

---

# UNIT 7    **NORMS AND GUIDELINES FOR CONTENT DEVELOPMENT**

---

## **Structure**

- 7.0 Objectives
- 7.1 Introduction
- 7.2 Need for Norms and Guidelines
- 7.3 Some General Guidelines
- 7.4 ISO/IEC Standards
  - 7.4.1 Standards Developed by ISO/ Related to Electronic Content
- 7.5 W3C Recommendations
  - 7.5.1 Web Content Accessibility Guidelines and Checkpoints
- 7.6 Electronic Text Encoding and Interchange (TEI)
- 7.7 Summary
- 7.8 Answers to Self Check Exercises
- 7.9 Keywords
- 7.10 References and Further Reading

---

## **7.0 OBJECTIVES**

---

In the previous unit you have learnt about the criteria for content development and its design aspects. The semantic and syntactic aspects of the context to be considered have also been discussed. There are guidelines and standards developed by international organisations for development of content. In this unit we will be exposing you to these standards and guidelines. After reading this unit you should be able to:

- discuss the need for guidelines for Content Development;
- focus on the general norms necessary; and
- discuss the important guidelines from bodies such as W3C and ISO.

---

## **7.1 INTRODUCTION**

---

The style of presentation of information in electronic forms is quite varied when compared to that of printed documents, which are referred to as traditional documents. The options and technologies available for adding multimedia components into documents have brought about a tremendous onslaught of documents, which are graphic intensive. While images, and other audio-visual methods are attractive and effective means of communication, there should be careful analysis and reasoning before adding on many multimedia effects. Documents today are available online and hence it becomes necessary to bear in mind the time taken for transmission as well as the traffic burden on networks. The content itself should be well organized keeping in mind the requirements and the level of understanding of the end users. There are several guidelines

given by bodies such as ISO (International Standards Organization) and W3C (WWW Consortium), which dictate the content and its presentation.

For those unfamiliar with accessibility issues pertaining to Web page design, it must be taken into account that the end users may be operating in contexts very different from the hosts environment. Also that the users

- may not be able to see, hear, move, or may not be able to process some types of information easily or at all.
- may have difficulty reading or comprehending text.
- may not have or be able to use a keyboard or mouse.
- may have a text-only screen, a small screen, or a slow Internet connection.
- may not speak or understand fluently the language in which the document is written.
- may be in a situation where their eyes, ears, or hands are busy or interfered with (e.g., driving to work, working in a loud environment, etc.).
- may have an early version of a browser, a different browser entirely, a voice browser, or a different operating system.

Content developers must consider these different situations during page design. While there are several situations to consider, each accessible design choice generally benefits the Web community as a whole and several disability groups as well. For example, by using style sheets to control font styles and eliminating the 'FONT' element, HTML authors will have more control over their pages, make those pages more accessible to people with low vision, and by sharing the style sheets, will often shorten page download times for all users. Guidelines discuss accessibility issues and provide accessible design solutions. They address typical scenarios (similar to the font style example) that may pose problems for users.

---

## **7.2 NEED FOR NORMS AND GUIDELINES**

---

Information explosion is the concern of information professionals who are trying to harness and manage information for its appropriate use. In the context of online and Internet documents 'explosion' seems to be an understatement. The information available is voluminous and the rate at which resources are added are even more daunting. If it is a matter of concern from end-users point of view it is equally of concern to the hosts who aim at having a target audience while hosting resources online. Once efforts are undertaken to develop content and host it, it should also reach its end-users. Retrieval is one of the main hassles in a distributed uncontrolled system of online resources. The search facility available through general Internet search engines is inadequate in harnessing all information available. It is not only the bulk of the information that is attributing to this. It is also the non-standard way of authoring and presentation that makes searching difficult. If each electronic resource were to have a 'title page' or the information that occurs in the title pages of the bibliographic documents, then it would be much easier to organize and locate information in an environment such as the Net. Thinking along the lines of standardized way of hosting electronic documents has lead to the formulation

of few standards and guidelines. Strictly speaking very few of these guidelines are meant for content aspects, rather, they deal more with issues in structuring of documents and presentation aspects.

### Self Check Exercise

- 1) What is the need for guidelines and norms for 'content development'?

.....

.....

.....

.....

---

## 7.3 SOME GENERAL GUIDELINES

---

Regular users of the Internet and other electronic documents often come across problems such as incompatibility of forms and formats, download time, difficulties in comprehending the content, incomplete information etc. Some general practices and ethics should be followed by content developers to assure reasonably well-written and organized resources for the end-users. According to the advice of Jakob Neilson, an ardent web content analyst, the following points should be taken care of by content developers:

- 1) To use well-tested and widely used packages in designing web pages. It is better to avoid recently released software, especially in the use of plug-ins such as Flash or PDF. A good rule is to not use upgrades immediately till they mature. Software vendors require time to fix their bugs and offer reasonably reliable versions. It also gives users time to upgrade.
- 2) As far as possible avoid Scripts in Web pages: Adding codes means there is also a risk of bugs (or faults). If scripts are a must then they should be tested and debugged on all combinations of platforms and browser versions.
- 3) It is advisable to have simple point-and-click web navigation: Special menu controls often fail, especially for users who move their mouse fast or who have motor skill impairments.
- 4) Always make the website as compact in size as possible: It is good to have spare server capacity so that your site can cope with traffic surges.
- 5) Server software must be robust: When selecting vendors, code quality must be a priority.

The above stated are guidelines (or precautions) to be followed by the host of the electronic documents. These however are not formal guidelines. The issues relating to the content and its organization are topics under consideration for arriving at standard practices by organisations such as the International Standards Organisation (ISO). The ISO 8879 deals with SGML, the Standard Mark up Language for the Web and has been adopted by Text Encoding and Interchange (TEI) initiative, which deals with guidelines for structuring of web documents.

**Self Check Exercise**

2) What are the general guidelines to be followed in developing content?

.....  
.....  
.....  
.....

---

**7.4 ISO/IEC STANDARDS**

---

For the past few years there has been an increasing awareness of the importance of the human interface to computer systems. It is no longer sufficient to provide reasonable functions at a good price unless the system can be easily and productively used. In an attempt to codify this usability and provide some consistency from interface to interface, several standards organizations are working to develop user interface software standards which they perceive will provide benefits such as:

- providing enhanced usability which results in increased satisfaction and productivity,
- providing some assurance of the users well being and lack of frustration,
- providing needed consistency within and between systems and applications,
- helping software procurers in their selection of acceptable products and
- helping to protect the employer from complaints and possible litigation.

**Self Check Exercise**

3) What are the perceived benefits of adhering to standards?

.....  
.....  
.....  
.....

**7.4.1 Standards Developed by ISO/ Related to Electronic Content**

To date there is no set of standards as such for the content of electronic documents but related guidelines deal with the issues of making user friendly products, accessibility of web documents and standards dealing with providing bibliographic details to electronic content.

**7.4.1.1 Standards Developed by ISO/TC 46/SC9 for Bibliographic References Including Electronic Documents**

- *ISO 690- 2:1997*, Information and documentation — Bibliographic references — Part 2: Electronic documents or parts thereof. *First edition.*
- *ISO 999:1996*, Information and documentation - Guidelines for the content, organization and presentation of indexes. *Second edition.*

- *ISO 3297:1998*, Information and documentation - International Standard Serial Number (ISSN). *Third edition*.
- *ISO 3901:2001*, Information and documentation - International Standard Recording Code (ISRC). *Second edition*.
- *ISO 5963:1985*, Documentation - Methods for examining documents, determining their subjects, and selecting indexing terms. *First edition*.
- *ISO 10324:1997*, Information and documentation - Holdings statements - Summary level. *First edition*.

#### **7.4.1.2 Standards Related to Design and Hosting of Electronic Documents with Multimedia Components**

These are under consideration and formulation stage by ISO.

**ISO 9241:** Ergonomic principles for visual display terminals.

**ISO/IEC 10741:** What happens to the cursor control when users interact with text editors.

**ISO/IEC 11581:** Usage and appropriateness of icons in the user interface.

**ISO/IEC 13251:** Collection of graphical symbols for office equipment.

**ISO 13407:** Designing user interfaces with humans in mind.

**ISO/IEC 14754:** Defines the basic gesture commands.

**ISO 14915:** Recommendations for multimedia controls and navigation.

**ISO/IEC 18019:** A standard for the design and preparation of software on user documentation.

**ISO/IEC 18035:** Icon symbols and functions for multimedia applications.

**ISO/IEC 18036:** Icon symbols and functions for World Wide Web browsers.

#### **7.4.1.2.1 Standards Related to Structuring of Web Documents Using SGML**

**ISO 8879 SGML:** Information Processing - Text and Office Systems - Standardized Generalized Markup Language (SGML), First Edition

---

## **7.5 W3C RECOMMENDATIONS**

---

The World Wide Web Consortium (W3C) was formalized on December 14, 1994. It's an independent, international organization made up of people from across the Internet and Web development communities—from individuals to representatives of major corporations like IBM, Microsoft, Netscape, and Sun Microsystems. The Consortium oversees the formal organization of HTML, as well as the various Web-related protocols and languages, including CSS, XML and XHTML. The Consortium's individual working groups focus on issues within a given technology. They make recommendations regarding the formal practices they believe should be put into use. The W3C recommendations—while based on those of the ISO—are not true standards. Web markup languages aren't ISO standards, but recommendations are often referred to as standards.

## 7.5.1 Web Content Accessibility Guidelines

The W3C guidelines are meant for Web Content Developers, both the authors and also the website architects. The guidelines include suggestions for content as well as its presentation on the Web. Though these guidelines basically deal with accessibility issues and are useful to the physically challenged, they also make the content more available to users, irrespective of which browser they are using or the PC they are using to view the pages. The recommendations also give guidelines for using multimedia components in the pages.

The guidelines are accompanied by a list of check points, which may be used for cross verification and check list for web content developers. The guidelines enlisted below are part of the W3C “Web Content Accessibility Guidelines 1.0”. The accompanying checklist of points for each guideline may be used to review a page or site for accessibility. For each checkpoint the web developer may indicate whether the checkpoint has been satisfied, has not been satisfied, or is not applicable.

### Priorities

Each checkpoint has a priority level assigned by the working group based on the checkpoint’s impact on accessibility.

#### [Priority 1]

A Web content developer **must** satisfy this checkpoint. Otherwise, one or more groups will find it impossible to access information in the document. Satisfying this checkpoint is a basic requirement for some groups to be able to use Web documents.

#### [Priority 2]

A Web content developer **should** satisfy this checkpoint. Otherwise, one or more groups will find it difficult to access information in the document. Satisfying this checkpoint will remove significant barriers to accessing Web documents.

#### [Priority 3]

A Web content developer **may** address this checkpoint. Otherwise, one or more groups will find it somewhat difficult to access information in the document. Satisfying this checkpoint will improve access to Web documents.

Some checkpoints specify a priority level that may change under certain (indicated) conditions.

#### *Guideline 1: Provide equivalent alternatives to auditory and visual content.*

This guideline emphasizes the importance of providing text equivalents of non-text content (images, pre-recorded audio, video). In cases where one cannot use images, movies, sounds, applets, etc. directly, still pages that include equivalent information to the visual or auditory content may be used. The equivalent information must serve the same purpose as the visual or auditory content. Thus, a text equivalent of an image of an upward arrow that links to a table of contents could be “Go to table of contents”. In some cases, an equivalent should also describe the appearance of visual content (e.g., for complex charts, billboards, or diagrams) or the sound of auditory content.

## Checkpoints

Provide a text equivalent for every non-text element (e.g., via “alt”, “longdesc”, or in element content). This includes: images, graphical representations of text (including symbols), image map regions, animations (e.g., animated GIFs), applets and programmatic objects, ascii art, frames, scripts, images used as list bullets, spacers, graphical buttons, sounds (played with or without user interaction), stand-alone audio files, audio tracks of video, and video. (Priority 1)

### **If multimedia is used:**

- until user agents can automatically read aloud the text equivalent of a visual track, provide an auditory description of the important information of the visual track of a multimedia presentation. (Priority 1)
- for any time-based multimedia presentation (e.g., a movie or animation), synchronize equivalent alternatives (e.g., captions or auditory descriptions of the visual track) with the presentation. (Priority 1)
- until user agents render text equivalents for client-side image map links, provide redundant text links for each active region of a client-side image map. (Priority 1)

### **Guideline 2. Don't rely on color alone.**

It must be ensured that text and graphics are understandable when viewed without colour. If colour alone is used to convey information, people who cannot differentiate between certain colours and users with devices that have non-colour or non-visual displays will not receive the information. When foreground and background colours are too close to the same hue, they may not provide sufficient contrast when viewed using monochrome displays or by people with different types of colour deficits.

## Checkpoints

- Ensure that all information conveyed with colour is also available without color, for example from context or markup.(Priority 1)
- Ensure that foreground and background colour combinations provide sufficient contrast when viewed by someone having color deficits or when viewed on a black and white screen. (Priority 2 for images, Priority 3 for text).

### **Guideline 3. Use markup and style sheets and do so properly.**

Mark up documents with the proper structural elements. Control presentation with style sheets rather than with presentation elements and attributes. Using markup improperly and not adhering to specifications hinders accessibility. Misusing markup for a presentation effect (e.g., using a table for layout or a header to change the font size) makes it difficult for users with specialized software to understand the organization of the page or to navigate through it. Furthermore, using presentation markup rather than structural markup to convey structure (e.g., constructing what looks like a table of data with an HTML PRE element) makes it difficult to render a page intelligibly to other devices. Content developers may be tempted to use (or misuse) constructs that achieve a desired formatting effect on older browsers. They must be aware that these practices cause accessibility problems and must consider whether the formatting

effect is so critical as to warrant making the document inaccessible to some users.

At the other extreme, content developers must not sacrifice appropriate markup because a certain browser or assistive technology does not process it correctly. For example, it is appropriate to use the TABLE element in HTML to mark up tabular information even though some older screen readers may not handle side-by-side text correctly. Using TABLE correctly and creating tables that transform gracefully makes it possible for software to render tables other than as two-dimensional grids.

### **Checkpoints**

- When an appropriate markup language exists, use markup rather than images to convey information. (Priority 2)
- Create documents that validate published formal grammars. (Priority 2)
- Use style sheets to control layout and presentation. (Priority 2)
- Use relative rather than absolute units in markup language attribute values and style sheet property values. (Priority 2)
- Use header elements to convey document structure and use them according to specification. (Priority 2)
- Mark up lists and list items properly. (Priority 2)
- Mark up quotations. Do not use quotation markup for formatting effects such as indentation. (Priority 2)

### **Guideline 4. Clarify natural language usage**

Use markup that facilitates pronunciation or interpretation of abbreviated or foreign text.

When content developers mark up natural language changes in a document, speech synthesizers and braille devices can automatically switch to the new language, making the document more accessible to multilingual users. Content developers should identify the predominant natural language of a document's content (through markup or HTTP headers). Content developers should also provide expansions of abbreviations and acronyms. In addition to helping assistive technologies, natural language markup allows search engines to find key words and identify documents in a desired language. Natural language markup also improves readability of the Web for all people, including those with learning disabilities and cognitive and physical disabilities.

### **Checkpoints**

- Clearly identify changes in the natural language of a document's text and any text equivalents (e.g., captions). (Priority 1)
- Specify the expansion of each abbreviation or acronym in a document where it first occurs. (Priority 3)
- Identify the primary natural language of a document. (Priority 3)



**Guideline 5. Create tables that transform gracefully**

Ensure that tables have necessary markup to be transformed by accessible browsers and other user agents. Tables should be used to mark up truly tabular information (“data tables”). Content developers should avoid using them to lay out pages (“layout tables”). Tables for any use also present special problems to users of screen readers. Some user agents allow users to navigate among table cells and access header and other table cell information. Unless marked-up properly, these tables will not provide user agents with the appropriate information.

**Checkpoints**

**If tables are used**

- for data tables, identify row and column headers. (Priority 1)
- for data tables that have two or more logical levels of row or column headers, use markup to associate data cells and header cells. (Priority 1)
- do not use tables for layout unless the table makes sense when linearized. Otherwise, if the table does not make sense, provide an alternative equivalent (which may be a linearized version). (Priority 2)
- if a table is used for layout, do not use any structural markup for the purpose of visual formatting. (Priority 2)
- provide summaries for tables. (Priority 3)
- provide abbreviations for header labels. (Priority 3)

**Guideline 6. Ensure that pages featuring new technologies transform gracefully.**

Ensure that pages are accessible even when newer technologies are not supported or are turned off. Although content developers are encouraged to use new technologies that solve problems raised by existing technologies, they should know how to make their pages still work with older browsers and people who choose to turn off features. One example is the use of frames. When frames are used a note should also be given as to which version of available browser should be used to view the page.

**Checkpoints**

- Organize documents so they may be read without style sheets. For example, when an HTML document is rendered without associated style sheets, it must still be possible to read the document. (Priority 1)
- Ensure that equivalents for dynamic content are updated when the dynamic content changes. (Priority 1)

**If applets and scripts are used**

- Ensure that pages are usable when scripts, applets, or other programmatic objects are turned off or not supported. If this is not possible, provide equivalent information on an alternative accessible page. (Priority 1)
- Ensure that dynamic content is accessible or provide an alternative presentation or page. (Priority 2)

- For scripts and applets, ensure that event handlers are input device-independent. (Priority 2)

***Guideline 7. Ensure user control of time-sensitive content changes.***

Some people with cognitive or visual disabilities are unable to read moving text quickly enough or at all. Movement can also cause such a distraction that the rest of the page becomes unreadable. Screen readers are unable to read moving text. When using utilities like web casters which use push technology to send information to the screens, they offer many features like giving running text (tickers) and various other animations like screen bursts to present information. But these features may distract users in their routine work. In the design of Web pages care should be taken to use animations and moving text so that either users could control the movements or the speed should be set such that it is convenient for the users.

**Checkpoints**

- Until user agents allow users to control flickering, avoid causing the screen to flicker. (Priority 1)
- Until user agents allow users to control blinking, avoid causing content to blink (i.e., change presentation at a regular rate, such as turning on and off). (Priority 2)
- Until user agents provide the ability to stop the refresh, do not create periodically auto-refreshing pages. (Priority 2)
- Until user agents provide the ability to stop auto-redirect, do not use markup to redirect pages automatically. Instead, configure the server to perform redirects. (Priority 2)

**If applets and scripts are used**

- Until user agents allow users to freeze moving content, avoid movement in pages. (Priority 2)

***Guideline 8. Ensure direct accessibility of embedded user interfaces.***

Ensure that the user interface follows principles of accessible design: device-independent access to functionality, keyboard operability, etc. When an embedded object has its “own interface”, the interface - like the interface to the browser itself - must be accessible. If the interface of the embedded object cannot be made accessible, an alternative accessible solution must be provided.

**Checkpoint**

- Make programmatic elements such as scripts and applets directly accessible or compatible with assistive technologies (Priority 1 if functionality is important and not presented elsewhere, otherwise Priority 2.)

***Guideline 9. Design for device-independence.***

Use features that enable activation of page elements via a variety of input devices. Device-independent access means that the user may interact with the user agent or document with a preferred input (or output) device - mouse, keyboard,

voice, etc. If, for example, a form control can only be activated with a mouse or other pointing device, someone who is using the page without sight, with voice input, or with a keyboard or who is using some other non-pointing input device will not be able to use the form. For example, providing text equivalents for image maps or images used as links makes it possible for users to interact with them without a pointing device.

### **Checkpoint**

- Provide client-side image maps instead of server-side image maps except where the regions cannot be defined with an available geometric shape. (Priority 1)

### **If you use applets and scripts**

- Ensure that any element that has its own interface can be operated in a device-independent manner. (Priority 2)
- For scripts, specify logical event handlers rather than device-dependent event handlers. (Priority 2)
- Create a logical tab order through links, form controls, and objects. (Priority 3)
- Provide keyboard shortcuts to important links (including those in client-side image maps), form controls, and groups of form controls. (Priority 3)

### ***Guideline 10. Use interim solutions.***

Use interim accessibility solutions so that assistive technologies and older browsers will operate correctly. For example, older browsers do not allow users to navigate to empty edit boxes. Older screen readers read lists of consecutive links as one link. These active elements are therefore difficult or impossible to access. Also, changing the current window or popping up new windows can be very disorienting to users who cannot see that this has happened.

*Note:* Web Content Guidelines: Working Group considers them to be valid and necessary to Web accessibility as of the publication of this document. However, the Working Group does not expect these checkpoints to be necessary in the future, once Web technologies have incorporated anticipated features or capabilities.

### **Checkpoint**

- Until user agents allow users to turn off spawned windows, do not cause pop-ups or other windows to appear and do not change the current window without informing the user. (Priority 2)

### **If forms are used**

- Until user agents support explicit associations between labels and form controls, for all form controls with implicitly associated labels, ensure that the label is properly positioned. (Priority 2)
- Until user agents (including assistive technologies) render adjacent links distinctly, include non-link, printable characters (surrounded by spaces) between adjacent links. (Priority 3)

### **If tables are used**

- Until user agents (including assistive technologies) render side-by-side text correctly, provide a linear text alternative (on the current page or some other) for all tables that lay out text in parallel, word-wrapped columns.

### **If forms are used**

- Until user agents handle empty controls correctly, include default, placeholder characters in edit boxes and text areas. (Priority 3)

### ***Guideline 11. Use W3C technologies and guidelines.***

Use W3C technologies (according to specification) and follow accessibility guidelines. Where it is not possible to use a W3C technology, or doing so results in material that does not transform gracefully, provide an alternative version of the content that is accessible.

The current guidelines recommend W3C technologies (e.g., HTML, CSS, etc.) for several reasons:

- W3C technologies include “built-in” accessibility features.
- W3C specifications undergo early review to ensure that accessibility issues are considered during the design phase.
- W3C specifications are developed in an open, industry consensus process.

Many non-W3C formats (e.g. Shockwave, etc.) require viewing with either plugins or stand-alone applications. Often, these formats cannot be viewed or navigated with standard user agents. Avoiding non-W3C and non-standard features (proprietary elements, attributes, properties, and extensions) will tend to make pages more accessible to more people using a wider variety of hardware and software. When inaccessible technologies (proprietary or not) must be used, equivalent accessible pages must be provided.

Even when W3C technologies are used, they must be used in accordance with accessibility guidelines. When using new technologies, ensure that they transform gracefully.

**Note.** Converting documents (from PDF, PostScript, RTF, etc.) to W3C markup languages (HTML, XML) does not always create an accessible document. Therefore, validate each page for accessibility and usability after the conversion process. If a page does not readily convert, either revise the page until its original representation converts appropriately or provide an HTML or plain text version.

### **Checkpoints**

#### **If all else fails**

- If, after best efforts, you cannot create an accessible page, provide a link to an alternative page that uses W3C technologies, which is accessible, has equivalent information (or functionality), and is updated as often as the inaccessible (original) page. (Priority 1)

- Use W3C technologies when they are available and appropriate for a task and use the latest versions when supported. (Priority 2)
- Avoid deprecated features of W3C technologies. (Priority 2)
- Provide information so that users may receive documents according to their preferences (e.g., language, content type, etc.) (Priority 3)

**Guideline 12.** *Provide context and orientation information.*

Provide context and orientation information to help users understand complex pages or elements. Grouping elements and providing contextual information about the relationships between elements can be useful for all users. Complex relationships between parts of a page may be difficult for people with cognitive disabilities and people with visual disabilities to interpret.

**Checkpoints**

**If frames are used**

- Title each frame to facilitate frame identification and navigation. (Priority 1)
- Divide large blocks of information into more manageable groups, where natural and appropriate. (Priority 2)
- Describe the purpose of frames and how frames relate to each other if it is not obvious by frame titles alone. (Priority 2)

**If forms are used**

- Associate labels explicitly with their controls. (Priority 2)

**Guideline 13.** *Provide clear navigation mechanisms.*

Provide clear and consistent navigation mechanisms - orientation information, navigation bars, a site map, etc. - to increase the likelihood that a person will find what they are looking for at a site.

**Checkpoints**

- Clearly identify the target of each link. (Priority 2)
- Provide metadata to add semantic information to pages and sites. (Priority 2)
- Provide information about the general layout of a site (e.g., a site map or table of contents). (Priority 2)
- Use navigation mechanisms in a consistent manner. (Priority 2)
- Provide navigation bars to highlight and give access to the navigation mechanism. (Priority 3)
- Group related links, identify the group (for user agents), and, until user agents do so, provide a way to bypass the group. (Priority 3)
- If search functions are provided, enable different types of searches for different skill levels and preferences. (Priority 3)

- Place distinguishing information at the beginning of headings, paragraphs, lists, etc. (Priority 3)
- Provide information about document collections (i.e., documents comprising multiple pages.). (Priority 3)
- Provide a means to skip over multi-line ASCII art. (Priority 3)

**Guideline 14.** *Ensure that documents are clear and simple.*

Ensure that documents are clear and simple so they may be more easily understood. Consistent page layout, recognizable graphics, and easy to understand language benefit all users. In particular, they help people with cognitive disabilities or who have difficulty reading.

Using clear and simple language promotes effective communication. Using clear and simple language also benefits people whose first language differs from your own, including those people who communicate primarily in sign language.

### **Checkpoints**

- Use the clearest and simplest language appropriate for a site's content. (Priority 1)
- Supplement text with graphic or auditory presentations where they will facilitate comprehension of the page. (Priority 3)
- Create a style of presentation that is consistent across pages. (Priority 3)

---

## **7.6 ELECTRONIC TEXT ENCODING AND INTERCHANGE (TEI)**

---

The Text Encoding Initiative (TEI) is an international cooperative research effort, the goal of which is to define a set of generic Guidelines for the representation of textual materials in electronic form. The project was sponsored and organized by three leading professional associations in the field: the Association for Computational Linguistics (ACL), the Association for Literary and Linguistic Computing (ALLC) and the Association for Computing and the Humanities (ACH). It has been funded primarily by the US National Endowment for the Humanities and by the European Union 3rd framework Programme for Linguistic Research and Engineering, but also with grants from the Mellon Foundation and from the Canadian Social Sciences and Humanities Research Council and the donation of time and expertise by the many members of the wider research community who have served on the TEI's Working Committees and Working Groups.

The TEI is strongly interested in text. It should not be understood literally as only text though. It is by no means confined to the use of electronic text as a stage in the production of paper documents. The TEI is equally concerned with both textual and non-textual resources in electronic form, whether as constituents of a research database or components of non-paper publications.

With the many different options available to publish, today information can be in any form. However information in any package is valuable, irrespective of

its physical form. As technology emerges, which is genuinely adequate to the task of integrating text, graphics and audio into a seamless information-bearing resource, the importance of an integrated vision becomes more apparent. TEI scheme enhances the content by facilitating the integration of multimedia technology into electronic resources. This is done by providing a description of information that is independent of realization or media of the resource.

Electronic resources are very heterogeneous in nature. Therefore the content is also diverse material, such as in newspapers, books, office memoranda, play scripts, publicity brochures, letters and diaries, transcribed lectures and interviews, TV and radio broadcasts, and unscripted conversations are integrated into a single body of material. Research needs impose that this integration be carried out with minimal loss of information, and at the same time with minimal complexity: in any case, the resulting 'text' is far removed from the conventional notion of a printed work. Electronic texts contain *markup* or encoding, which makes explicit, various features of the text, so that they can be efficiently processed. Printed texts adopt a variety of similarly-motivated conventions (use of typeface, organization of the carrier medium etc), but these are not so readily processible as the tags of a formal markup scheme.

TEI project is concerned with two things basically:

- **what** textual features should be encoded (i.e. made explicit) in an electronic text and
- **how** that encoding should be represented for loss-free, platform-independent, interchange.

The TEI guidelines mainly deal with any text in electronic form. They are meant for facilitating text processing. TEI explicitly represents the different features of text, which in turn makes it easy to manipulate the processing through computer programs. A set of tags or markers are specified, which may be inserted in the electronic representation of the text, in order to mark the text structure and other textual features of interest. Without such explicit markers, many important features remain difficult to locate by mechanical means such as computer programs, and thus difficult to process effectively. The process of inserting such explicit markers for implicit textual features is often called 'markup' or 'tagging', and the term *encoding scheme* or *markup language* denotes the rules, which govern the use of markup in a set of encodings.

The Guidelines formulated are intended for use in interchange between individuals and research groups using different programs and computer systems over a broad range of applications. Since they contain an inventory of the features most often found useful for text processing, the Guidelines also provide help to those creating texts in electronic form. They can also be used for the local storage of text that is to be processed with multiple software packages requiring different input formats.

The Guidelines apply to texts in any natural language, of any date, in any literary genre or text type, without restriction on form or content. They treat both continuous materials ('running text') and discontinuous materials such as dictionaries and linguistic corpora. Though principally directed to the needs of the scholarly research community, the Guidelines are not restricted to esoteric academic applications. They should also be useful for librarians who maintain

and document electronic materials, as well as for publishers and others creating or distributing electronic texts. Although they focus on problems of representing in electronic form texts that already exist in traditional media, these Guidelines should also be useful for the creation of electronic texts. They are adequate to, but not limited by, existing practices.

Electronic resources in the global networked environment has brought new challenges in information processing. With the possibilities of access and usage in diverse locations, standardization has become a major issue. Standardization of the way in which information is stored and represented has become a key in activities like generating secondary services and information locating tools. The main advantages of standarization of representation facilitates the following:

- Re- usability of Information once created
- True inter-operability with respect to different platforms and applications and multiple languages.
- Global networking and seamless integration of resources

Standardizing text representation also helps development of new applications while maintaining upward compatibility between the newer versions and the old ones. By standardizing at the level of text representation, it is hoped to retain the flexibility needed to develop new applications, while ensuring that old ones continue to function. TEI *Guidelines* aim to provide not a set of normative rules for particular applications, but rather a modular and extensible framework, within which particular application-specific norms can be defined.

**Self Check Exercise**

4) What is the basic objective of Text encoding?

.....

.....

.....

.....

---

**7.7 SUMMARY**

---

The basic objective of content development is, it should reach target audience. It should be comprehended by the end users. To achieve successful and meaningful transfer of ideas there are basically two factors when electronic sources are considered. One is the technological compatibility. This emphasizes on using the mainstream applications which are widely used and available. And other is regarding the content itself and the ideas contained within. While no particular guideline deals with particularly ‘thought content’ the guidelines described in this chapter generally give direction to design of electronic documents, structuring and presentation of information.

---

**7.8 ANSWERS TO SELF CHECK EXERCISES**

---

1) It is important to follow norms and guidelines as standardization increases the accessibility to users. Moreover the searches on Internet could be made more precise by following standards.



- 2) While using the electronic documents often the user faces problems with forms and formats. Generally the concepts should be made clear in the content and there should be thought flow and continuity in ideas. The underlining documents should be well structured and presentation aspects such as using bold fonts, underlining important points and use of colour should be taken care of.
- 3) Adhering to standards ensure 'quality' in products. In the parlance of electronic and online resources, standards help in providing enhanced usability which results in increased satisfaction and productivity and also provide assurance of the users well being in using the products. They provide consistency within and between systems and applications. The standards also help software customers in their selection of acceptable products. They protect the employer from complaints and possible litigation.
- 4) Text encoding deals with the structuring of Web documents using markup tags. The guidelines for text encoding are concerned with both what textual features should be encoded (i.e. made explicit) in an electronic text, and how that encoding should be represented for loss-free, platform-independent, interchange.

---

## 7.9 KEYWORDS

---

- Accessible** : Content is accessible when it can be used by someone with a disability.
- Applet** : A program inserted into a Web page.
- Braille** : Braille uses six raised dots in different patterns to represent letters and numbers to be read by people who are blind with their fingertips.
- Content developer** : Someone who authors Web pages or designs Web sites.
- Device independent** : Users must be able to interact with a user agent (and the document it renders) using the supported input and output devices of their choice and according to their needs. Input devices may include pointing devices, keyboards, braille devices, head wands, microphones, and others. Output devices may include monitors, speech synthesizers, and braille devices.
- Element** : An element is a syntactic construct and more generally means a type of content (such as video or sound) or a logical construct (such as a header or list). An element that causes text characters to be part of the document is called a text element.
- Image** : A graphical presentation.
- Link text** : The rendered text content of a link.

- Natural Language** : Spoken, written, or signed human languages such as French, Japanese, American Sign Language, and braille. The natural language of content may be indicated with the “lang” attribute in HTML and the “xml:lang” attribute in XML.
- Navigation Mechanism** : A navigation mechanism is any means by which a user can navigate a page or site. Some typical mechanisms include:
- Navigation Bars** : A navigation bar is a collection of links to the most important parts of a document or site.
- Site Maps** : A site map provides a global view of the organization of a page or site.
- Tables of Contents** : A table of contents generally lists (and links to) the most important sections of a document.
- Style Sheets** : A style sheet is a set of statements that specify presentation of a document. Style sheets may have three different origins: they may be written by content providers, created by users, or built into user agents. In CSS, the interaction of content provider, user, and user agent style sheets is called the *cascade*.
- Tabular Information** : When tables are used to represent logical relationships among data - text, numbers, images, etc., that information is called “tabular information” and the tables are called “data tables”. The relationships expressed by a table may be rendered visually (usually on a two-dimensional grid), aurally (often preceding cells with header information), or in other formats.
- User Agent** : Software to access Web content, including desktop graphical browsers, text browsers, voice browsers, mobile phones, multimedia players, plug-ins, and some software assistive technologies used in conjunction with browsers such as screen readers, screen magnifiers, and voice recognition software.

---

## 7.10 REFERENCES AND FURTHER READING

---

Jakob Neilson. Content Creation for Average People  
<http://www.useit.com/alertbox/20001001.html> [as on 12/01/2002]

ISO Standards search page. [http://www.iso.org/iso/en/standards\\_Search.StandardsQueryForm](http://www.iso.org/iso/en/standards_Search.StandardsQueryForm) [as on 17/01/2002]

ISO standard for Information and Documentation- Identification and Description  
<http://www.nlc-bnc.ca/iso/tc46sc9/> [as on 21/12/2001]

Web Content Accessibility Guidelines 1.0. W3C Recommendation 5-May-1999. <http://www.w3.org/TR/WAI-WEBCONTENT/> (Status: This document has been reviewed by W3C Members and other interested parties and has been endorsed by the Director as a W3C Recommendation. It is a stable document and may be used as reference material or cited as a normative reference from another documents [as on 17/01/2002])

Lou, Burnard. An Introduction to the Text Encoding Initiative. *Oxford University Computing Services* July 1995. <http://www.hcu.ox.ac.uk/TEI/Papers/J31/WHAT.htm>. [as on 16/02/2002]

IBM ease of use ISO IEC standards. At [http://www-3.ibm.com/ibm/easy/eou\\_ext.nsf/Publish/583](http://www-3.ibm.com/ibm/easy/eou_ext.nsf/Publish/583) [as on 13/12/2001]

Sperberg -McQueen, C.M., and Lou, Burnard. Guidelines for Electronic Text Encoding and Interchange. Text Encoding Initiative Chicago, Oxford, 1994.

Neilson, Jakob: Poor Code Quality Contaminates Users' Conceptual Models <http://www.useit.com/alertbox/20011028.html> [as on 14/12/2001]

Tufted, Edward R.. *Envisioning Information*. Graphics Press 1990.

Workman, Richard Saul. *Information Architects*. Graphics Press. 1996.

Hackos, JoAnn. *Content Management for Dynamic Delivery*. John Wiley & Sons. New York. 2002.

Goldfarb, Charles. *SGML Handbook*. IBM Almaden Research Center, San Jose. 1992.