Modern technical publishing methodologies

Introduction

First, we need to understand what a PUBLICATION is. Examples include:

- Operations Manuals. Usually subdivided into Parts A,B, C, D etc.
- Maintenance manuals.
- Minimum Equipment lists.
- Training manuals

Re-usable, retrievable chunks of information

In modern technical publishing, a PUBLICATION is a collection of discrete pieces of information. Different standards call these pieces of information different things; in DITA, they are called **topics**; in S1000D, they are called data modules. The principle is very similar.

The topics stand alone (i.e. they are said to be "context free") so that they may be reused - in different publications; in different contexts; and on different delivery devices (printed; mobile; EFBs etc.)

Minimalism

A topic consists of a title and some content. By the principle of minimalism, the content is just sufficient to describe the topic - no more, no less.

Another way to describe this is: "the topic title describes the topic, the whole topic, and nothing but the topic". Thus the topic "Engine fire in flight" does not include any information about "Engine fire on ground" because these two procedures are different and apply in different situations; they are two separate topics.

Adhering to the principle of minimalism has the benefit of making the information more readable and easier to understand.

Not like a DMS

At this point, we often hear "but surely this is what a document management system (DMS) does !

Actually, it's quite different. DMSs manage PUBLICATIONS. What DMSs do very effectively is manage *completed* publications. DMSs trace their origin back to the age-old problem of "where did I save that document?" In the DMS world, a publication may be just about anything you have stored on a disk - Microsoft® Office files (Word®; Excel®, PowerPoint® etc.); CAD designs made in AutoCad®; vector graphics generated by Adobe® Illustrator®; and also, XML files. The DMS catalogues these documents, allows users to establish taxonomic classifications and easily retrieve documents.

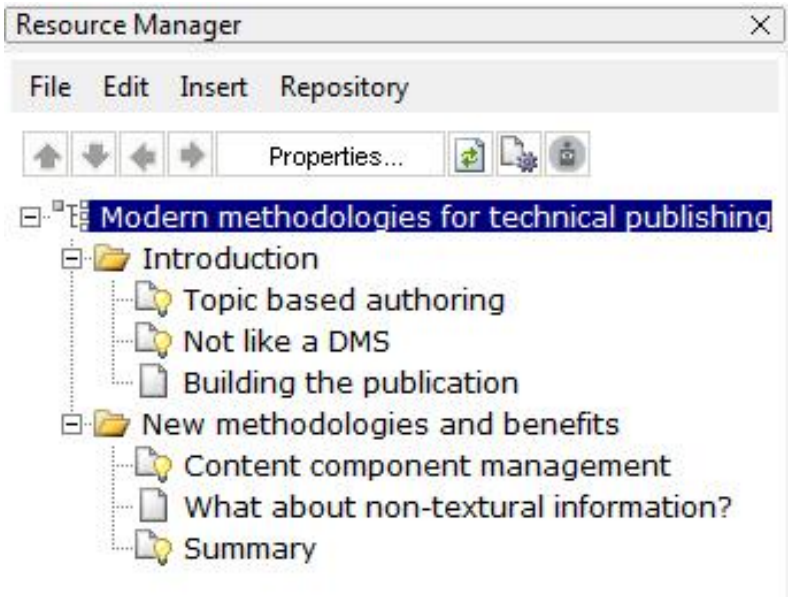
The key point about DMSs is that they manage completed publications; they do not build publications from topics of retrievable information modules. Put another way, SharePoint® and its competitors treat the publication as an 'atom' - a seemingly indivisible unit.

Building the publication

DITA maps are structures that assemble topics into PUBLICATIONS. These may be output in a wide range of formats: PDF; HTML; RTF (Word); CHM (Windows® help) etc.

A sample DITA map is shown below:

Figure 1:Representative DITA map. Individual topics may be unique to the publication being built or shared amongst many different publications.



A simple way to build publications using re-usable as well as unique content is to have a separate DITA map for each publication. This is simple to understand and maintain.

However, it is by no means the only way, as DITA provides other more sophisticated methods for reusing content. For example, *conditional publication* is provided. Suppose we have two groups of people, called *specialists* and *users*. Both groups of people need some common information, so these topics will appear in both a Specialist publication and a User publication.

We need the Specialist publication to include technical procedures and specifications that are not relevant to a User. Therefore, those topics (and illustrations etc) are marked for selective publishing in a Specialist version of the publication only.

Content component management

Content component management systems (CCMSs) address a completely different problem to that addressed by a DMS.

The CCMS addresses the problem of how we build the document.

First point to note is that the value of a CCMS is realised with complex (though not necessarily long) documents. If your entire documentation set consists of twenty documents each of half a page, with no common material, managed by one person, you probably don't need a CCMS to build them. For that matter, you won't need a DMS to manage completed publications either.

Using a DITA based system we address issues such as:

1. The topic named *Engine fire in flight* may be required in more than one PUBLICATION - for example in an Operations Manual Part B; and in a training class.

   Instead of authoring separately - or copying and pasting from one publication into the other, the topic is shared or re-used between the two (or more) publications.

2. This produces three immediate benefits:

   - Cutting out the duplicated effort means a cost saving.
   - When the topic needs to be updated, it's done ONCE; another cost saving.
   - The information is presented consistently. It will be the same in both publications. This produces a quality and safety benefit; it reduces risk.

3. Using DITA/XML, the content is separated from the style of presentation. This produces more benefits:

   - The author is freed up from having to worry about layout. Using traditional word processors, authors spend large amounts of time trying to make the publication look attractive. When the authors are highly paid engineers or flight operations staff, this makes no economic sense.
   - The author can't "tweak" the layout. This means that the layout is always consistent. The layout is defined for the organisation by an authorised person, and only changed when there are sound business reasons for doing so.
   - The output is formatted automatically, no matter how long the publication is. This is done using *machine time* which is cheap, rather than people time, which is expensive.

4. The system can be used to generate output from a single common source - the collection of DITA topics - into a wide variety of target environments, such as:

   - PDF (for printing)
   - Electronic Flight Bags (EFBs)
   - Web sites (HTML)
   - Apple® iPad®

- Android® based devices (Samsung® Galaxy Tab; Motorola® etc.)

For each target platform, the output is formatted automatically, no matter how long the publication is, using inexpensive machine time rather than people time.

In contrast to a DMS, which treats publications as atoms, a CCMS regards publications as molecules consisting of 'atoms' which are discrete topics.

A CCMS offers sophisticated management tools. Figure 2 below shows the manifest listing for a topic. This allows the topic owner to verify where the topic is being used by other publications and thereby to establish the impact of making edits to the topic. Figure 3 shows the audit of modification records. These records will contain date and time stamps plus the ID of the user committing the modification, plus any user comments made at check-out time and/or check-in time.

Figure 2: manifest listing for a topic shows which DITA maps include that topic as a reference.

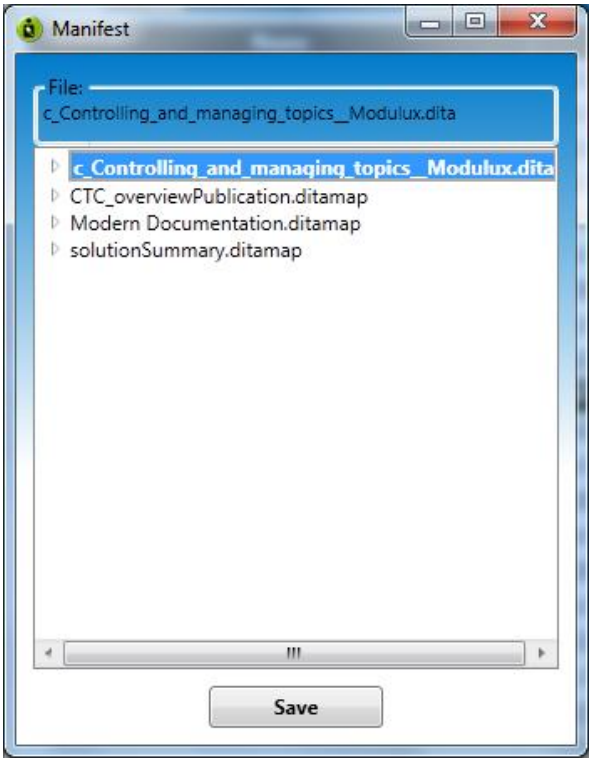
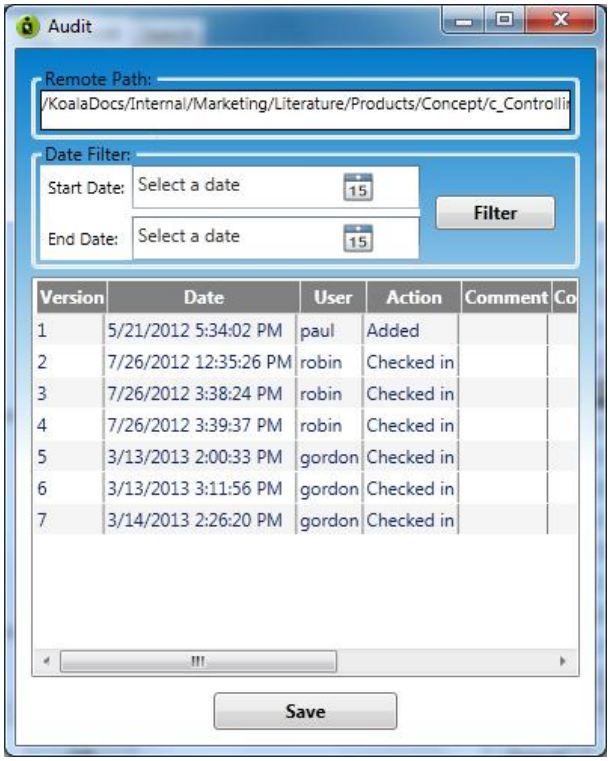


Figure 3: Audit list displays the modification record for that topic. Screenshots from Modulux® by Koala



What about non-textural information?

Illustrations

Today, most non textural information consists of illustrations or diagrams.

As in the case of text, it is vital that every user can use the same version of the illustration. This is achieved exactly in the same way as with textural topics. Instead of copying and pasting as when using a word processor, the illustration is *referenced* by every topic that needs it. A single up-to-date copy of the illustration is maintained, and is faithfully rendered in each publication.

The illustration itself may feature *hotspots* that link parts of the diagram to textural descriptions such as procedures, parts lists etc.

Illustrations built using the Scalar Vector Graphics language (SVG - an XML based language) are not pixel dependent and can thus be scaled to a very wide range of devices - even smart phones.

Audio and video segments

Multi-media components can be handled in the same way, in that they are referenced from within the DITA map. Rendering these as publications means that a suitable playing environment has to be provided. For example, if the publication is to be built as a web page, the rendering environment can make use of open-source jQuery libraries for playback.

Summary

1. CCMSs and DMSs do quite different things. DMSs manage completed *publications* and regard publications as indivisible 'atoms'. The function of the CCMS is to build the publication. CCMSs manage topics and other re-usable components and are used to build publications.

2. CCMSs are particularly useful in building large documents such as are required in most regulated industries. DMSs are useful where large numbers of manuals must be managed as completed publications.

3. CCMS's using standards such as DITA regard publications as being molecular, consisting of collections of 'atomic' components. In DITA, the basic unit of information is the *topic*. Other units of information include illustrations.

4. The advantages of topic based authoring include:

    - Cost savings - mainly by reducing duplicated effort.

    - Other costs are saved because it is much easier to correct small units of information than a very large manual of hundreds of pages. Similar savings will also be realised because proof reading is also made easier.

    - Increased consistency. This means reduced downstream costs due to incorrect information being presented, and since publications are more readable they cause less mistakes to be made. Exposure to both commercial and legal risks is reduced.

Gordon Dennis