

The background features a large, stylized graphic of a face. The face is composed of thick, rounded lines in shades of light blue and light green. The eyes are represented by two solid light blue circles. The mouth is a thick, light green pen nib pointing downwards and to the right. The overall style is clean and modern.

The DITA business case

Maximizing content value

Sarah O'Keefe

Bill Swallow

September 17, 2018

Executive summary

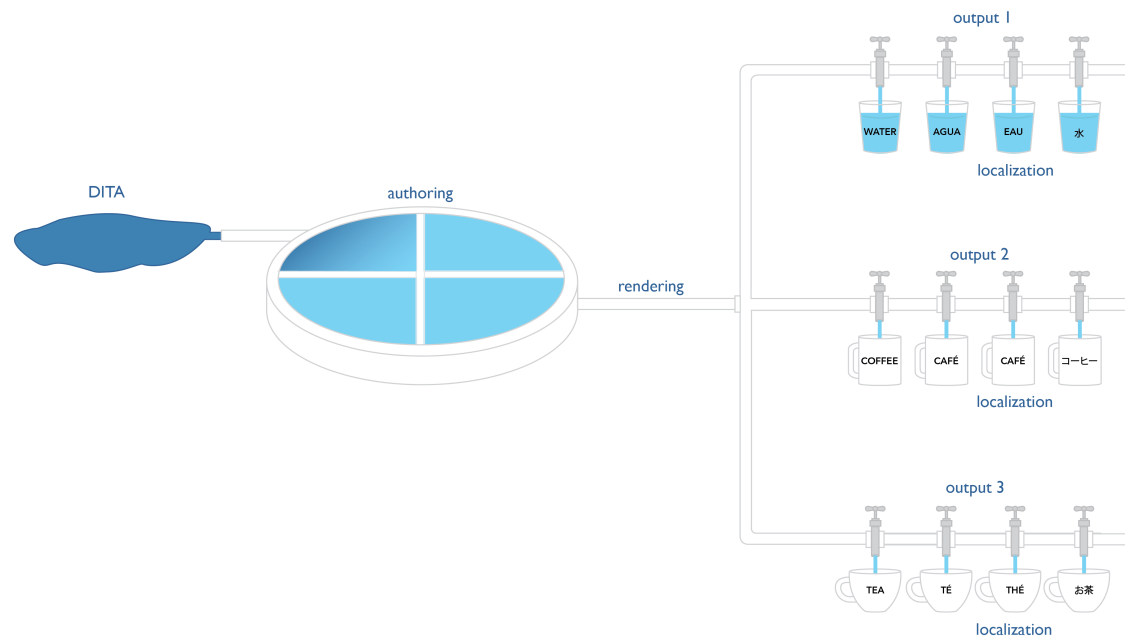
Companies require content to support ever-increasing requirements, including:

- Delivering content in multiple formats
- Meeting compliance requirements
- Accelerating time to market
- Handling content variants
- Delivering translated content on a limited budget

This white paper describes the business justifications for investing in the Darwin Information Typing Architecture (DITA)—an open source XML standard—as a foundation for content management.

Need for information in multiple formats

DITA XML files are encoded in plain text markup that looks similar to HTML. From DITA XML, you can render content into many formats, including PDF, HTML, Word, Markdown, InDesign, man pages, JSON, WordPress, and SCORM. Each publishing pipeline can also be set up to support multiple languages. Once you configure the publishing pipeline, the process of generating output is automated.



Automation can provide the following benefits:

- Your organization can ensure that all information streams are in sync.
- Instead of repetitive manual processes to set up and format each content delivery type, you configure the publishing pipeline once and then let it run.

- In a well-designed pipeline, the changes required for localization are isolated into configuration files, which makes it straightforward to add new languages or make adjustments for a specific language.
- Starting from XML gives you the flexibility to add new content delivery formats as needed.
- You can integrate the publishing pipeline into other enterprise tools, such as source control, QA, and build systems, to further automate the content updates.

Compliance requirements

DITA is useful to ensure compliance with regulatory and legal frameworks. The DITA content model is configurable, and you can validate content against the required model. You can use DITA validation to ensure that the document you are submitting contains the required information, in the specified order. Editors can then concentrate on the quality of the text rather than double-checking the technical compliance against the document standard.

For example, you can set up a DITA content model for a medical journal article to require an abstract, author information including credentials and affiliations, and required content containers in a specified order (introduction, method, results, and discussion, for example). Without these required components, the article is not valid. An editor would read the abstract to ensure that it is an accurate representation of the article.

You can configure DITA to require, allow, and disallow content components in different contexts. Inside DITA tools, authors have guided editing; the software tells them which elements are required and allowed at different points in a document.

Accelerating time to market

DITA can improve content velocity (time to market) in the following ways:

- Content reuse means less content to manage. Content is written once and then referenced into multiple other locations.
- Automated formatting means less authoring effort and faster delivery of content, in all languages.
- Connected content. You can extract information from its source and automatically push it into the DITA content. For example, you can extract product specifications from an engineering product database instead of copying and pasting. If the database is updated, the document changes accordingly.

In addition to efficiency in your publishing workflow, you can also assess reduced time to market and how the earlier availability of product content might affect product delivery. If a product sells a modest \$1 million per year, then each week of availability is worth about \$20,000. If you can deliver your content sooner and thereby accelerate the delivery of the product or reduce the delays in shipping localized versions, you can potentially get your revenue faster.

The most common business case for accelerated time to market is in reducing the wait for localized content.

Supporting content variants

Variants allow you to eliminate content redundancy and provide targeted information to support product variants (for example, product models with both shared and different features). This strategy can also help you meet customer requirements for personalized documentation.

In DITA, you can set up content variants with multiple facets, such as platform, customer, audience, and product. This allows you to create a huge number of possible variants from a single set of source files. You can then use the tagged content to publish each variant or support a dynamic versioning approach.

Localization cost savings

If you localize your content, DITA can streamline the translation and formatting work:

- *Translation effort:* DITA offers a robust content reuse framework. Instead of copying and pasting reusable text, the content is written only once and is included by reference as needed. This means that the text is also translated once (per language), regardless of how often it appears in the published content. Reducing the total number of words translated is a powerful way to reduce the overall cost of localization. In addition, it improves the quality of the information by ensuring consistency throughout the document. (After information is copied and pasted, copies inexorably diverge over time.)
- *Formatting costs:* Typically, 30–50 percent of total localization cost in a traditional workflow is for formatting. After the files are translated from the source language into the target language, the files must be reformatted to accommodate text expansion and pagination changes. In a DITA workflow, that formatting is automated. The DITA content files do not contain formatting information. To generate final formatted content, you apply formatting stylesheets to the translated DITA files. These stylesheets are configured ahead of time to support the languages you need. An organization that spends \$1,000,000 per year on localization will save \$300,000–\$500,000 of that amount by eliminating formatting costs.

The DITA ecosystem

DITA adoption has grown steadily since its public debut. The types of companies employing DITA for their content have also grown beyond software to include life sciences, manufacturing, financial, and more. Although DITA has its roots in product and technical content, organizations are also using it for marketing, training, and other important customer-facing content. As a result, there are numerous DITA software vendors. Many organizations use a DITA-specific content management system (CMS) to manage their files. In addition to typical file management, a DITA CMS typically provides:

- Workflows for authoring, editing, approval, and publishing
- Robust search and “where-used” reporting
- Lightweight authoring environments for part-time content contributors
- An Application Programming Interface (API) for connecting to other systems

An API typically allows you to share and retrieve content from other enterprise systems, such as those for asset management, web content, and source control.

Next steps

DITA can be a powerful solution to tricky and costly content problems. Does DITA seem like a potential fit for your organization? Use our [XML ROI Calculator](#) to estimate potential savings from formatting automation, or [contact us](#) to explore how DITA can help you maximize your content's business value.

About the authors

Sarah O'Keefe, Chief Executive Officer, founded Scriptorium Publishing to work at the intersection of content, technology, and publishing.



Today, she leads an organization known for expertise in solving business-critical content problems with a special focus on product and technical content.

Sarah identifies and assesses new trends and their effects on the industry. Her analysis is widely followed on Scriptorium's blog and in other publications. As an experienced public speaker, she is in demand at conferences worldwide. In 2016, MindTouch named her as an "unparalleled" content strategy influencer. Sarah holds a BA from Duke University and is bilingual in English and German.

Follow Sarah on Twitter at [@sarahokeefe](#).

Bill Swallow, Director of Operations, specializes in global content strategy.



Bill partners with enterprise content owners to design and build content systems that solve complex information management and localization problems. Often, projects require centralizing and streamlining content development across multiple teams, departments, and regions. Bill expertly balances high-level business priorities with the specific needs of content contributors. Between 2010 and 2017, MindTouch identified Bill as a leading influencer in technical communication, content strategy, and content experience. Bill holds a BS in Communication from Rensselaer Polytechnic Institute.

Follow Bill on Twitter at [@billswallow](#).