
Software Documentation

Ian Sommerville
Lancaster University, UK

Introduction

All large software development projects, irrespective of application, generate a large amount of associated documentation. For moderately sized systems, the documentation will probably fill several filing cabinets; for large systems, it may fill several rooms. A high proportion of software process costs is incurred in producing this documentation. Furthermore, documentation errors and omissions can lead to errors by end-users and consequent system failures with their associated costs and disruption. Therefore, managers and software engineers should pay as much attention to documentation and its associated costs as to the development of the software itself.

The documents associated with a software project and the system being developed have a number of associated requirements:

1. They should act as a communication medium between members of the development team.
2. They should be a system information repository to be used by maintenance engineers.
3. They should provide information for management to help them plan, budget and schedule the software development process.
4. Some of the documents should tell users how to use and administer the system.

Satisfying these requirements requires different types of document from informal working documents through to professionally produced user manuals. Software engineers are usually responsible for producing most of this documentation although professional technical writers may assist with the final polishing of externally released information.

My goals here are to describe the documentation which may be produced during the software process, to give some hints on ways of writing effective documents and to describe processes involved in producing these documents. I start by discussing different types of documentation that may be produced in a software project. I then cover the important topic of document quality and discuss document structure, documentation standards and effective writing style. Finally, I cover processes of preparing, producing and managing documents.

My focus in this paper is on documentation that is intended to be printed and so is delivered on paper or in a format such as PDF which may be viewed on a screen or locally printed. Many systems now also have associated hypertext help systems. Producing these systems requires a different set of skills from producing paper documentation and I only discuss these briefly here.

Process and Product Documentation

For large software projects, it is usually the case that documentation starts being generated well before the development process begins. A proposal to develop the system may be produced in response to a request for tenders by an external client or in response to other business strategy documents. For some types of system, a comprehensive requirements document may be produced which defines the features required and expected behavior of the system. During the development process itself, all sorts of different documents may be produced – project plans, design specifications, test plans etc.

It is not possible to define a specific document set that is required – this depends on the contract with the client for the system, the type of system being developed and its expected lifetime, the culture and size of the company developing the system and the development schedule that it expected. However, we can generally say that the documentation produced falls into two classes:

1. *Process documentation* These documents record the process of development and maintenance. Plans, schedules, process quality documents and organizational and project standards are process documentation.
2. *Product documentation* This documentation describes the product that is being developed. System documentation describes the product from the point of view of the engineers developing and maintaining the system; user documentation provides a product description that is oriented towards system users.

Process documentation is produced so that the development of the system can be managed. Product documentation is used after the system is operational but is also essential for management of the system development. The creation of a document, such as a system specification, may represent an important milestone in the software development process.

Process documentation

Effective management requires the process being managed to be visible. Because software is intangible and the software process involves apparently similar cognitive tasks rather than obviously different physical tasks, the only way this visibility can be achieved is through the use of process documentation.

Process documentation falls into a number of categories:

1. *Plans, estimates and schedules* These are documents produced by managers which are used to predict and to control the software process.
2. *Reports* These are documents which report how resources were used during the process of development.
3. *Standards* These are documents which set out how the process is to be implemented. These may be developed from organizational, national or international standards.

4. *Working papers* These are often the principal technical communication documents in a project. They record the ideas and thoughts of the engineers working on the project, are interim versions of product documentation, describe implementation strategies and set out problems which have been identified. They often, implicitly, record the rationale for design decisions.
5. *Memos and electronic mail messages* These record the details of everyday communications between managers and development engineers.

The major characteristic of process documentation is that most of it becomes outdated. Plans may be drawn up on a weekly, fortnightly or monthly basis. Progress will normally be reported weekly. Memos record thoughts, ideas and intentions which change.

Although of interest to software historians, much of this process information is of little real use after it has gone out of date and there is not normally a need to preserve it after the system has been delivered. However, there are some process documents that can be useful as the software evolves in response to new requirements.

For example, test schedules are of value during software evolution as they act as a basis for re-planning the validation of system changes. Working papers which explain the reasons behind design decisions (design rationale) are also potentially valuable as they discuss design options and choices made. Access to this information helps avoid making changes which conflict with these original decisions. Ideally, of course, the design rationale should be extracted from the working papers and separately maintained. Unfortunately this hardly ever happens.

Product documentation

Product documentation is concerned with describing the delivered software product. Unlike most process documentation, it has a relatively long life. It must evolve in step with the product which it describes. Product documentation includes user documentation which tells users how to use the software product and system documentation which is principally intended for maintenance engineers.

User Documentation

Users of a system are not all the same. The producer of documentation must structure it to cater for different user tasks and different levels of expertise and experience. It is particularly important to distinguish between end-users and system administrators:

1. End-users use the software to assist with some task. This may be flying an aircraft, managing insurance policies, writing a book, etc. They want to know how the software can help them. They are not interested in computer or administration details.

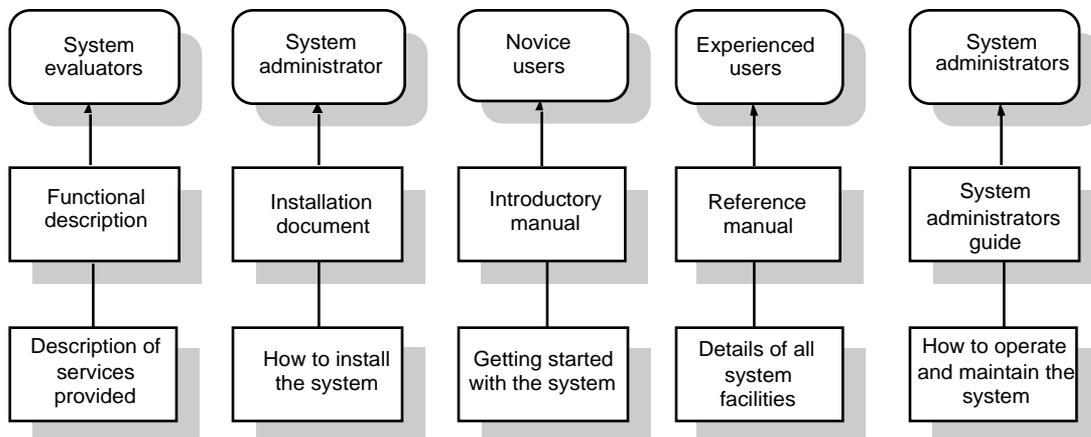


Figure 1: Different types of user documentation

2. System administrators are responsible for managing the software used by end-users. This may involve acting as an operator if the system is a large mainframe system, as a network manager if the system involves a network of workstations or as a technical guru who fixes end-users software problems and who liaises between users and the software supplier.

To cater for these different classes of user and different levels of user expertise, there are at least 5 documents (or perhaps chapters in a single document) which should be delivered with the software system (Figure1).

The *functional description* of the system outlines the system requirements and briefly describes the services provided. This document should provide an overview of the system. Users should be able to read this document with an introductory manual and decide if the system is what they need.

The *system installation document* is intended for system administrators. It should provide details of how to install the system in a particular environment. It should contain a description of the files making up the system and the minimal hardware configuration required. The permanent files which must be established, how to start the system and the configuration dependent files which must be changed to tailor the system to a particular host system should also be described. The use of automated installers for PC software has meant that some suppliers see this document as unnecessary. In fact, it is still required to help system managers discover and fix problems with the installation.

The *introductory manual* should present an informal introduction to the system, describing its 'normal' usage. It should describe how to get started and how end-users might make use of the common system facilities. It should be liberally illustrated with examples. Inevitably beginners, whatever their background and experience, will make mistakes. Easily discovered information on how to recover from these mistakes and restart useful work should be an integral part of this document.

The *system reference manual* should describe the system facilities and their usage, should provide a complete listing of error messages and should describe

how to recover from detected errors. It should be complete. Formal descriptive techniques may be used. The style of the reference manual should not be unnecessarily pedantic and turgid, but completeness is more important than readability.

A more general *system administrator's guide* should be provided for some types of system such as command and control systems. This should describe the messages generated when the system interacts with other systems and how to react to these messages. If system hardware is involved, it might also explain the operator's task in maintaining that hardware. For example, it might describe how to clear faults in the system console, how to connect new peripherals, etc.

As well as manuals, other, easy-to-use documentation might be provided. A quick reference card listing available system facilities and how to use them is particularly convenient for experienced system users. On-line help systems, which contain brief information about the system, can save the user spending time in consultation of manuals although should not be seen as a replacement for more comprehensive documentation. I briefly discuss on-line documentation in a later section in this paper.

System Documentation

System documentation includes all of the documents describing the system itself from the requirements specification to the final acceptance test plan. Documents describing the design, implementation and testing of a system are essential if the program is to be understood and maintained. Like user documentation, it is important that system documentation is structured, with overviews leading the reader into more formal and detailed descriptions of each aspect of the system.

For large systems that are developed to a customer's specification, the system documentation should include:

1. The requirements document and an associated rationale.
2. A document describing the system architecture.
3. For each program in the system, a description of the architecture of that program.
4. For each component in the system, a description of its functionality and interfaces.
5. Program source code listings. These should be commented where the comments should explain complex sections of code and provide a rationale for the coding method used. If meaningful names are used and a good, structured programming style is used, much of the code should be self-documenting without the need for additional comments. This information is now normally maintained electronically rather than on paper with selected information printed on demand from readers.
6. Validation documents describing how each program is validated and how the validation information relates to the requirements.

7. A system maintenance guide which describes known problems with the system, describes which parts of the system are hardware and software dependent and which describes how evolution of the system has been taken into account in its design.

A common system maintenance problem is ensuring that all representations are kept in step when the system is changed. To help with this, the relationships and dependencies between documents and parts of documents should be recorded in a document management system as discussed in the final part of this paper.

For smaller systems and systems that are developed as software products, system documentation is usually less comprehensive. This is not necessarily a good thing but schedule pressures on developers mean that documents are simply never written or, if written, are not kept up to date. These pressures are sometimes inevitable but, in my view, at the very least you should always try to maintain a specification of the system, an architectural design document and the program source code.

Unfortunately, documentation maintenance is often neglected. Documentation may become out of step with its associated software, causing problems for both users and maintainers of the system. The natural tendency is to meet a deadline by modifying code with the intention of modifying other documents later.

Often, pressure of work means that this modification is continually set aside until finding what is to be changed becomes very difficult indeed. The best solution to this problem is to support document maintenance with software tools which record document relationships, remind software engineers when changes to one document affect another and record possible inconsistencies in the documentation. Such a system is described by Garg and Scacchi (1990).

Document Quality

Unfortunately, much computer system documentation is badly written, difficult to understand, out-of-date or incomplete. Although the situation is improving, many organizations still do not pay enough attention to producing system documents which are well-written pieces of technical prose.

Document quality is as important as program quality. Without information on how to use a system or how to understand it, the utility of that system is degraded. Achieving document quality requires management commitment to document design, standards, and quality assurance processes. Producing good documents is neither easy nor cheap and many software engineers find it more difficult than producing good quality programs.

Document structure

The document structure is the way in which the material in the document is organized into chapters and, within these chapters, into sections and sub-sections. Document structure has a major impact on readability and usability

Component	Description
Identification data	Data such as a title and identifier that uniquely identifies the document.
Table of contents	Chapter/section names and page numbers.
List of illustrations	Figure numbers and titles
Introduction	Defines the purpose of the document and a brief summary of the contents
Information for use of the documentation	Suggestions for different readers on how to use the documentation effectively.
Concept of operations	An explanation of the conceptual background to the use of the software.
Procedures	Directions on how to use the software to complete the tasks that it is designed to support.
Information on software commands	A description of each of the commands supported by the software.
Error messages and problem resolution	A description of the errors that can be reported and how to recover from these errors.
Glossary	Definitions of specialized terms used.
Related information sources	References or links to other documents that provide additional information
Navigational features	Features that allow readers to find their current location and move around the document.
Index	A list of key terms and the pages where these terms are referenced.
Search capability	In electronic documentation, a way of finding specific terms in the document.

Figure 2: Suggested components in a software user document

and it is important to design this carefully when creating documentation. As with software systems, you should design document structures so that the different parts are as independent as possible. This allows each part to be read as a single item and reduces problems of cross-referencing when changes have to be made.

Structuring a document properly also allows readers to find information more easily. As well as document components such as contents lists and indexes, well-structured documents can be skim read so that readers can quickly locate sections or sub-sections that are of most interest to them.

The IEEE standard for user documentation (IEEE, 2001) proposes that the structure of a document should include the components shown in Figure 2. The standard makes clear that these are desirable or essential features of a document but makes clear that the ways in which these components are provided depends on the designers of the documentation. Some (such as a table of contents) are

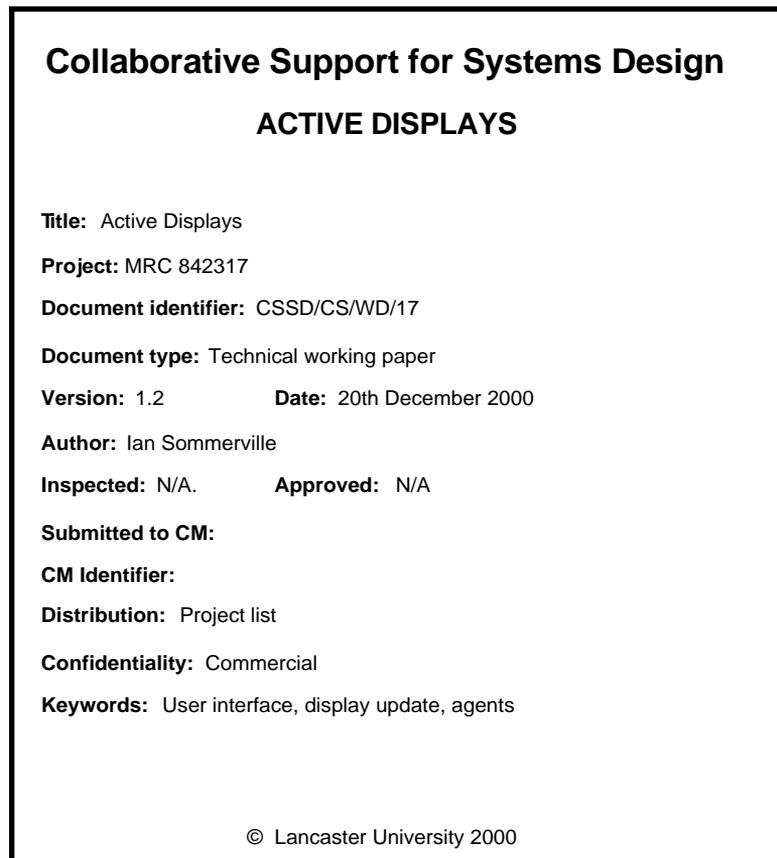


Figure 3: A sample document cover page

clearly separate sections; other components such as navigational features will be found throughout the document.

As I discuss in the next section, this IEEE standard is a generic standard and, if the use of this standard is mandated, then all of these components must be included. However, many organizations will use the standard as a guide and will not necessarily include all of the components shown in Figure 2. In such circumstances, there are some minimal structuring guidelines that I believe should always be followed:

1. All documents, however short, should have a cover page which identifies the project, the document, the author, the date of production, the type of document, configuration management and quality assurance information, the intended recipients of the document, and the confidentiality class of the document. It should also include information for document retrieval (an abstract or keywords) and a copyright notice. Figure 3 is an example of a possible front cover format.

2. Documents which are more than a few pages long should be divided into chapters with each chapter structured into sections and sub-sections. A contents page should be produced listing these chapters, sections and sub-sections. A consistent numbering scheme for chapters, sections and sub-sections should be defined and chapters should be individually page numbered (the page number should be *chapter-page*). This simplifies document change as individual chapters may be replaced without re-printing the whole document.
3. If a document contains a lot of detailed, reference information it should have an index. A comprehensive index allows information to be discovered easily and can make a badly written document usable. Without an index, reference documents are virtually useless.
4. If a document is intended for a wide spectrum of readers who may have differing vocabularies, a glossary should be provided which defines the technical terms and acronyms used in the document.

Document structures are often defined in advance and set out in documentation standards. This has the advantage of consistency although it can cause problems. Standards may not be appropriate in all cases and an unnatural structure may have to be used if standards are thoughtlessly imposed.

Documentation Standards

Documentation standards act as a basis for document quality assurance. Documents produced according to appropriate standards have a consistent appearance, structure and quality. I have already introduced the IEEE standard for user documentation in the previous section and will discuss this standard in more detail shortly. However, it is not only standards that focus on documentation that are relevant. Other standards that may be used in the documentation process are:

1. *Process standards* These standards define the process which should be followed for high-quality document production.
2. *Product standards* These are standards which govern the documents themselves.
3. *Interchange standards* It is increasingly important to exchange copies of documents via electronic mail and to store documents in databases. Interchange standards ensure that all electronic copies of documents are compatible.

Standards are, by their nature, designed to cover all cases and, consequently, can sometimes seem unnecessarily restrictive. It is therefore important that, for each project, the appropriate standards are chosen and modified to suit that particular project. Small projects developing systems with a relatively short expected lifetime need different standards from large software projects where the software may have to be maintained for 10 or more years.

Process standards

Process standards define the approach to be taken in producing documents. This generally means defining the software tools which should be used for document production and defining the quality assurance procedures which ensure that high-quality documents are produced.

Document process quality assurance standards must be flexible and must be able to cope with all types of document. In some cases, where documents are simply working papers or memos, no explicit quality checking is required. However, where documents are formal documents, that is, when their evolution is to be controlled by configuration management procedures, a formal quality process should be adopted. Figure 3 illustrates one possible process.

Drafting, checking, revising and re-drafting is an iterative process which should be continued until a document of acceptable quality is produced. The acceptable quality level will depend on the document type and the potential readers of the document.

Product standards

Product standards apply to all documents produced in the course of the software development. Documents should have a consistent appearance and, documents of the same class should have a consistent structure. Document standards are project-specific but should be based on more general organizational standards.

Examples of product standards which should be developed are:

1. *Document identification standards* As large projects typically produce thousands of documents, each document must be uniquely identified. For formal documents, this identifier may be the formal identifier defined by the configuration manager. For informal documents, the style of the document identifier should be defined by the project manager.
2. *Document structure standards* As discussed in the previous section, there is an appropriate structure for each class of document produced during a software project. Structure standards should define this organization. They should also specify the conventions used for page numbering, page header and footer information, and section and sub-section numbering.
3. *Document presentation standards* Document presentation standards define a 'house style' for documents and they contribute significantly to document consistency. They include the definition of fonts and styles used in the document, the use of logos and company names, the use of color to highlight document structure, etc.
4. *Document update standards* As a document is changed to reflect changes in the system, a consistent way of indicating these changes should be used. These might include the use of different colors of cover to indicate a new document version and the use of change bars to indicate modified or deleted paragraphs.

Document standards should apply to all project documents and to the initial drafts of user documentation. In many cases, however, user documentation has to be presented in a form appropriate to the user rather than the project and it should be recast into that form during the production process.

Interchange standards

Document interchange standards are important as more and more documents are produced in electronic format as well as or instead of on paper. For documentation that is delivered with a software system, the Adobe Portable Document Format (PDF) is now very commonly used. However, when documents are exchanged by the development team and drafts are circulated within an organization these are often in the format of whatever word processor is used (often Microsoft Word).

Assuming that the use of a standard word processor and graphical editing system is mandated in the process standards, interchange standards define the conventions for using these tools. The use of interchange standards, allows documents to be transferred electronically and re-created in their original form.

Interchange standards are more than simply an agreement to use a common version of a system for document production. Examples of interchange standards include the use of an agreed standard macro set if a text formatting system is used for document production or the use of a standard style sheet for a word processor. Interchange standards may also limit the fonts and text styles used because of differing printer and display capabilities.

The IEEE standard for user documentation

The first IEEE standard for user documentation (IEEE, 1987) was produced in 1987 and, at the time of writing, a new draft of this standard is being prepared for publication (IEEE, 2001).

Like all standards, this standard encapsulates wisdom and experience about software documentation and proposes a structure for user documentation. Using this structure as a basis, the standard discusses the content of software user documentation and proposes formatting standards for these documents.

I have already covered the documentation structure proposed by the latest version of the standard. To illustrate the formatting advice in the standard, here are some quotations from the current draft standard of good practice:

The documentation should be provided in media and formats that allow its use by those with vision, hearing or other physical limitation.

A description of how to print the electronic documentation should be included in both the electronic and the printed documentation.

Because some users cannot distinguish between colors, documentation should provide text cues rather than using colors such as red and green as the only way to convey meaning.

Warnings, cautions and notes shall be displayed in a consistent format that is readily distinguishable from ordinary text or instructional steps.

Documentation formats for user-entered commands or codes shall clearly distinguish between literals (to be input exactly as shown) and variables (to be selected by the user).

Illustrations that accompany text should appear adjacent to their first reference in the text so that the associated text and illustration can be viewed simultaneously.

You can see from these that the standard is helpful without being proscriptive and therefore different conventions used by different companies and organizations can be accommodated.

Like all standards, this documentation standard has to be adapted to the local situation where it is used. These should instantiate the advice in the standard to the local situation and define the specific structures and formats that should be used.

Writing style

Standards and quality assessment are essential if good documentation is to be produced but document quality is fundamentally dependent on the writer's ability to construct clear and concise technical prose. In short, good documentation requires good writing.

Writing documents well is neither easy nor is it a single stage process. Written work must be written, read, criticized and then rewritten until a satisfactory document is produced. Technical writing is a craft rather than a science but some broad guide-lines about how to write well are:

1. *Use active rather than passive tenses* It is better to say 'You should see a flashing cursor at the top left of the screen' rather than 'A flashing cursor should appear at the top left of the screen'.
2. *Use grammatically correct constructs and correct spelling* To boldly go on splitting infinitives (like this) and to misspell words (like misspell) irritates many readers and reduces the credibility of the writer in their eyes. Unfortunately, English spelling is not standardized and both British and American readers are sometimes irrational in their dislike of alternative spellings.
3. *Do not use long sentences which present several different facts* It is better to use a number of shorter sentences. Each sentence can then be assimilated on its own. The reader does not need to maintain several pieces of information at one time to understand the complete sentence.
4. *Keep paragraphs short* As a general rule, no paragraph should be made up of more than seven sentences. Our capacity for holding immediate information is limited. In short paragraphs, all of the concepts in the paragraph can be maintained in our short-term memory which can hold about 7 chunks of information.
5. *Don't be verbose* If you can say something in a few words do so. A lengthy description is not necessarily more profound. Quality is more important than quantity.

6. *Be precise and define the terms which you use* Computing terminology is fluid and many terms have more than one meaning. If you use terms like *module* or *process* make sure that your definition is clear. Collect definitions in a glossary.
7. *If a description is complex, repeat yourself* It is often a good idea to present two or more differently phrased descriptions of the same thing. If readers fail to completely understand one description, they may benefit from having the same thing said in a different way.
8. *Make use of headings and sub-headings.* These break up a chapter into parts which may be read separately. Always ensure that a consistent numbering convention is used.
9. *Itemize facts wherever possible.* It is usually clearer to present facts in a list rather than in a sentence. Use textual highlighting (italics or underlining) for emphasis.
10. *Do not refer to information by reference number alone* Give the reference number and remind the reader what that reference covered. For example, rather than say 'In section 1.3 ...' you should say 'In section 1.3, which described management process models, ...'

Documents should be inspected in the same way as programs. During a document inspection, the text is criticized, omissions pointed out and suggestions made on how to improve the document. In this latter respect, it differs from a code inspection which is an error finding rather than an error correction mechanism.

As well as personal criticism, you can also use grammar checkers which are incorporated in word processors. These checkers find ungrammatical or clumsy uses of words. They identify long sentences and paragraphs and the use of passive rather than active tenses. These checkers are not perfect and sometimes they use outmoded style rules or rules which are specific to one country. Nevertheless, because they often check style as you are typing, they can help identify phrases which could be improved.

On-line documentation

It is now normal to provide some on-line documentation with delivered software systems. This can range from simple 'read me' files that provide very limited information about the software through interactive hypertext-based help systems to a complete on-line suite of system documentation. Most commonly, however, hypertext-based help systems are provided. These may be based on a specialized hypertext system or may be HTML-based and rely on web browsers for access.

The main advantage with on-line documentation is, of course, its accessibility. It is not necessary for users to find manuals, there is no possibility of picking up out-of-date documentation and built-in search facilities can be used to locate information quickly.

However, on-line hypertext systems have several disadvantages which mean that they should be used to supplement rather than replace paper-based documentation. These are:

- They lack ‘browsability’ so that readers cannot easily skim through them to find the information they need. We often find it difficult to characterize the information we want from documentation although we can recognize it when we find it. Browsing is the key mechanism that we use when searching in this way. Browsing also offers opportunities for serendipitous discovery of system facilities that were unknown.
- Screens, at least with monitors normally used in 2001, have a much lower resolution than paper and hence it is more difficult and tiring to read a document on the screen rather than on paper.
- It is very easy for users to get lost in hypertext-based help systems and they consequently find it difficult to navigate to where they want to go.

When designing screen-based documentation, you should always bear these problems in mind. Consequently, although both screen-based and paper-based documents should be well-written, different designs are needed for electronic and paper documentation. Because of the differences between screens and paper, simply converting a word processor document to HTML (say) rarely produces high-quality on-line documentation.

The design of screen-based documents is a major topic in its own right and I don’t have space to discuss it here. Interested readers can find a brief introduction to this topic in Chapter 15 of my book on software engineering (Sommerville, 2001) and a much more comprehensive account in the book on HCI design by Dix et al (1998).

Document Preparation

Document preparation is the process of creating a document and formatting it for publication. Figure 3 shows the document preparation process as being split into 3 stages namely document creation, polishing and production. Modern word processing systems are now integrated packages of software tools that support all parts of this process. However, it is still the case that for the highest-quality documents, it is best to use separate tools for some preparation processes rather than the built-in word processor functions. The three phases of preparation and associated support facilities are:

1. *Document creation* The initial input of the information in the document. This is supported by word processors and text formatters, table and equation processors, drawing and art packages.
2. *Document polishing* This process involves improving the writing and presentation of the document to make it more understandable and readable. This involves finding and removing spelling, punctuation and grammatical errors, detecting clumsy phrases and removing redundancy in the text. The process may be supported by tools such as

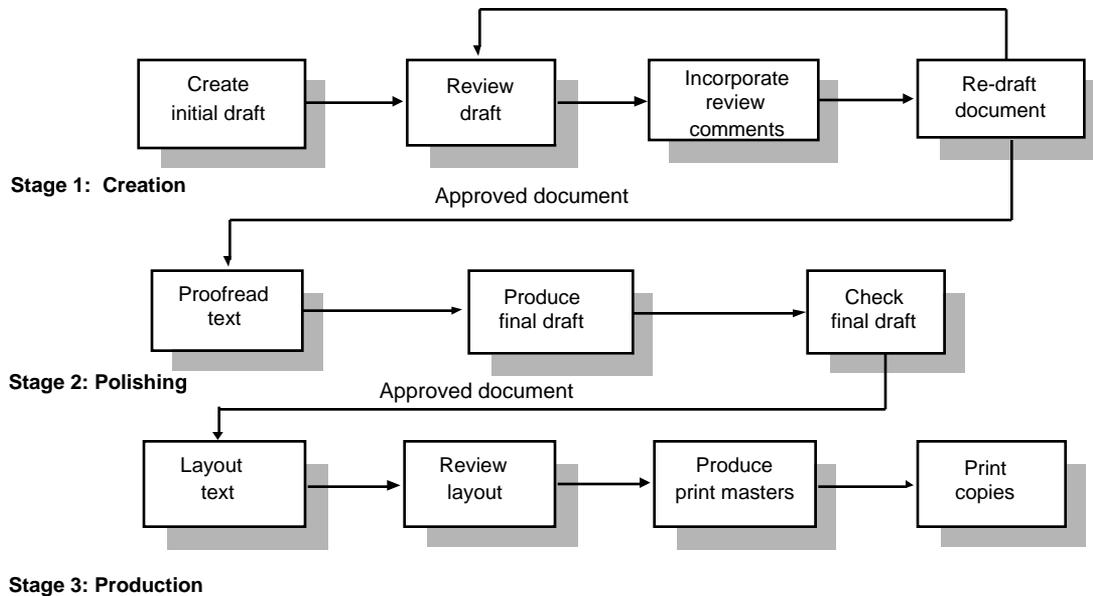


Figure 3 Stages of document preparation

on-line dictionaries, spelling checkers, grammar and style checkers and style checkers.

3. *Document production* This is the process of preparing the document for professional printing. It is supported by desktop-publishing packages, artwork packages and type styling programs.

As well as these tools to support the document production process, configuration management systems, information retrieval systems and hypertext systems may also be used to support document maintenance, retrieval and management.

Modern word processing systems are screen based and combine text editing and formatting. The image of the document on the user's terminal is, more or less, the same as the final form of the printed document. Finished layout is immediately obvious. Errors can be corrected and layout improved before printing the document. However, programmers who already use an editor for program preparation may sometimes prefer to use a separate editor and text formatting system.

Text formatting systems such as Latex interpret a layout program specified by the document writer. Layout commands (often chosen from a standard, definable command set) are interspersed with the text of the document. The text formatter processes these commands and the associated text and lays the document out according to the programmer's instructions. The distinction between these systems and word processors is illustrated in Figure 4.

Text formatting systems can look ahead at the text to be laid out so can make better layout decisions than word processing systems whose working context is more restricted. Because the commands are really a programming language,

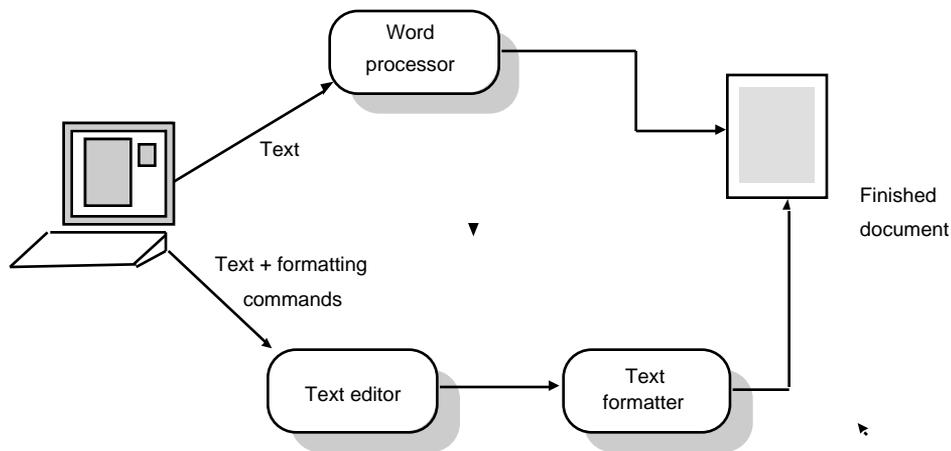


Figure 4: Word processors and text formatters

programmers often prefer them to word processors but other, non-technical users usually find them more difficult to use.

The major disadvantage of text processors, once their programming has been mastered, is that they do not provide an immediate display of the output they produce. The user must process the text (this may take several minutes) then display the output using a preview package. If an error is discovered, it cannot be fixed immediately. The original source must be modified and the preview process repeated. Thus, although they can result in higher quality documents, most users find text formatters more inconvenient than word processors.

The final stage of document production is a skilled task which, for documents with large print runs, should be left to professional printers. However, desktop publishing (DTP) systems and graphics systems that support scanning and processing photographs and artwork are now widely available. These have revolutionized document production. DTP systems partially automate the layout of text and graphics. They allow very fine-grain control over the layout and look of a document and can be used by engineers to produce finished system documentation.

The advantage of using a desktop publishing system is that the cost of producing high-quality documents is reduced because some of the steps in the production process are eliminated. Even documents which are produced in small numbers can be produced to a high standard. The disadvantage of using desktop publishing systems is that they do not automate the skills of the graphic designer. Their seductive ease-of-use means that they are accessible to unskilled users who may produce unattractive and badly designed documents.

An enormous number of documents are produced in the course of a project and these need to be managed so that the right version of the document is available when required. If a project is distributed, copies of documents will be produced and stored at different locations and it is very important to maintain a 'master file' of documents which contains the definitive versions of each document. This helps minimize a very common problem that arises when users of a

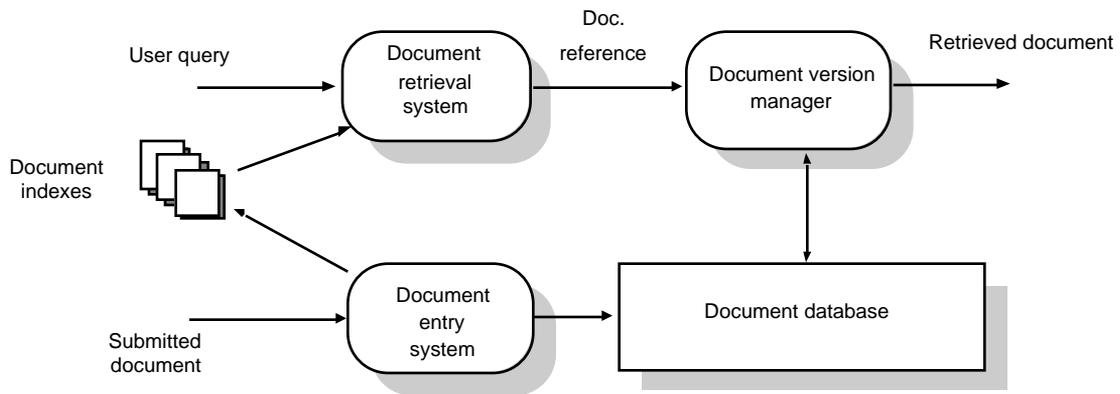


Figure 5: Document management

document make mistakes because they are not working from the current version of a document.

Each document should have a unique record and this can be used as a key in a document database record. However, retrieval by other fields such as the title and author should also be supported.

The basic problem with managing documents using a file system to store the documents and a database management system to maintain document information is that users have to be disciplined in the way they use they system. They must ensure that they check out a copy of the document from the system each time they need it rather than use a local copy on their computer or the copy that they have printed. In practice, achieving this level of discipline is difficult and errors are always likely to arise.

In very large projects, specialized document management systems may be used that integrate the storage of the documents and the maintenance of document information (Figure 5). Document management software allows related documents to be linked, maintains records of who has checked out documents, may support the compression and de-compression of document text and provides indexing and information retrieval facilities so that documents can be found. Document management systems may also include version management facilities so that different document versions may be maintained.

References

Dix, A., Finley, J., Abowd, G. and Beale, R. 1998. *Human-Computer Interaction, 2nd ed.* London: Prentice-Hall.

Garg, P. K. and Scacchi, W. 1990. 'A hypertext system to maintain software life-cycle document'. *IEEE Software*, 7 (3), 90–8.

IEEE, 1987. *IEEE Standard for Software User Documentation*, IEEE-Std1063-1987. New York: Institute of Electrical and Electronics Engineers.

IEEE, 2001. *Draft Standard for Software User Documentation*, IEEE-Std1063/D5.1. 2001. New York: Institute of Electrical and Electronics Engineers.

Sommerville, I. 2001. *Software Engineering, 6th Edition*. Harlow, UK: London: Pearson Education Ltd.

Further reading

The Art of Technical Documentation 2nd ed. 1998. K. Haramundanis, Woburn-MA: Butterworth-Heinemann.

This book is primarily aimed at technical writers and prospective technical writers and is not specific to software documentation. However, it includes good general advice on document presentation and style.

Writing for Science and Engineering: Papers, Presentations and Reports. 2000. H. Silyn-Roberts. Woburn-MA: Butterworth-Heinemann.

Again, this is a general book on documentation. It is more detailed than Haramundanis's book and includes very specific advice and checklists on structure and style.

A Style Guide for the Computer Industry. 1996, Mountain View, CA: Sun Microsystems.

This book includes specific information and advice on writing style for software and hardware documents. Includes information on writing for international audiences. It was originally developed for people involved in writing documentation for Sun software and hardware systems.

Designing and Writing Online Documentation. 2nd ed. W. Horton, New York: Wiley Publishers Inc.

This is based on a hypermedia model of on-line documentation as discussed in this paper. It includes good general advice but it is rather outdated as it predates the general use of the WWW so does not mention HTML-based systems.

Acknowledgements

This paper is a revised version of Chapter 30 from my book *Software Engineering, 4th edition*, published by Addison Wesley in 1992. Thanks to the publishers (now Pearson Education Ltd) for permission to re-publish this material.