

Planning for DITA Success: Part Two

Deploying DITA: Step-by-step

A JUSTSYSTEMS WHITEPAPER

Introducing DITA as a key element of your content strategy introduces your company to a standard that helps you deliver just-in-time content to your users across your delivery channels. DITA is a documentation revolution that merges authoring, managing, and delivering information in a new paradigm.

You can specialize DITA to meet your specific business goals.

DITA is not a product that constrains how you work; it is a standard that encapsulates the best practices of the industry allowing you to incorporate those practices and ease pain points in satisfying the need for documentation to support your products. Your documentation team will face many questions in the course of adopting the DITA standard.

As an XML standard, DITA is extensible; you can specialize DITA to meet your specific business goals and needs. Because DITA is topic-based, you may consider using a component content management system. DITA is output-agnostic, you can generate a variety of output formats from the same set of topics. It's up to your team to determine what the business needs are and how you're going to satisfy them. All this affects the success of a DITA implementation.

Congratulations on taking the time to investigate what comes before you begin writing documentation. Putting the time and energy into these preliminaries will improve your results.

Before you begin, here are some considerations that you need to take on board:

- Even though the primary stakeholder is your documentation team, you need to build and maintain collaboration between all potential stakeholders. These may be regulatory or legal advisors who have a say on what is said in your manuals and how documentation is delivered to your customers. Consider your IT department who may find themselves supporting an unfamiliar technology. Who reviews your documentation? They need to learn the new processes. Look at your business, who would benefit from being involved; who would bring value to the transition? What business goals do you have and who is responsible for those goals?
- To truly understand, recognize, and appreciate the savings brought about through the transition to DITA, document each phase and record the costs and savings achieved.
- If DITA is new to your publication team, you will need to consume training all along the process. Plan for this. The team will need to be educated on the standard, the tools, and the new processes.
- Before any DITA endeavor, even a piecemeal one, address these issues at a high level. Consider how your organization's people, processes, and technology interact in the pursuit of authoring and publishing. Which methods are effective? Which are a liability? The transition to DITA can be

Engage your stakeholders, however remote from the documentation team.

a fresh start, a chance to identify and eliminate bad habits and inefficiencies. Use this guide to making these valuable assessments and begin the journey to DITA on the right path.

There are four large steps to transitioning to a DITA-based publishing model. This paper discusses each of these steps:

1. Developing a conversion/migration strategy.
2. Selecting the appropriate tools and workflow.
3. Configuring the tools for your publishing needs.
4. Converting legacy content to DITA.
5. Going live with DITA driving your documentation publication needs.

Step One: Define Your Migration Strategy

Before undertaking a project of this magnitude, it's important that you first define your business goals. Doing this allows you to determine the value of the transition and identify the pain points that you are looking to solve. Perhaps localization or branching product lines are part of your growth path.

Understanding the business goals helps you identify the priorities for the transition to DITA authoring and what your conversion requirements will be. Are you seeking to align your content with industry standard documentation practices such as topic-based content, minimalism, or single-sourcing and reusing content intelligently? Are you simply looking to bring a new degree of consistency into your instructions for use? Is localization a pain point in your documentation budget?

To prepare for developing a content strategy that builds on the business goals, you'll need to perform an inventory of your content. The inventory is a key element in the process of selecting the path you'll take to DITA implementation. The inventory provides the information you need to understand how much work is between your current documentation and your business goals for documentation.

Content Strategy

Before you embark on the transition to a DITA publication model, you need a comprehensive and flexible content strategy. A content strategy defines how your content creation and distribution satisfies your business goals and requirements. Once you know where you're going, you combine this with your inventory, which tells you where you're coming from, to construct a migration strategy. Your migration strategy is the who does what when and how. It includes any content conversion plans.

Your inventory and your content strategy prioritize the content that will be migrated to DITA. Your conversion strategy defines how you'll go about this. Your

strategy will depend, in part, on when you plan to clean content up: before, during, or after migration to XML and DITA. Your conversion strategy, as part of your migration plan, needs to include how you'll handle metadata and tagging in the transitioned content.

Your content strategy also drives whether you specialize DITA for your long-term requirements.

Step Two: Select Your Tools

A DITA configuration consists, minimally, of DITA authoring software and a publishing mechanism. Some authoring tools provide the publishing mechanism through the DITA Open Toolkit (DITA-OT), an open-source resource, or directly through custom code. Note that some custom processing is less DITA specialization-friendly.

Beyond these two required pieces, a DITA configuration can consist of any of these optional components:

- **Content Management.** Many organizations turn to a content management system (CMS) or component content management system (CCMS) to store the topics and other content generated as part of the process of authoring documentation. A CMS or CCMS provides easier storage, version control, and search capabilities making it simple to store, use, and reuse content. Be sure to ensure your authoring and publishing environments can work in conjunction with your chosen content management system. If you specialize DITA, choose a CMS/CCMS that supports modifications to the standard.

Some systems integrate localization and workflow management.

- **Localization Management.** If you publish in multiple languages, you'll need some sort of localization management functionality to manage the relationship between the original and localized content, as well as be able to automatically reuse localized content. The complexity of this system will depend on your localization workflow.
- **Workflow Management.** A workflow management system automates the routing of information. In the case of content creation and publishing, it replaces the manual email request for a review, edit, or approval of content, with a notification that is part of a consistent, pre-defined process that is applied to each topic and map.

DITA Authoring Software

When you're introducing so much new technology to your writers, you have to take care to ensure that you take their experience into consideration. If the authoring tool is easy to use, it will meet with an enthusiastic reception. An easy-to-learn interface leads to a shorter learning curve, minimizing training costs, accelerating productivity, and resulting in quick wins that reach beyond the core documentation team.

A tool as simple as a plain text editor, like Notepad, can be used to author DITA, but a robust, full-featured, DITA-oriented XML authoring tool enhances the author experience and is a more production-worthy tool.

The tool, itself, has to support your implementation of the DITA standard; if you've specialized or if you're using anything other than a file system for topic storage, you need to have software that works as seamlessly as possible.

Consider the following when choosing a DITA-aware editor:

- **What is the experience of writing with the tool?** Writers should be concentrating on the content, not the structure or the layout. A tool that protects the writers from the mechanics of XML authoring keeps them from being distracted from their core job. Choose an authoring tool that includes, or allows you to build in, templates for the topic types so that writers are guided through the topic structure and do not have to memorize the DITA specification.
- **Does the tool support special elements and attributes?** To add to the ease of use, look for a tool that easily integrates and aligns with your metadata values and enables the visualization of any reused content embedded using content references and conditional text. Look for a tool that allows you to customize the interface to include or align the metadata values with your typical usage.
- **Does the tool create valid DITA content that satisfies your business needs?** The editing tool you choose should support DITA compliance through validation. Beyond that, the tool and the DITA content should work the way you require and expect. Does the editor and output processor deliver that or does the tool use non-DITA compliant delimiters and reference indicators?
- **How easily does the tool support building and maintaining DITA maps?** Rather than manually editing the map in XML, choose a tool that has a graphical interface for map manipulations. This illustrates the hierarchy of the map structure as a tree, similar to a table of contents, and allows for easy, direct organization of the map structure.

- **How easy is it to generate deliverables?** Does the editor include a publishing engine? Can the author preview or generate output as it will be used? Ideally, these options are available via a menu option rather than exiting the application to drive the process through command line access or a second application.
- **Does the editor have rich word processing features?** Writers are accustomed to a rich toolset for authoring and they should be given access to a robust interface and toolset. This includes features such as spell checking, revision marks, and graphics handling and visualization. More mature offerings include these and other word processing features.

Content Contribution Tools

Some tools include a pared down version of themselves available through a web interface that gives contributors, who may not be experts with content creation or word processing, a way to add content to the repository without having to learn DITA or the full-featured editing tool.

If subject matter experts regularly contribute content that is to be used in a DITA topic and they author in a tool such as Microsoft Word, before you buy, test your authoring tool's ability to intelligently copy and paste content from a word processor into DITA. Tools vary in their ability to incorporate non-DITA content into DITA topics.

Publishing DITA

Depending on the editor you choose, you may find yourself using an external tool for publishing your DITA content. Because it is standards-based, DITA can be transformed, or published, using a variety of techniques. XSL is the XML language for transforming content. XSL consists of three components:

- XSLT – a transformation language, a straight-forward learning adventure
- XPath – an expression language for addressing parts of an XML document
- XSL-FO – a vocabulary of formatting objects with their associated formatting properties, a challenge to master

Authors can publish via XSL-FO for PDF or print or an XSL stylesheet for transforming content to HTML or Windows HTML Help via the Microsoft Help Workshop. These transformations rely on the DITA content, as defined by the map being processed, being assembled with all the topics with their content references resolved and conditional content filtered. This requires a tool that can interpret DITA's advanced features for DITA-unaware output tools, such as PDF and HTML.

The most popular way of achieving output with DITA is through the DITA-OT, an open source project that keeps pace with the changes in the DITA standard.

Delving into the DITA-OT to produce customizations can be challenging as it requires a detailed understanding of XSL processors for transformations and conversions. Having experience with Apache Ant, a scripting/build environment that helps create and run projects, can make developing and distributing configurations of the DITA-OT easier. Some modifications may require Java programming capabilities to set up and configure applications on your system. Technical writers may find this level of technical skill daunting.

There are resources in the DITA community who can help you stylize your output to meet your branding and content requirements.

Content Management

DITA does not impose a requirement for either a CMS or CCMS. Investing in a system for managing the components (such as topics, images, or maps) gives you a great deal of power over the management of the documents you offer and enhance your implementation of DITA.

Unlike traditional technical writing tools, DITA stores content as individual topic and resource files. Depending on the size of your output document, and its complexity, the number of individual files may grow significantly, becoming ungainly and difficult to search and maintain.

DITA works when stored in a file structure on a server located centrally for authors and contributors. Using a CMS or CCMS is more practical and production-worthy, especially when you have extensive or complex production requirements. This includes specialization and localization.

Using source control software gives you some rudimentary control over the content but misses many of the advanced features that make working with DITA easier.

Ideally, you want all your projects stored in the same location to facilitate sharing. A CMS or CCMS provides a central location with additional benefits. Depending on the features provided, the content management solution chosen should include some of the following capabilities:

- Search at the topic level, particularly full-text searching.
- Integration with your authoring tool.
- Is extensible, allowing you to influence the actions when content is checked in or out.
- Is configurable, allowing you to change the user interface.
- Link awareness and management. Use this to track where a piece of content is used so that you understand the effect of changes to that topic or graphic.

- DITA-specific topic grouping showing the relationship between topics that share a map.
- Handles XML as a distinct data type, recognizing XML features.
- Integration of workflow features or can be integrated with an external workflow system to manage the process of developing and delivering content.

Localization Management

Switching to DITA has the potential for dramatic localization savings if the process is carefully managed. Some CMS and CCMS incorporate the ability to assist with localization management. There are a variety of tools, methods, and techniques for handling localization requirements. When evaluating your toolset, consider the following:

- **How will content be exchanged between authors and translators?** Are the translators connected to the content repository or do sets of topics have to be bundled and transmitted to them? If the translators are not connected, can the process of bundling and shipping content to localized be automated? If there are topics that have not changed, is there a way to mark them to avoid redundancy of effort?
- **How will content be passed to translators?** Is there a workflow management system that connects the authors and translators? Is there a localization-specific workflow system in place?
- **Do the translators have the tools to accelerate turnaround?** This, in particular, refers to tools like translation memory which tags content as previously localized and then automatically suggests the translation based on past experience.

Workflow Management

Workflow management systems route information and tasks from one individual to another based on their roles and assignments in a process. The provides confidence in the process and also standardizes the notifications passed to individuals identified for a particular document or process. Many workflow management systems include reporting, which enables you to track where there are bottlenecks in a process allowing you to address those issues quickly.

Workflow systems ensure that individuals are assigned the right task at the right time. These systems help shepherd content through the process of creation, approval, and publishing. The system should present a relatively simple mechanism for setting unique workflow rules.

Step Three: Configuring Your Tools

Nothing fits right out of the box. There are compromises and adjustments required to achieve a tailored fit. Traditional authoring and publishing tools require the setup of styles, templates, macros, and conditional text settings; DITA is similar.

Unlike a traditional tool, DITA is comprised of different components that are configured independently of each other but need to be coordinated with each other so that all your tools are speaking the same language.

Configure the Authoring Interface

Just as the DITA standard has a degree of generality, allowing it to conform to broad industry standards for instructions for use and learning content, the authoring tools accommodate a variety of authoring needs. Just as you would adjust the style and layout of a non-DITA authoring tool, you should be able to configure a DITA authoring tool for unique elements and attributes. The ease of configuring an authoring environment varies with the tool and can be quite an undertaking. The need for and ease of this should be part of your tool evaluation.

Changing the interface of your selected authoring tool is a much safer choice than changing the DITA DTD or schema. Changing the interface does not affect the display of documents created using DITA. Nor does this type of change affect the stylesheets.

Plan how you will use metadata elements and attributes: will you embed them in topics or in the document map? Set up conditional text for your required attributes and values. If your editor provides dialog boxes for editing metadata elements or attributes, trim down the offerings to match your requirements and simplify the interface for your contributors.

Look for an editor that will allow you to add menu items or dialogs for inserting specialized elements. Configure the new stylesheets to display the new elements. If, for example, authors repeatedly enter the same metadata values, create checkboxes, radio buttons, and drop-down lists configured to default to the most common or safest value. Hide or remove commands for unused DITA elements and domains. Finally, modify the topic templates to accommodate the requirements of your implementation, adding specialized elements and hiding unused elements.

Do You Need to Specialize DITA?

An important part of configuring your tools may involve configuring, or specializing, DITA. One of the drawbacks to using a standard is that it may not include the detailed requirements that you've identified for your content. DITA is a rare standard as it is extensible, allowing specialization. The DITA-OT supports the inclusion of specialization plug-ins to process your specialized content.

When you specialize DITA, you can create new elements and attributes. These will be based on existing elements and attributes. The specific topic types in DITA, such as concept and task, are specializations of the core topic type.

Why would you specialize?

One reason for specializing DITA is to reduce the elements available to your contributors. This, for example, is similar to the specialization of the general task into the strict task, both available in the DITA standard. This type of specialization is useful if you want to facilitate learning DITA.

One way of simplifying the DITA elements, is to add industry- or company-level semantic value to the elements by renaming them.

If the DITA standard does not include elements that you've identified as part of your requirements, for example part numbers, you can specialize elements to create new elements that satisfy your specific needs.

As DITA is developed, new needs are recognized and addressed. Until DITA 1.3, organizations wanting learning topics and elements were faced with specializing those topics and elements.

When the DITA specification is specialized, the output tools need to be updated, as well. The DITA-OT, which is the base for many tools, handles the output processing and needs to be modified with the addition of code that handles the specialization.

Before you begin, check the DITA focus area maintained by OASIS for the latest leverageable resources that industry leaders have contributed to, an appropriate specialization may already exist. An inquiry on the DITA mailing list (groups.io) may connect you to someone doing similar work already.

Before you decide on specialization, you need to consider the following questions:

- **Do your contributors require additional semantic help?** In this case, you can specialize the generic DITA element names to provide clear indications as to what, exactly, is to be entered into the structure at any given point. A specialized schema or DTD helps both the organization and the writer by enforcing consistency and guiding the writer along with rules designed for consistency, clarity, and searchability.
- **How do you plan to use and reuse content?** DITA is designed for reuse and does not require specialization to accomplish this. But, if your reuse involves sharing content with another company, you will either have to share your specialization with them, or generalize your content when exchanging it with another company. Generalizing content is simply transforming it back down to the original elements and attributes, losing

the semantic value but simplifying the exchange. If you plan to use DITA for a purpose that requires specialization, such as a knowledgebase, you may want to create specific topic types, such as policy, specification, or how-to, that are not part of the DITA specification.

- **How complex are your output requirements?** If you have specific output requirements for elements and attributes within your content, you need to make it findable. Semantic tagging is useful in filtering and formatting for different outputs. This way, the stylesheet can apply the required manipulation to the correct element. For example, you can specialize an unordered list into a checklist, generating checkboxes instead of bullets at rendering.

Specialization requires coordination between all the tools in your solution. Keeping all the tools up-to-date with the specialization represents an on-going cost that must be taken into consideration. Be sure, before you commit to specialization, that all your tools can handle the changes you plan to make to DITA and the DITA-OT.

Best Practices for Specialization

- **Understand your content goals.** Through your content audit and careful planning, get clear on your content requirements, not as they are today, but as you would like content to be after the move to topic-based authoring.
- **Have clear motivations.** There are many reasons for specialization, all of which need to be clearly articulated to make effective design choices.
- **Build on success.** Find organizations that have already specialized in a similar way. Gain ground by creating a relationship and leveraging their work or collaborating with them on specializing the standard. Learn from their experience and benefit from their work.
- **Extend your DITA knowledge.** Many aspects of working with DITA are not apparent without digging into the specification, the DITA-OT, and experiencing where DITA is strong or weak for your particular industry, implementation, or staffing. Designate a DITA strategist on your team, someone whose job includes developing or connecting with the required expertise for specialization. That person would be responsible for ensuring that the end result aligns with proven architectural principles.

Apply Your Brand and Customize Output

Organizations wanting to present their brand as part of their content output will customize the DITA-OT stylesheets. The DITA Open Toolkit provide generic starting points for a variety of output options. The best approach to these modifications depends on your strategy and the details of the change.

For example, simple Web-based changes can be accomplished through some basic CSS modifications. On the other end of the scale, changing print requirements, such as the formatting of tables, could require some XSL-FO programming.

Specialized DITA makes customizing the stylesheet critically important to get the elements and attributes you've created to appear as you want.

Configure Other Tools

If you've chosen a CMS/CCMS, a localization management system, or a workflow management system, these will require work to configure them to operate with your business goals, content, processes, and other tools. Speak with the vendors to find out what is entailed in bringing these tools into compliance with your needs.

Step Four: Convert Content

What to do with legacy content is one of the hurdles of conversion and transitioning to DITA. This step is best done using the comprehensive inventory of your content, done in step one, to inform this process. You have the option, here, to start fresh and not actively convert your content but to leverage it as a resource for new documents, new topics, and a new structure.

If you decide to convert your legacy content, you then have options on how to handle this. Do you want to use writer resources for conversion or do you want to farm it out to a business that specializes in converting content? Do you need to replicate your manuals for certification reasons?

Do you want to change the content to align your DITA-based content with industry practices such as minimalism and topic-based authoring? Do you want to leverage the DITA topic types and provide semantic indications through elements and topic structure? Do you want to restructure to make better use of reusable content topics and other elements?

Has your content been localized? Are you wanting to stay as close as possible to the legacy content to save time and money on localization?

What are your plans for handling issues such as variations in descriptions, instructions, notes, and images uncovered during the conversion process? Will you incorporate post-conversion clean-up or clean-up before conversion? What parts of your current corpus will you convert to satisfy the requirements of your products and users? Are they prioritized in your plan?

Are you using a CMS or CCMS? How are you going to import content into the system? Who will build the books and publish them? Is this purely legacy content

or are you satisfying a new product release? Have you run a pilot project to ensure each step of your planned procedures work and meet expectations?

A pilot project can result in an improved knowledge of DITA on the part of those on the pilot team and this is then shared with the broader organization. Is the team converting the legacy content or are they working with a conversion vendor? If they are working with a vendor, do you have clear and concrete service agreements with your provider? The pilot team is responsible for mission critical documents such as, the conversion plan, timeline, and test data to demonstrate the conversion process.

Do you have an up-to-date style guide? Does it need to be updated to accommodate the transition to DITA? How are the changes communicated to the pilot team and then to the larger organization? Who will review the converted content against the style guide to ensure it is consistent? Who reviews the converted content against the conversion specification?

Make sure there is a project manager shepherding your conversion process through the many pain points you need to identify and solve. You'll need someone to keep the project on track and the team focused on the conversion and transition to DITA.

During this pilot project, you'll also be testing your choices made, such as DITA specialization, output customization, and content management strategy. You'll track and resolve issues discovered while test driving the documentation during the pilot project.

In the end, the right technique for your situation will answer the following factors:

- **How much content do you need to convert?** The greater the volume of your legacy corpus you plan to convert, as it is, the greater the need for automation in the process. Are you making a clean break with the past? Are you constrained by localization costs and forced to convert everything word-for-word? Can you identify reuse opportunities that will reduce your conversion costs?

Focus on the documents with the highest ROI for conversion. Prioritize documents with the greatest strategic value.

There is no need to convert everything at once, prioritize and plan your conversion.

- **How consistent is the legacy formatting?** How strict are your styles and templates in your legacy system? The stricter the styles and templates, the easier the conversion. Few conversions are one-step processes without some confusion and manual effort.

Recognize that DITA represents a new paradigm for authoring and publishing. To be effective, DITA itself is strictly structured and your new content, your published content going forward, will reflect that and be noticed as an improvement by your customers.

- **How topic-oriented is your current content construction?** Do you already write in a topic-oriented fashion, dividing content into discreet topics that focus on a specific aspect of the product? It's the rare situation where the content is already structured and ready for an easy conversion. Some revision may be necessary, either before or after conversion. There are tools and vendors who can create individual topic files automatically from the heading structure of your documents.

Some companies implement the topic-based structure before venturing into a conversion. By cleaning the content up before converting, the writers have an opportunity to familiarize themselves with the new paradigm for authoring.

- **Look ahead at how rich you want the converted documents to be.** One time saving strategy for conversion is to simply use the unspecialized generic topic as the destination. This eliminates the need to identify the specific topic type, concept, reference, task, troubleshooting, glossary, and the newer learning topic types.

Step Five: Go Live With DITA

It's time to step out of the sandbox, to end the pilot program, to launch DITA and complete the transition. You may start out with some projects fully written in DITA and others, pre-transition, being written to new styles and standards.

Unlike traditional publishing, DITA content is topic-based and each topic is stored in its own file. In traditional authoring and publishing, writers work on a document and are responsible for the full breadth of content, or they have a complicated sharing process where multiple writer contribute to a single file. DITA resolves this. Individual writers are responsible for a subset of the final project result.

A typical transition consists of creating an outline, that will become your map file, scoping the files that are required for the project. The early creation of all the associated files makes planning and output generation throughout the project easier. You can create placeholder topics and distribute these to authors to fill out and complete. A placeholder topic consists of the title, a short description, and the structure of the topic.

The right tools and technology have a huge effect on the success of any DITA deployment. DITA and its associated tools can be implemented in phases and adapted over time.

JUSTSYSTEMS.

<http://www.justsystems.com>

North America:

+1.866.793.1542

contactsales-na@justsystems.com

International:

+1.604.697.8705

About JustSystems

JustSystems is a leading global software provider with decades of successful innovation in office productivity, information management, and consumer and enterprise software. With over 2,500 customers worldwide, the company is continuing a global expansion strategy based on its xfy enterprise software, XMetaL content lifecycle solutions, and its pioneering work in enabling XBRL financial reporting technologies. JustSystems is one of the 2008 KMWorld 100 Companies that Matter in Knowledge Management, a 2008 EContent 100 member, and was recognized on the 2008 KMWorld Trend-Setting Product list for XMetaL. Major strategic partnerships include IBM, Oracle, and EMC.

© Copyright JustSystems

All rights reserved. JustSystems, the JustSystems logo, and xfy are trademarks or registered trademarks of JustSystems Corporation in the United States and other countries. XMetaL is a trademark of JustSystems Canada, Inc in the United States and other countries. Other company, product, and service names may be trademarks or service marks of others.