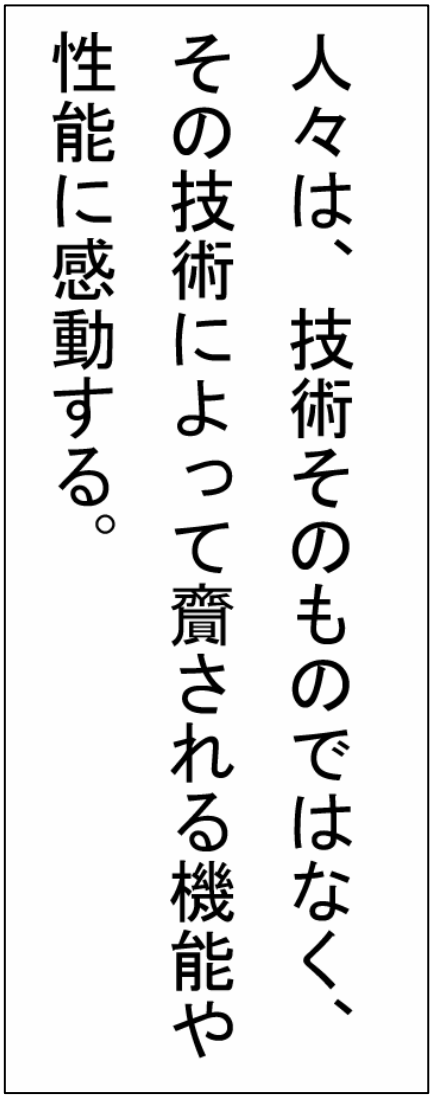
   </hiddentext>

</htx>
```

© ISO/IEC 2007 – All rights reserved --`,,```,,,,````-`-`,,`,,`,`,,`---
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

## I.2 Example 2

The following example shows how three lines in Japanese language can be formatted within HTX.




```
<paragraph angle="-90" baseline="90">
  <line shape="rect" coords="236, 20, 296,
854">
     <char conf="95%">人々は</char>
     <char conf="70%">、
        <altchar conf="40%">,</altchar>
     </char>
     <char conf="80%">技術</char>
     <char conf="90%">そのものではなく</char>
     <char conf="70">、
        <altchar conf="40%">,</altchar>
     </char>
  </line>

  <line shape="rect" coords="136, 20, 196,
888">
     <char conf="90%">その技術によって</char>
     <char conf="70%">齎
        <altchar conf="50%">斎</altchar>
     </char>
     <char conf="90%">される</char>
     <char conf="80%">機能</char>
     <char conf="90%">や</char>
  </line>
  </line shape="rect" coords="36, 20, 96,
452">
     <word conf="80%"><char>性能
  </char></word>
     <char conf="90%">に感動する</char>
     <char conf="70%">。
        <altchar conf="50%">.</altchar>
     </char>
  </line>
</paragraph>
```

## I.3 Example 3

This example shows an HTX that has been automatically generated from an OCR process' XML output by use of an Extensible Stylesheet Language Transformation (XSLT). It contains text formatting information and


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

character confidences as provided by the OCR process. The included reference to HTX2HTML.xsl enables an automatic conversion of this HTX to HTML within a web browser that supports XSLT.




**Figure I.1 — Example of a scanned document page**

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet href="HTX2HTML.xsl" type="text/xsl"?>
<htx xmlns="http://www.jpeg.org/hiddentext/htx" res="300"
     xmlns:xhtml=http://www.w3.org/1999/xhtml
     xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance
```

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
 xsi:schemaLocation="http://www.jpeg.org/hiddentext/htx htx.xsd">

<xhtml:head lang="de" dir="ltr">
    <xhtml:title/>
    <xhtml:style type="text/css" media="Screen" title="" lang="de" dir="ltr" id="style">
            style, param                          { display: none; }
            body, paragraph, line                 { display: block; }
            word, altword, char, altchar, snippet   { display: inline; }

            *.format-bold { font-weight:bold; }
            *.format-italic { font-style:italic; }
            *.format-subscript { vertical-align:sub; }
            *.format-superscript { vertical-align:super; }
            *.format-smallcaps { font-variant:small-caps; }
            *.format-underline { text-decoration:underline; }
            *.format-strikeout { text-decoration:line-through; }
            *.format-ffTimes-New-Roman { font-family:Times New Roman; }
            *.format-fs14 { font-size:14pt; }
            *.format-ffArial { font-family:Arial; }
            *.format-fs12 { font-size:12pt; }
            *.format-color14942208 { color:rgb(0, 0, 228); }
            *.format-fs32 { font-size:32pt; }
            *.format-fs10 { font-size:10pt; }
            </xhtml:style>
</xhtml:head>

<hiddentext>

    <region shape="rect" coords="1744, 566, 2196, 634">
        <paragraph>
            <line shape="rect" coords="1759, 578, 2179, 630">
                <word class="format-ffTimes-New-Roman format-fs14 ">
                    <char conf="100%" shape="rect" coords="1759, 579, 1793, 618">E</char>
                    <char conf="100%" shape="rect" coords="1798, 591, 1824, 630">g</char>
                    <char conf="100%" shape="rect" coords="1825, 592, 1854, 630">y</char>
                    <char conf="100%" shape="rect" coords="1855, 591, 1884, 630">p</char>
                    <char conf="100%" shape="rect" coords="1888, 582, 1905, 618">t</char>
                    <char conf="100%" shape="rect" coords="1905, 579, 1923, 618"> </char>
                </word>
                <word class="format-ffTimes-New-Roman format-fs14 ">
                    <char conf="100%" shape="rect" coords="1923, 579, 1958, 618">T</char>
                    <char conf="77%" shape="rect" coords="1962, 591, 1985, 618">r</char>
                    <char conf="100%" shape="rect" coords="1987, 591, 2013, 619">a</char>
                    <char conf="100%" shape="rect" coords="2016, 592, 2044, 619">v</char>
                    <char conf="100%" shape="rect" coords="2046, 591, 2068, 619">e</char>
                    <char conf="49%" shape="rect" coords="2071, 579, 2085, 618">l</char>
                    <char conf="49%" shape="rect" coords="2087, 579, 2101, 618">l</char>
                    <char conf="100%" shape="rect" coords="2103, 578, 2117, 618">i</char>
                    <char conf="100%" shape="rect" coords="2121, 591, 2150, 618">n</char>
                    <char conf="100%" shape="rect" coords="2153, 591, 2179, 630">g</char>
                </word>
            </line>
        </paragraph>
    </region>

    <region shape="rect" coords="360, 832, 2052, 1758">
        <paragraph>
            <line shape="rect" coords="379, 843, 888, 891">
                <word class="format-ffArial format-fs12 format-bold ">
                    <char conf="100%" shape="rect" coords="379, 844, 409, 880">B</char>
```


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <char conf="100%" shape="rect" coords="413, 853, 440, 881">o</char>
                    <char conf="100%" shape="rect" coords="444, 853, 471, 881">o</char>
                    <char conf="100%" shape="rect" coords="475, 844, 499, 880">k</char>
                    <char conf="100%" shape="rect" coords="503, 844, 510, 880">i</char>
                    <char conf="100%" shape="rect" coords="517, 853, 541, 880">n</char>
                    <char conf="100%" shape="rect" coords="546, 853, 571, 891">g</char>
                    <char conf="100%" shape="rect" coords="571, 853, 591, 881"> </char>
                </word>
                <word class="format-ffArial format-fs12 format-bold ">
                    <char conf="100%" shape="rect" coords="591, 853, 615, 881">c</char>
                    <char conf="100%" shape="rect" coords="619, 853, 646, 881">o</char>
                    <char conf="100%" shape="rect" coords="650, 853, 674, 880">n</char>
                    <char conf="100%" shape="rect" coords="678, 843, 696, 880">f</char>
                    <char conf="100%" shape="rect" coords="698, 844, 705, 880">i</char>
                    <char conf="100%" shape="rect" coords="711, 853, 728, 880">r</char>
                    <char conf="100%" shape="rect" coords="731, 853, 770, 880">m</char>
                    <char conf="100%" shape="rect" coords="774, 853, 798, 881">a</char>
                    <char conf="100%" shape="rect" coords="801, 845, 816, 881">t</char>
                    <char conf="100%" shape="rect" coords="820, 844, 827, 880">i</char>
                    <char conf="100%" shape="rect" coords="833, 853, 860, 881">o</char>
                    <char conf="100%" shape="rect" coords="864, 853, 888, 880">n</char>
                </word>
            </line>
        </paragraph>
        <paragraph>
            <line shape="rect" coords="379, 1015, 723, 1059">
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="379, 1016, 408, 1052">D</char>
                    <char conf="84%" shape="rect" coords="413, 1025, 437, 1053">e</char>
                    <char conf="100%" shape="rect" coords="441, 1025, 464, 1053">a</char>
                    <char conf="100%" shape="rect" coords="470, 1025, 484, 1052">r</char>
                    <char conf="100%" shape="rect" coords="484, 1015, 500, 1052"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="85%" shape="rect" coords="500, 1015, 531, 1053">C</char>
                    <char conf="100%" shape="rect" coords="537, 1026, 557, 1053">u</char>
                    <char conf="100%" shape="rect" coords="562, 1025, 584, 1053">s</char>
                    <char conf="100%" shape="rect" coords="587, 1017, 599, 1053">t</char>
                    <char conf="100%" shape="rect" coords="602, 1025, 626, 1053">o</char>
                    <char conf="100%" shape="rect" coords="631, 1025, 665, 1052">m</char>
                    <char conf="84%" shape="rect" coords="672, 1025, 696, 1053">e</char>
                    <char conf="100%" shape="rect" coords="701, 1025, 715, 1052">r</char>
                    <char conf="100%" shape="rect" coords="719, 1047, 723, 1059">,</char>
                </word>
            </line>
        </paragraph>
        <paragraph>
            <line shape="rect" coords="375, 1130, 2036, 1178">
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="375, 1141, 411, 1167">w</char>
                    <char conf="84%" shape="rect" coords="413, 1140, 437, 1168">e</char>
                    <char conf="84%" shape="rect" coords="437, 1140, 453, 1168"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="453, 1141, 489, 1167">w</char>
                    <char conf="100%" shape="rect" coords="491, 1140, 514, 1168">a</char>
                    <char conf="100%" shape="rect" coords="521, 1140, 541, 1167">n</char>
                    <char conf="100%" shape="rect" coords="546, 1132, 558, 1168">t</char>
                    <char conf="100%" shape="rect" coords="558, 1132, 573, 1168"> </char>
                </word>
```

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="573, 1132, 585, 1168">t</char>
                        <char conf="100%" shape="rect" coords="588, 1140, 612, 1168">o</char>
                        <char conf="100%" shape="rect" coords="612, 1131, 632, 1168"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="632, 1131, 636, 1167">i</char>
                        <char conf="100%" shape="rect" coords="643, 1140, 663, 1167">n</char>
                        <char conf="100%" shape="rect" coords="668, 1130, 683, 1167">f</char>
                        <char conf="100%" shape="rect" coords="683, 1140, 707, 1168">o</char>
                        <char conf="100%" shape="rect" coords="712, 1140, 726, 1167">r</char>
                        <char conf="100%" shape="rect" coords="728, 1140, 762, 1167">m</char>
                        <char conf="100%" shape="rect" coords="762, 1140, 782, 1168"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="782, 1141, 806, 1178">y</char>
                        <char conf="100%" shape="rect" coords="808, 1140, 832, 1168">o</char>
                        <char conf="100%" shape="rect" coords="838, 1141, 858, 1168">u</char>
                        <char conf="100%" shape="rect" coords="858, 1140, 877, 1168"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="877, 1140, 901, 1168">o</char>
                        <char conf="100%" shape="rect" coords="904, 1130, 919, 1167">f</char>
                        <char conf="100%" shape="rect" coords="919, 1130, 932, 1168"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="932, 1141, 956, 1178">y</char>
                        <char conf="100%" shape="rect" coords="958, 1140, 982, 1168">o</char>
                        <char conf="100%" shape="rect" coords="988, 1141, 1008, 1168">u</char>
                        <char conf="100%" shape="rect" coords="1015, 1140, 1029, 1167">r</char>
                        <char conf="100%" shape="rect" coords="1029, 1140, 1043, 1168"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1043, 1140, 1065, 1168">s</char>
                        <char conf="100%" shape="rect" coords="1071, 1141, 1091, 1168">u</char>
                        <char conf="100%" shape="rect" coords="1097, 1140, 1119, 1168">c</char>
                        <char conf="100%" shape="rect" coords="1122, 1140, 1144, 1168">c</char>
                        <char conf="82%" shape="rect" coords="1147, 1140, 1171, 1168">e</char>
                        <char conf="100%" shape="rect" coords="1173, 1140, 1195, 1168">s</char>
                        <char conf="100%" shape="rect" coords="1198, 1140, 1220, 1168">s</char>
                        <char conf="100%" shape="rect" coords="1223, 1130, 1238, 1167">f</char>
                        <char conf="100%" shape="rect" coords="1240, 1141, 1260, 1168">u</char>
                        <char conf="100%" shape="rect" coords="1268, 1131, 1272, 1167">l</char>
                        <char conf="100%" shape="rect" coords="1272, 1131, 1292, 1168"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1292, 1140, 1306, 1167">r</char>
                        <char conf="84%" shape="rect" coords="1308, 1140, 1332, 1168">e</char>
                        <char conf="100%" shape="rect" coords="1335, 1140, 1357, 1168">s</char>
                        <char conf="82%" shape="rect" coords="1361, 1140, 1385, 1168">e</char>
                        <char conf="100%" shape="rect" coords="1390, 1140, 1405, 1167">r</char>
                        <char conf="96%" shape="rect" coords="1405, 1141, 1428, 1167">v</char>
                        <char conf="100%" shape="rect" coords="1430, 1140, 1453, 1168">a</char>
                        <char conf="100%" shape="rect" coords="1457, 1132, 1469, 1168">t</char>
                        <char conf="100%" shape="rect" coords="1474, 1131, 1478, 1167">i</char>
                        <char conf="100%" shape="rect" coords="1483, 1140, 1507, 1168">o</char>
                        <char conf="100%" shape="rect" coords="1513, 1140, 1533, 1167">n</char>
                        <char conf="100%" shape="rect" coords="1542, 1162, 1546, 1167">.</char>
                        <char conf="100%" shape="rect" coords="1546, 1131, 1564, 1168"> </char>
                    </word>
```


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1564, 1131, 1596, 1167">Y</char>
                        <char conf="100%" shape="rect" coords="1600, 1140, 1624, 1168">o</char>
                        <char conf="100%" shape="rect" coords="1630, 1141, 1650, 1168">u</char>
                        <char conf="100%" shape="rect" coords="1656, 1140, 1670, 1167">r</char>
                        <char conf="100%" shape="rect" coords="1670, 1131, 1688, 1168"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1688, 1131, 1708, 1167">h</char>
                        <char conf="100%" shape="rect" coords="1714, 1140, 1738, 1168">o</char>
                        <char conf="100%" shape="rect" coords="1741, 1132, 1753, 1168">t</char>
                        <char conf="82%" shape="rect" coords="1756, 1140, 1780, 1168">e</char>
                        <char conf="100%" shape="rect" coords="1785, 1131, 1789, 1167">l</char>
                        <char conf="100%" shape="rect" coords="1789, 1131, 1810, 1168"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1810, 1131, 1814, 1167">i</char>
                        <char conf="100%" shape="rect" coords="1819, 1140, 1841, 1168">s</char>
                        <char conf="100%" shape="rect" coords="1841, 1131, 1860, 1168"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 format-underline ">
                        <char conf="100%" shape="rect" coords="1860, 1131, 1887, 1167">P</char>
                        <char conf="100%" shape="rect" coords="1890, 1141, 1914, 1171">y</char>
                        <char conf="38%" shape="rect" coords="1917, 1140, 1931, 1167">r</char>
                        <char conf="56%" shape="rect" coords="1933, 1140, 1956, 1168">a</char>
                        <char conf="100%" shape="rect" coords="1962, 1140, 1996, 1167">m</char>
                        <char conf="100%" shape="rect" coords="2004, 1131, 2008, 1167">i</char>
                        <char conf="100%" shape="rect" coords="2014, 1131, 2036, 1168">d</char>
                    </word>
                </line>
                <line shape="rect" coords="375, 1187, 1966, 1227">
                    <word class="format-ffArial format-fs12 format-underline ">
                        <char conf="100%" shape="rect" coords="375, 1188, 408, 1224">V</char>
                        <char conf="100%" shape="rect" coords="412, 1188, 416, 1224">i</char>
                        <char conf="91%" shape="rect" coords="421, 1197, 445, 1225">e</char>
                        <char conf="100%" shape="rect" coords="447, 1198, 483, 1224">w</char>
                        <char conf="100%" shape="rect" coords="483, 1188, 501, 1225"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 format-underline ">
                        <char conf="100%" shape="rect" coords="501, 1188, 532, 1224">R</char>
                        <char conf="91%" shape="rect" coords="535, 1197, 559, 1225">e</char>
                        <char conf="100%" shape="rect" coords="562, 1197, 584, 1225">s</char>
                        <char conf="100%" shape="rect" coords="588, 1197, 612, 1225">o</char>
                        <char conf="37%" shape="rect" coords="617, 1197, 631, 1224">r</char>
                        <char conf="70%" shape="rect" coords="632, 1189, 644, 1225">t</char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="676, 1187, 708, 1225">G</char>
                        <char conf="100%" shape="rect" coords="715, 1188, 719, 1224">i</char>
                        <char conf="100%" shape="rect" coords="724, 1198, 747, 1224">z</char>
                        <char conf="100%" shape="rect" coords="750, 1197, 773, 1225">a</char>
                        <char conf="100%" shape="rect" coords="780, 1219, 784, 1224">.</char>
                        <char conf="100%" shape="rect" coords="784, 1188, 803, 1225"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="803, 1188, 835, 1224">Y</char>
                        <char conf="100%" shape="rect" coords="839, 1197, 863, 1225">o</char>
                        <char conf="100%" shape="rect" coords="868, 1198, 888, 1225">u</char>
                        <char conf="100%" shape="rect" coords="888, 1188, 910, 1225"> </char>
                    </word>
```

--`,,```,,,,````-`-`,,`,,`,`,,`---
© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="910, 1188, 932, 1225">b</char>
                        <char conf="100%" shape="rect" coords="936, 1197, 960, 1225">o</char>
                        <char conf="100%" shape="rect" coords="964, 1197, 988, 1225">o</char>
                        <char conf="100%" shape="rect" coords="994, 1188, 1015, 1224">k</char>
                        <char conf="84%" shape="rect" coords="1017, 1197, 1041, 1225">e</char>
                        <char conf="100%" shape="rect" coords="1044, 1188, 1066, 1225">d</char>
                        <char conf="100%" shape="rect" coords="1066, 1188, 1086, 1225"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1086, 1197, 1109, 1225">a</char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1109, 1189, 1127, 1225"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1127, 1189, 1139, 1225">t</char>
                        <char conf="100%" shape="rect" coords="1140, 1198, 1176, 1224">w</char>
                        <char conf="100%" shape="rect" coords="1180, 1188, 1184, 1224">i</char>
                        <char conf="100%" shape="rect" coords="1191, 1197, 1211, 1224">n</char>
                        <char conf="100%" shape="rect" coords="1211, 1197, 1232, 1225"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1232, 1197, 1246, 1224">r</char>
                        <char conf="100%" shape="rect" coords="1247, 1197, 1271, 1225">o</char>
                        <char conf="100%" shape="rect" coords="1275, 1197, 1299, 1225">o</char>
                        <char conf="100%" shape="rect" coords="1304, 1197, 1338, 1224">m</char>
                        <char conf="100%" shape="rect" coords="1338, 1187, 1357, 1225"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1357, 1187, 1372, 1224">f</char>
                        <char conf="100%" shape="rect" coords="1373, 1197, 1387, 1224">r</char>
                        <char conf="100%" shape="rect" coords="1389, 1197, 1413, 1225">o</char>
                        <char conf="100%" shape="rect" coords="1417, 1197, 1451, 1224">m</char>
                        <char conf="100%" shape="rect" coords="1451, 1188, 1472, 1225"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1472, 1188, 1495, 1225">3</char>
                        <char conf="100%" shape="rect" coords="1498, 1188, 1512, 1225">/</char>
                        <char conf="100%" shape="rect" coords="1514, 1189, 1537, 1224">7</char>
                        <char conf="100%" shape="rect" coords="1540, 1188, 1554, 1225">/</char>
                        <char conf="100%" shape="rect" coords="1555, 1188, 1578, 1224">2</char>
                        <char conf="100%" shape="rect" coords="1584, 1188, 1606, 1225">0</char>
                        <char conf="100%" shape="rect" coords="1612, 1188, 1634, 1225">0</char>
                        <char conf="100%" shape="rect" coords="1639, 1188, 1662, 1225">6</char>
                        <char conf="100%" shape="rect" coords="1662, 1188, 1680, 1225"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1680, 1189, 1692, 1225">t</char>
                        <char conf="100%" shape="rect" coords="1695, 1197, 1719, 1225">o</char>
                        <char conf="100%" shape="rect" coords="1719, 1188, 1740, 1225"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1740, 1188, 1753, 1224">1</char>
                        <char conf="100%" shape="rect" coords="1764, 1188, 1787, 1225">6</char>
                        <char conf="100%" shape="rect" coords="1790, 1188, 1804, 1225">/</char>
                        <char conf="100%" shape="rect" coords="1806, 1189, 1829, 1224">7</char>
                        <char conf="100%" shape="rect" coords="1832, 1188, 1846, 1225">/</char>
                        <char conf="100%" shape="rect" coords="1847, 1188, 1870, 1224">2</char>
                        <char conf="100%" shape="rect" coords="1876, 1188, 1898, 1225">0</char>
```


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <char conf="100%" shape="rect" coords="1904, 1188, 1926, 1225">0</char>
                    <char conf="100%" shape="rect" coords="1931, 1188, 1954, 1225">6</char>
                    <char conf="100%" shape="rect" coords="1962, 1219, 1966, 1224">.</char>
                </word>
            </line>
            <line shape="rect" coords="377, 1245, 1220, 1293">
                <word class="format-ffArial format-fs12 format-bold format-italic ">
                    <char conf="100%" shape="rect" coords="377, 1246, 410, 1282">P</char>
                    <char conf="100%" shape="rect" coords="410, 1246, 425, 1282">l</char>
                    <char conf="100%" shape="rect" coords="425, 1255, 450, 1283">e</char>
                    <char conf="100%" shape="rect" coords="452, 1255, 477, 1283">a</char>
                    <char conf="100%" shape="rect" coords="479, 1255, 506, 1283">s</char>
                    <char conf="100%" shape="rect" coords="508, 1255, 533, 1283">e</char>
                    <char conf="100%" shape="rect" coords="533, 1255, 549, 1283"> </char>
                </word>
                <word class="format-ffArial format-fs12 format-bold format-italic ">
                    <char conf="100%" shape="rect" coords="549, 1255, 574, 1283">a</char>
                    <char conf="100%" shape="rect" coords="577, 1255, 599, 1282">r</char>
                    <char conf="100%" shape="rect" coords="596, 1255, 618, 1282">r</char>
                    <char conf="100%" shape="rect" coords="616, 1246, 631, 1282">i</char>
                    <char conf="100%" shape="rect" coords="632, 1256, 659, 1282">v</char>
                    <char conf="100%" shape="rect" coords="658, 1255, 683, 1283">e</char>
                    <char conf="100%" shape="rect" coords="683, 1246, 699, 1283"> </char>
                </word>
                <word class="format-ffArial format-fs12 format-bold format-italic ">
                    <char conf="100%" shape="rect" coords="699, 1246, 727, 1283">b</char>
                    <char conf="100%" shape="rect" coords="728, 1256, 759, 1293">y</char>
                    <char conf="100%" shape="rect" coords="759, 1246, 771, 1283"> </char>
                </word>
                <word class="format-ffArial format-fs12 format-bold format-italic ">
                    <char conf="100%" shape="rect" coords="771, 1246, 798, 1282">4</char>
                </word>
                <word class="format-ffArial format-fs12 format-bold format-italic ">
                    <char conf="100%" shape="rect" coords="798, 1246, 810, 1283"> </char>
                </word>
                <word class="format-ffArial format-fs12 format-bold format-italic ">
                    <char conf="100%" shape="rect" coords="810, 1255, 841, 1292">p</char>
                    <char conf="100%" shape="rect" coords="844, 1275, 852, 1282">.</char>
                    <char conf="100%" shape="rect" coords="858, 1255, 900, 1282">m</char>
                    <char conf="100%" shape="rect" coords="902, 1275, 910, 1282">.</char>
                    <char conf="100%" shape="rect" coords="910, 1247, 932, 1283"> </char>
                </word>
                <word class="format-ffArial format-fs12 format-bold format-italic ">
                    <char conf="100%" shape="rect" coords="932, 1247, 947, 1283">t</char>
                    <char conf="100%" shape="rect" coords="947, 1246, 974, 1282">h</char>
                    <char conf="100%" shape="rect" coords="978, 1255, 1003, 1283">e</char>
                    <char conf="100%" shape="rect" coords="1003, 1245, 1020, 1283"> </char>
                </word>
                <word class="format-ffArial format-fs12 format-bold format-italic ">
                    <char conf="100%" shape="rect" coords="1020, 1245, 1041, 1282">f</char>
                    <char conf="100%" shape="rect" coords="1035, 1246, 1050, 1282">i</char>
                    <char conf="100%" shape="rect" coords="1049, 1255, 1071, 1282">r</char>
                    <char conf="100%" shape="rect" coords="1068, 1255, 1095, 1283">s</char>
                    <char conf="100%" shape="rect" coords="1098, 1247, 1113, 1283">t</char>
                    <char conf="100%" shape="rect" coords="1113, 1246, 1128, 1283"> </char>
                </word>
                <word class="format-ffArial format-fs12 format-bold format-italic ">
                    <char conf="100%" shape="rect" coords="1128, 1246, 1159, 1283">d</char>
                    <char conf="100%" shape="rect" coords="1158, 1255, 1183, 1283">a</char>
                    <char conf="100%" shape="rect" coords="1183, 1256, 1214, 1293">y</char>
```

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <char conf="100%" shape="rect" coords="1216, 1277, 1220, 1282">.</char>
                </word>
            </line>
        </paragraph>
        <paragraph>
            <line shape="rect" coords="380, 1360, 1688, 1407">
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="380, 1361, 384, 1397">I</char>
                    <char conf="100%" shape="rect" coords="393, 1370, 413, 1397">n</char>
                    <char conf="100%" shape="rect" coords="413, 1370, 433, 1398"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="433, 1370, 455, 1398">c</char>
                    <char conf="100%" shape="rect" coords="458, 1370, 481, 1398">a</char>
                    <char conf="100%" shape="rect" coords="484, 1370, 506, 1398">s</char>
                    <char conf="82%" shape="rect" coords="510, 1370, 534, 1398">e</char>
                    <char conf="82%" shape="rect" coords="534, 1370, 552, 1398"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="552, 1370, 576, 1398">o</char>
                    <char conf="100%" shape="rect" coords="579, 1360, 594, 1397">f</char>
                    <char conf="100%" shape="rect" coords="594, 1360, 610, 1398"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="610, 1370, 632, 1407">p</char>
                    <char conf="100%" shape="rect" coords="637, 1370, 651, 1397">r</char>
                    <char conf="100%" shape="rect" coords="652, 1370, 676, 1398">o</char>
                    <char conf="100%" shape="rect" coords="682, 1361, 704, 1398">b</char>
                    <char conf="100%" shape="rect" coords="710, 1361, 714, 1397">l</char>
                    <char conf="82%" shape="rect" coords="719, 1370, 743, 1398">e</char>
                    <char conf="100%" shape="rect" coords="748, 1370, 782, 1397">m</char>
                    <char conf="100%" shape="rect" coords="787, 1370, 809, 1398">s</char>
                    <char conf="100%" shape="rect" coords="809, 1370, 827, 1398"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="827, 1370, 851, 1398">o</char>
                    <char conf="100%" shape="rect" coords="856, 1370, 870, 1397">r</char>
                    <char conf="100%" shape="rect" coords="870, 1370, 886, 1398"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="886, 1370, 908, 1407">q</char>
                    <char conf="100%" shape="rect" coords="916, 1371, 936, 1398">u</char>
                    <char conf="82%" shape="rect" coords="941, 1370, 965, 1398">e</char>
                    <char conf="100%" shape="rect" coords="968, 1370, 990, 1398">s</char>
                    <char conf="100%" shape="rect" coords="993, 1362, 1005, 1398">t</char>
                    <char conf="100%" shape="rect" coords="1010, 1361, 1014, 1397">i</char>
                    <char conf="100%" shape="rect" coords="1019, 1370, 1043, 1398">o</char>
                    <char conf="100%" shape="rect" coords="1049, 1370, 1069, 1397">n</char>
                    <char conf="100%" shape="rect" coords="1074, 1370, 1096, 1398">s</char>
                    <char conf="100%" shape="rect" coords="1103, 1392, 1107, 1404">,</char>
                    <char conf="100%" shape="rect" coords="1107, 1370, 1130, 1404"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="1130, 1370, 1152, 1407">p</char>
                    <char conf="100%" shape="rect" coords="1157, 1361, 1161, 1397">l</char>
                    <char conf="82%" shape="rect" coords="1166, 1370, 1190, 1398">e</char>
                    <char conf="100%" shape="rect" coords="1194, 1370, 1217, 1398">a</char>
                    <char conf="100%" shape="rect" coords="1221, 1370, 1243, 1398">s</char>
                    <char conf="82%" shape="rect" coords="1247, 1370, 1271, 1398">e</char>
                    <char conf="82%" shape="rect" coords="1271, 1360, 1288, 1398"> </char>
                </word>
```


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1288, 1360, 1303, 1397">f</char>
                        <char conf="82%" shape="rect" coords="1303, 1370, 1327, 1398">e</char>
                        <char conf="82%" shape="rect" coords="1330, 1370, 1354, 1398">e</char>
                        <char conf="100%" shape="rect" coords="1360, 1361, 1364, 1397">l</char>
                        <char conf="100%" shape="rect" coords="1364, 1360, 1382, 1398"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1382, 1360, 1397, 1397">f</char>
                        <char conf="100%" shape="rect" coords="1398, 1370, 1412, 1397">r</char>
                        <char conf="84%" shape="rect" coords="1414, 1370, 1438, 1398">e</char>
                        <char conf="82%" shape="rect" coords="1442, 1370, 1466, 1398">e</char>
                        <char conf="82%" shape="rect" coords="1466, 1362, 1482, 1398"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1482, 1362, 1494, 1398">t</char>
                        <char conf="100%" shape="rect" coords="1497, 1370, 1521, 1398">o</char>
                        <char conf="100%" shape="rect" coords="1521, 1370, 1539, 1398"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1539, 1370, 1561, 1398">c</char>
                        <char conf="100%" shape="rect" coords="1564, 1370, 1587, 1398">a</char>
                        <char conf="100%" shape="rect" coords="1594, 1361, 1598, 1397">l</char>
                        <char conf="100%" shape="rect" coords="1605, 1361, 1609, 1397">l</char>
                        <char conf="100%" shape="rect" coords="1609, 1361, 1630, 1398"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="1630, 1371, 1650, 1398">u</char>
                        <char conf="100%" shape="rect" coords="1655, 1370, 1677, 1398">s</char>
                        <char conf="100%" shape="rect" coords="1684, 1371, 1688, 1397">:</char>
                    </word>
                </line>
            </paragraph>
            <paragraph>
                <line shape="rect" coords="1011, 1475, 1471, 1523">
                    <word class="format-ffArial format-fs12 format-color14942208 format-bold ">
                        <char conf="100%" shape="rect" coords="1011, 1482, 1034, 1506">+</char>
                        <char conf="100%" shape="rect" coords="1039, 1482, 1063, 1506">+</char>
                        <char conf="100%" shape="rect" coords="1069, 1476, 1085, 1512">1</char>
                        <char conf="100%" shape="rect" coords="1094, 1476, 1118, 1512">2</char>
                        <char conf="66%" shape="rect" coords="1136, 1475, 1149, 1523">(</char>
                        <char conf="100%" shape="rect" coords="1153, 1476, 1177, 1513">3</char>
                        <char conf="100%" shape="rect" coords="1180, 1476, 1206, 1512">4</char>
                        <char conf="67%" shape="rect" coords="1208, 1475, 1221, 1523">)</char>
                        <char conf="100%" shape="rect" coords="1241, 1476, 1257, 1512">1</char>
                        <char conf="100%" shape="rect" coords="1257, 1476, 1280, 1512"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 format-color14942208 format-bold ">
                        <char conf="100%" shape="rect" coords="1280, 1476, 1304, 1512">2</char>
                        <char conf="100%" shape="rect" coords="1309, 1476, 1333, 1513">3</char>
                        <char conf="100%" shape="rect" coords="1333, 1476, 1349, 1512"> </char>
                    </word>
                    <word class="format-ffArial format-fs12 format-color14942208 format-bold ">
                        <char conf="100%" shape="rect" coords="1349, 1476, 1375, 1512">4</char>
                        <char conf="100%" shape="rect" coords="1378, 1477, 1402, 1513">5</char>
                        <char conf="100%" shape="rect" coords="1402, 1476, 1420, 1513"> </char>
                    </word>
```

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                <word class="format-ffArial format-fs12 format-color14942208 format-bold ">
                    <char conf="100%" shape="rect" coords="1420, 1476, 1444, 1513">6</char>
                    <char conf="74%" shape="rect" coords="1448, 1477, 1471, 1512">7</char>
                </word>
            </line>
        </paragraph>
        <paragraph>
            <line shape="rect" coords="376, 1591, 1009, 1638">
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="376, 1591, 422, 1627">W</char>
                    <char conf="84%" shape="rect" coords="424, 1600, 448, 1628">e</char>
                    <char conf="84%" shape="rect" coords="448, 1600, 464, 1628"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="464, 1601, 500, 1627">w</char>
                    <char conf="100%" shape="rect" coords="504, 1591, 508, 1627">i</char>
                    <char conf="100%" shape="rect" coords="512, 1600, 534, 1628">s</char>
                    <char conf="100%" shape="rect" coords="540, 1591, 560, 1627">h</char>
                    <char conf="100%" shape="rect" coords="560, 1591, 579, 1628"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="579, 1601, 603, 1638">y</char>
                    <char conf="100%" shape="rect" coords="605, 1600, 629, 1628">o</char>
                    <char conf="100%" shape="rect" coords="635, 1601, 655, 1628">u</char>
                    <char conf="100%" shape="rect" coords="655, 1600, 674, 1628"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="674, 1600, 697, 1628">a</char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="697, 1600, 716, 1628"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="716, 1600, 739, 1638">g</char>
                    <char conf="100%" shape="rect" coords="745, 1600, 759, 1627">r</char>
                    <char conf="84%" shape="rect" coords="761, 1600, 785, 1628">e</char>
                    <char conf="100%" shape="rect" coords="788, 1600, 811, 1628">a</char>
                    <char conf="100%" shape="rect" coords="815, 1592, 827, 1628">t</char>
                    <char conf="100%" shape="rect" coords="827, 1591, 846, 1628"> </char>
                </word>
                <word class="format-ffArial format-fs12 ">
                    <char conf="100%" shape="rect" coords="846, 1591, 866, 1627">h</char>
                    <char conf="100%" shape="rect" coords="872, 1600, 896, 1628">o</char>
                    <char conf="100%" shape="rect" coords="902, 1591, 906, 1627">l</char>
                    <char conf="100%" shape="rect" coords="913, 1591, 917, 1627">i</char>
                    <char conf="100%" shape="rect" coords="922, 1591, 944, 1628">d</char>
                    <char conf="100%" shape="rect" coords="950, 1600, 973, 1628">a</char>
                    <char conf="100%" shape="rect" coords="976, 1601, 1000, 1638">y</char>
                    <char conf="100%" shape="rect" coords="1004, 1591, 1009, 1627">!</char>
                </word>
            </line>
        </paragraph>
        <paragraph>
            <line shape="rect" coords="379, 1706, 657, 1753">
                <word class="format-ffArial format-fs12 ">
                    <char conf="88%" shape="rect" coords="379, 1706, 405, 1742">B</char>
                    <char conf="84%" shape="rect" coords="410, 1715, 434, 1743">e</char>
                    <char conf="100%" shape="rect" coords="437, 1715, 459, 1743">s</char>
                    <char conf="100%" shape="rect" coords="462, 1707, 474, 1743">t</char>
                    <char conf="100%" shape="rect" coords="474, 1707, 492, 1743"> </char>
                </word>
```


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <word class="format-ffArial format-fs12 ">
                        <char conf="100%" shape="rect" coords="492, 1715, 506, 1742">r</char>
                        <char conf="84%" shape="rect" coords="508, 1715, 532, 1743">e</char>
                        <char conf="100%" shape="rect" coords="535, 1715, 558, 1753">g</char>
                        <char conf="100%" shape="rect" coords="563, 1715, 586, 1743">a</char>
                        <char conf="100%" shape="rect" coords="592, 1715, 606, 1742">r</char>
                        <char conf="100%" shape="rect" coords="608, 1706, 630, 1743">d</char>
                        <char conf="100%" shape="rect" coords="635, 1715, 657, 1743">s</char>
                    </word>
                </line>
            </paragraph>
        </region>

        <region shape="rect" coords="364, 1862, 830, 2018">
            <paragraph>
                <line shape="rect" coords="380, 1892, 814, 2014">
                    <word class="format-ffTimes-New-Roman format-fs32 format-bold format-italic ">
                        <char conf="1%" shape="rect" coords="380, 1892, 412, 1978">1</char>
                        <char conf="15%" shape="rect" coords="402, 1894, 460, 1986">2</char>
                        <char conf="100%" shape="rect" coords="450, 1914, 570, 1994">^</char>
                        <char conf="28%" shape="rect" coords="586, 1906, 628, 2002">!</char>
                        <char conf="100%" shape="rect" coords="610, 1908, 814, 2014">^</char>
                    </word>
                </line>
            </paragraph>
        </region>

        <region shape="rect" coords="362, 2028, 596, 2080">
            <paragraph>
                <line shape="rect" coords="378, 2036, 579, 2075">
                    <word class="format-ffArial format-fs10 ">
                        <char conf="100%" shape="rect" coords="378, 2036, 401, 2066">B</char>
                        <char conf="100%" shape="rect" coords="404, 2043, 425, 2067">o</char>
                        <char conf="100%" shape="rect" coords="429, 2036, 448, 2067">b</char>
                        <char conf="100%" shape="rect" coords="448, 2036, 464, 2067"> </char>
                    </word>
                    <word class="format-ffArial format-fs10 ">
                        <char conf="100%" shape="rect" coords="464, 2036, 487, 2066">B</char>
                        <char conf="100%" shape="rect" coords="490, 2043, 511, 2067">o</char>
                        <char conf="100%" shape="rect" coords="515, 2036, 534, 2067">b</char>
                        <char conf="100%" shape="rect" coords="538, 2036, 557, 2067">b</char>
                        <char conf="100%" shape="rect" coords="559, 2044, 579, 2075">y</char>
                    </word>
                </line>
            </paragraph>
        </region>

    </hiddentext>

</htx>
```

--`,,```,,,,````-`-`,,`,,`,`,,`---



© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

The following XSLT can be used to translate HTX for viewing in a web browser. It is not the intension of the HTX to replace the original image. Translation to HTML might be used to review the OCR results, or to help creating JPM viewers that provide copy & paste operations of the hidden text underneath the page image.

```
<?xml version="1.0" encoding="ISO-8859-1"?>

<!--
    XSL Script for translating from HTX to HTML.

    Tested with
        - Xalan-C++ 1.10.0
        - Internet Explorer 6
        - Mozilla Firefox 1.5


    Author: Andreas Hirth, LuraTech Imaging GmbH, www.luratech.com

    Todo:
        - supporting polygonic bounding shapes

    Note:
        - altchar and altword are skipped
        - absolute positioning is used for every element with coordinates
-->

<xsl:stylesheet version="1.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xhtml="http://www.w3.org/1999/xhtml"
  xmlns:xalan="http://xml.apache.org/xalan"
  xmlns:htx="http://www.jpeg.org/hiddentext/htx"
  exclude-result-prefixes="xsl xsi xhtml xalan htx">

    <xsl:output method="html" indent="yes" xalan:indent-amount="4"/>

    <!-- Make the character enclosing <span> tags be written without linebreak
    or whitespaces in between. This is important because the browser will put
    a blank between each character otherwise. -->
    <xsl:strip-space elements="*"/>

    <!-- Make a <title> tag from a xhtml <title> tag > -->
    <xsl:template match="xhtml:head/xhtml:title">
        <title><xsl:value-of select="text()"/></title>
    </xsl:template>

    <!-- Make a <style> tag from a xhtml <style> tag > -->
    <xsl:template match="xhtml:head/xhtml:style">
        <style type="text/css"><xsl:value-of select="text()"/></style>
    </xsl:template>

    <!-- Translation entry point -->
    <xsl:template match="/">
        <html>
            <head>
                <xsl:apply-templates select="//xhtml:title"/>
                <xsl:apply-templates select="//xhtml:style"/>
            </head>
            <body>
                <xsl:apply-templates select="htx:htx"/>
            </body>
        </html>
    </xsl:template>
```

--`,,```,,,,````-`-`,,`,,`,`,,`---



© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
    <!-- There is nothing to do with the root element -->
    <xsl:template match="htx:htx">
        <xsl:apply-templates/>
    </xsl:template>

    <!-- Skip annotations -->
    <xsl:template match="htx:annotations"/>

    <!-- There is nothing to do with the hiddentext element -->
    <xsl:template match="htx:hiddentext">
        <xsl:apply-templates/>
    </xsl:template>

    <!-- Translate the <region> element into a <div> tag with core attributes and positioning -->
    <xsl:template match="htx:region">
        <xsl:element name="div">
            <xsl:call-template name="outputCoreAttributes"/>
            <xsl:call-template name="outputPositionAttributes"/>

            <xsl:apply-templates/>
        </xsl:element>
    </xsl:template>

    <!-- Translate the <paragraph> element into a <p> tag with core attributes and positioning -->
    <xsl:template match="htx:paragraph">
        <xsl:element name="p">
            <xsl:call-template name="outputCoreAttributes"/>
            <xsl:call-template name="outputPositionAttributes"/>

            <xsl:apply-templates/>
        </xsl:element>
    </xsl:template>

    <!-- Translate the <line> element into a <span> tag with core attributes and positioning -->
    <xsl:template match="htx:line">
        <xsl:element name="span">
            <xsl:call-template name="outputCoreAttributes"/>
            <xsl:call-template name="outputPositionAttributes"/>

            <xsl:apply-templates/>
        </xsl:element>
    </xsl:template>

    <!-- Translate the <word> element into a <span> tag with core attributes and positioning -->
    <xsl:template match="htx:word">
        <xsl:element name="span">
            <xsl:call-template name="outputCoreAttributes"/>
            <xsl:call-template name="outputPositionAttributes"/>

            <xsl:apply-templates/>
        </xsl:element>
    </xsl:template>

    <!-- Translate the <char> element into a <span> tag with core attributes and
    special positioning -->
    <xsl:template match="htx:char">
        <xsl:element name="span">
            <xsl:call-template name="outputCoreAttributes"/>
            <xsl:call-template name="outputCharacterPositionAttributes"/>
```

--`,,```,,,,````-`-`,,`,,`,`,,`---

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
            <xsl:value-of select="text()"/>
        </xsl:element>
    </xsl:template>

    <!-- Do not translate any other elements, like <altword> or <altchar> -->
    <xsl:template match="*"/>

    <!-- Parse and output the core attributes of an element -->
    <xsl:template name="outputCoreAttributes">
        <xsl:if test="@class">
            <xsl:attribute name="class"><xsl:value-of select="@class"/></xsl:attribute>
        </xsl:if>

        <xsl:if test="@lang">
            <xsl:attribute name="lang"><xsl:value-of select="@lang"/></xsl:attribute>
        </xsl:if>

        <xsl:if test="@xml:lang">
            <xsl:attribute name="xml:lang"><xsl:value-of select="@xml:lang"/></xsl:attribute>
        </xsl:if>

        <xsl:if test="@dir">
            <xsl:attribute name="dir"><xsl:value-of select="@dir"/></xsl:attribute>
        </xsl:if>

        <xsl:if test="@id">
            <xsl:attribute name="id"><xsl:value-of select="@id"/></xsl:attribute>
        </xsl:if>
    </xsl:template>

    <!-- Parse and output positioning attributes -->
    <xsl:template name="outputPositionAttributes">
        <xsl:choose>
            <xsl:when test="@shape='poly'">
                <xsl:message>Polygonic shapes are not supported yet. No positioning attributes are
written for element &lt;<xsl:value-of select="name()"/>&gt;. </xsl:message>
            </xsl:when>
            <xsl:otherwise>
                <xsl:call-template name="outputRelativeCoordinatesAttribute"/>
            </xsl:otherwise>
        </xsl:choose>
    </xsl:template>

    <!-- Parse and output positioning attributes especialy for characters -->
    <xsl:template name="outputCharacterPositionAttributes">
        <xsl:choose>
            <xsl:when test="@shape='poly'">
                <xsl:message>Polygonic shapes are not supported yet. No positioning attributes are
written for element <xsl:value-of select="name()"/></xsl:message>
            </xsl:when>
            <xsl:otherwise>
                <xsl:call-template name="outputRelativeCharacterCoordinatesAttribute"/>
            </xsl:otherwise>
        </xsl:choose>
    </xsl:template>

    <!-- Calculate and output positioning attributes from coords. Must be recalculated to be
    relative to the nearest parent with coordinates. -->
    <xsl:template name="outputRelativeCoordinatesAttribute">
        <xsl:variable name="horRes">
```


--`,,```,,,,````-`-,,`,,`,`,,`---
© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
            <xsl:choose>
                <xsl:when test="/htx:htx[(position()=1) and (@res)]">
                    <xsl:variable name="res"><xsl:value-of
select="/htx:htx[1]/@res"/></xsl:variable>
                    <xsl:choose>
                        <xsl:when test="contains($res, ',')">
                            <xsl:value-of select="normalize-space(substring-before($res, ','))"/>
                        </xsl:when>
                        <xsl:otherwise>
                            <xsl:value-of select="$res"/>
                        </xsl:otherwise>
                    </xsl:choose>
                </xsl:when>
                <xsl:otherwise>0</xsl:otherwise>
            </xsl:choose>
        </xsl:variable>
        <xsl:variable name="vertRes">
            <xsl:choose>
                <xsl:when test="/htx:htx[(position()=1) and (@res)]">
                    <xsl:variable name="res"><xsl:value-of
select="/htx:htx[1]/@res"/></xsl:variable>
                    <xsl:choose>
                        <xsl:when test="contains($res, ',')">
                            <xsl:value-of select="normalize-space(substring-after($res, ','))"/>
                        </xsl:when>
                        <xsl:otherwise>
                            <xsl:value-of select="$res"/>
                        </xsl:otherwise>
                    </xsl:choose>
                </xsl:when>
                <xsl:otherwise>0</xsl:otherwise>
            </xsl:choose>
        </xsl:variable>

        <xsl:variable name="left">
            <xsl:variable name="coord"><xsl:value-of select="number(normalize-space(substring-
before(@coords, ',')))"/></xsl:variable>
            <xsl:choose>
                <xsl:when test="number($horRes) &gt; 0">
                    <xsl:value-of select="$coord div $horRes"/>
                </xsl:when>
                <xsl:otherwise>
                    <xsl:value-of select="$coord"/>
                </xsl:otherwise>
            </xsl:choose>
        </xsl:variable>
        <xsl:variable name="top">
            <xsl:variable name="coord"><xsl:value-of select="number(normalize-space(substring-
before(substring-after(@coords, ','), ',')))"/></xsl:variable>
            <xsl:choose>
                <xsl:when test="number($vertRes) &gt; 0">
                    <xsl:value-of select="$coord div $vertRes"/>
                </xsl:when>
                <xsl:otherwise>
                    <xsl:value-of select="$coord"/>
                </xsl:otherwise>
            </xsl:choose>
        </xsl:variable>
```

--`,,```,,,,````-`-`,,`,,`,`,,`---
© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
        <xsl:if test="(@shape='rect') and (string-length($left)>0) and (string-length($top)>0)">
            <xsl:variable name="parentLeft">
                <xsl:choose>
                    <xsl:when test="ancestor::*[(@shape='rect') and (@coords)]">
                        <xsl:variable name="coord"><xsl:value-of select="number(normalize-
space(substring-before(ancestor::*[(@shape='rect') and (@coords)][1]/@coords,
',')))"/></xsl:variable>
                        <xsl:choose>
                            <xsl:when test="number($horRes) &gt; 0">
                                <xsl:value-of select="$coord div $horRes"/>
                            </xsl:when>
                            <xsl:otherwise>
                                <xsl:value-of select="$coord"/>
                            </xsl:otherwise>
                        </xsl:choose>
                    </xsl:when>
                    <xsl:otherwise>0</xsl:otherwise>
                </xsl:choose>
            </xsl:variable>

            <xsl:variable name="parentTop">
                <xsl:choose>
                    <xsl:when test="ancestor::*[(@shape='rect') and (@coords)]">
                        <xsl:variable name="coord"><xsl:value-of select="number(normalize-
space(substring-before(substring-after(ancestor::*[(@shape='rect') and (@coords)][1]/@coords, ','),
',')))"/></xsl:variable>
                        <xsl:choose>
                            <xsl:when test="number($vertRes) &gt; 0">
                                <xsl:value-of select="$coord div $vertRes"/>
                            </xsl:when>
                            <xsl:otherwise>
                                <xsl:value-of select="$coord"/>
                            </xsl:otherwise>
                        </xsl:choose>
                    </xsl:when>
                    <xsl:otherwise>0</xsl:otherwise>
                </xsl:choose>
            </xsl:variable>

            <xsl:variable name="unit">
                <xsl:choose>
                    <xsl:when test="(number($horRes) &gt; 0) and (number($vertRes) &gt;
0)">in</xsl:when>
                    <xsl:otherwise>px</xsl:otherwise>
                </xsl:choose>
            </xsl:variable>
            <xsl:attribute name="style">position:absolute; left:<xsl:value-of select="$left -
$parentLeft"/><xsl:value-of select="$unit"/>; top:<xsl:value-of select="$top -
$parentTop"/><xsl:value-of select="$unit"/>;</xsl:attribute>
        </xsl:if>
    </xsl:template>

    <!-- Calculate and output positioning attributes from coords especialy for characters.
    Must be recalculated to be relative to the nearest parent with coordinates. The top coordinate
    is used from the parent due to get strait lines because the bounding boxes of characters in
    HTML are equal height for large and small letters of a font. -->
    <xsl:template name="outputRelativeCharacterCoordinatesAttribute">
        <xsl:variable name="horRes">
            <xsl:choose>
                <xsl:when test="/htx:htx[(position()=1) and (@res)]">
```


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <xsl:variable name="res"><xsl:value-of
select="/htx:htx[1]/@res"/></xsl:variable>
                    <xsl:choose>
                        <xsl:when test="contains($res, ',')">
                            <xsl:value-of select="normalize-space(substring-before($res, ','))"/>
                        </xsl:when>
                        <xsl:otherwise>
                            <xsl:value-of select="$res"/>
                        </xsl:otherwise>
                    </xsl:choose>
                </xsl:when>
                <xsl:otherwise>0</xsl:otherwise>
            </xsl:choose>
        </xsl:variable>
        <xsl:variable name="vertRes">
            <xsl:choose>
                <xsl:when test="/htx:htx[(position()=1) and (@res)]">
                    <xsl:variable name="res"><xsl:value-of
select="/htx:htx[1]/@res"/></xsl:variable>
                    <xsl:choose>
                        <xsl:when test="contains($res, ',')">
                            <xsl:value-of select="normalize-space(substring-after($res, ','))"/>
                        </xsl:when>
                        <xsl:otherwise>
                            <xsl:value-of select="$res"/>
                        </xsl:otherwise>
                    </xsl:choose>
                </xsl:when>
                <xsl:otherwise>0</xsl:otherwise>
            </xsl:choose>
        </xsl:variable>

        <xsl:variable name="left">
            <xsl:variable name="coord"><xsl:value-of select="number(normalize-space(substring-
before(@coords, ',')))"/></xsl:variable>
            <xsl:choose>
                <xsl:when test="number($horRes) &gt; 0">
                    <xsl:value-of select="$coord div $horRes"/>
                </xsl:when>
                <xsl:otherwise>
                    <xsl:value-of select="$coord"/>
                </xsl:otherwise>
            </xsl:choose>
        </xsl:variable>
        <xsl:variable name="top">
            <xsl:variable name="coord"><xsl:value-of select="number(normalize-space(substring-
before(substring-after(@coords, ','), ',')))"/></xsl:variable>
            <xsl:choose>
                <xsl:when test="number($vertRes) &gt; 0">
                    <xsl:value-of select="$coord div $vertRes"/>
                </xsl:when>
                <xsl:otherwise>
                    <xsl:value-of select="$coord"/>
                </xsl:otherwise>
            </xsl:choose>
        </xsl:variable>

        <xsl:if test="(string-length($left)>0) and (string-length($top)>0)">
            <xsl:variable name="parentLeft">
                <xsl:choose>
```

--`,,```,,,,````-`-`,,`,,`,`,,`---

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

```
                    <xsl:when test="ancestor::*[(@shape='rect') and (@coords)]">
                        <xsl:variable name="coord"><xsl:value-of select="number(normalize-
space(substring-before(ancestor::*[(@shape='rect') and (@coords)][1]/@coords,
',')))"/></xsl:variable>
                        <xsl:choose>
                            <xsl:when test="number($horRes) &gt; 0">
                                <xsl:value-of select="$coord div $horRes"/>
                            </xsl:when>
                            <xsl:otherwise>
                                <xsl:value-of select="$coord"/>
                            </xsl:otherwise>
                        </xsl:choose>
                    </xsl:when>
                    <xsl:otherwise>0</xsl:otherwise>
                </xsl:choose>
            </xsl:variable>

            <xsl:variable name="parentTop">
                <xsl:choose>
                    <xsl:when test="ancestor::*[(@shape='rect') and (@coords)]">
                        <xsl:variable name="coord"><xsl:value-of select="number(normalize-
space(substring-before(substring-after(ancestor::*[(@shape='rect') and (@coords)][1]/@coords, ','),
',')))"/></xsl:variable>
                        <xsl:choose>
                            <xsl:when test="number($vertRes) &gt; 0">
                                <xsl:value-of select="$coord div $vertRes"/>
                            </xsl:when>
                            <xsl:otherwise>
                                <xsl:value-of select="$coord"/>
                            </xsl:otherwise>
                        </xsl:choose>
                    </xsl:when>
                </xsl:choose>
            </xsl:variable>

            <xsl:variable name="finalTop">
                <xsl:choose>
                    <xsl:when test="string-length($parentTop) > 0">0</xsl:when>
                    <xsl:otherwise>
                        <xsl:value-of select="$top"/>
                    </xsl:otherwise>
                </xsl:choose>
            </xsl:variable>

            <xsl:variable name="unit">
                <xsl:choose>
                    <xsl:when test="(number($horRes) &gt; 0) and (number($vertRes) &gt;
0)">in</xsl:when>
                    <xsl:otherwise>px</xsl:otherwise>
                </xsl:choose>
            </xsl:variable>
            <xsl:attribute name="style">position:absolute; left:<xsl:value-of select="$left -
$parentLeft"/><xsl:value-of select="$unit"/>; top:<xsl:value-of
select="number($finalTop)"/><xsl:value-of select="$unit"/>;</xsl:attribute>
        </xsl:if>
    </xsl:template>
</xsl:stylesheet>
```


© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

--`,,```,,,,````-`-`,,`,,`,`,,`---
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

**ICS 35.040**

Price based on 46 pages

© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale