An Introduction to the Darwin Information Typing Architecture (DITA) and its benefits for technical training & documentation

A new XML content standard for blended online, mobile and instructor-led technical training and documentation

The importance of Structured Content Management in technical training

To compete in a demanding market, manufacturing and engineering companies must deliver information and training to their workforce and their clients wherever and whenever required.

This is driving smart instructional designers to create compelling training materials that can handle the increasing complexity and blending of technical training formats, from instructor-led sessions to web-based training, mobile learning and beyond.

Instructional designers who create technical training courses in large engineering, manufacturing and process-based organizations obtain the majority of their input materials from a wide variety of sources. Typically, these include engineering, design and marketing departments - all of which use different file formats and different content packaging and exchange standards such as DocBook, ePub, IETM or S1000D.

These formats are also different from those commonly used within learning and training tools where we typically find SCORM, IMS and AICC.

This lack of a common interoperability standard restricts the type of documents that training departments can receive from their technical counterparts to simple formats such as Word, PDF and HTML.
Consequently, the ‘cut and paste’ approach is still the most commonly used strategy in producing technical documentation for online training delivery. This results in high levels of content replication, low levels of content reuse and the high costs of engaging subject matter experts (SMEs) in never-ending proof-reading cycles.

Recognizing these inefficiencies, organizations need an improved interoperability format to ease the flow of information between knowledge hubs and to foster content reuse across departments.

**DITA XML: a solution for technical training**

The new Darwin Information Typing Architecture (DITA) ([http://dita.xml.org](http://dita.xml.org)) Learning & Training specialization addresses the gap between technical documentation and training. It brings new life and content interoperability across and within organizations managing large volumes of technical documentation and training materials and rapid production cycles.

The standard is issued by OASIS, the US-headquartered global consortium for open standards in the Information Society, which aims to boost the convergence of technical and training documentation within engineering and process oriented industries, including the semiconductor, automotive, financial, pharmaceutical, translation and defense industries.

In common with other XML standards DITA clearly separates information content from presentation style, but its innovation is in the clever use of object orientation and specialization within XML typing.

The DITA specification encourages greater flexibility in learning documentation publishing and reuse by organizing digital materials and documents into small, tightly focused ‘topics’, specialized by subject area or vertical market.

Each topic can be organized and linked into a variety of different publications. This is in contrast to formats such as *ePub* and *DocBook* where a piece of information is sealed into a book or a chapter, where it cannot easily be reused.

DITA working groups have released many “specializations” relevant to a variety of industries including semiconductors, machinery and language translation. DITA’s contribution to learning reached a milestone in late 2011 with the release of the new specialization for Learning & Training Documentation.

eXact learning solutions has always been active within the communities which develop learning standards, such as ADL, SCORM and IMS. As a result of a recent upgrade, the award-winning eXact LCMS and eXact Mobile platforms now support XML DITA content authoring and management, in addition to the other learning standards (such as SCORM, AICC or IMS Common Cartridge).

DITA’s features offer a significant return on investment for organizations because content reuse and structured writing significantly reduce the time needed for authoring and content production.
Why a new standard for education and publishing?

Rich media and mobile learning are important to any manufacturing and engineering organization supporting its workforce with just-in-time training and performance support.

The production, management and delivery of effective and up-to-date training materials can be extremely complex. Getting it out ‘right’ and getting it out on time are the most effective ways for training departments to help their companies compete globally.

In planning their contents for technical training, instructional designers receive up to 90% of input materials from very different sources, typically tied to product engineering and design departments and often in formats that are not directly compatible with learning and training tools.

I don’t want to write, review, translate, and repurpose the same content multiple times! „

To date, content duplication and repeated creation have often been the only ways to cope with the misalignment. This introduces opportunities for errors and relies on expensive and time consuming content review and proof reading phases to correct them. It also makes content re-use hard to achieve. This increases costs for content production and repurposing across the different units of an organization. This level of complexity is mainly due to the ineffective answers that some publishing standards and technologies - currently in place in most organizations - provide to “multiple source - multiple output” scenarios that the same organizations need to implement. A better interoperability strategy and a solution to diverse and dispersed knowledge hubs within the organization are the right answers to those needs.

Benefits of a “multi-source/output” strategy

An effective “multi-source/output” approach is essential for organizations to:

• Reduce the cost of the whole set of training-related processes in the organization
• Create and maintain large and structured product documentation in a cost-effective way, reducing time and cost of knowledge transfer between SMEs (Products and Services, IT) and knowledge consumers (Sales and Marketing, Customer Care)
• Protect the IP of knowledge and documentation, avoiding duplication of content across the different channels through which it circulates during the production of documentation, training materials and marketing messages
• Explore new ways to deliver content other than via printed materials but still drive everything from the same source;
• Enable a competency-based system for the allocation of courses for each discipline based on students’ gaps and learning goals.

In short, “multi-source/output” provides fully blended knowledge and learning processes at a reasonable cost.

The publishing market has identified, in DITA the standard modality to structure documentation in a way that makes “single source - multiple output” a reality and
traditional publishing mechanisms can be kept while also addressing new media and new distribution channels.

About DITA
DITA builds content reuse into the authoring process. It defines an architecture for designing, writing, managing and publishing information in print and on the Web.

The standard is advanced through an open process by the OASIS DITA Technical Committee, a group that encourages participation from developers and users.

DITA’s features are based on decades of research into methods for technical documentation. These include modularity, structured writing, information typing, minimalism, inheritance, specialization, simplified XML, single-source, topic-based, ready-made metadata, conditional processing, component publishing, task-orientation, content reuse, multi-channel and translation-compatibility.

History
DITA traces its conceptual ancestry to before the 1960s. In that decade, B.F. Skinner used mainframe computers as ‘teaching machines’ and developed precepts for ‘programmed learning’ that required learning materials to be segregated into ‘chunks’. In the mid-1960s, Robert Horn developed information mapping techniques. Common ‘information types’ were identified in standard documents such as user manuals, annual reports, policy and procedure manuals. Identifying standard information types is at the heart of DITA.

In March 2001, IBM introduced DITA as a series of developer Works articles about a new version of XML for documentation. DITA was an attempt to make a simplified XML start kit - one designed from the outset to encourage the reuse of small content components. DITA’s goal was to formalize information typing practices, both print and online, and also enable an extensible typing architecture through specialization of base topics. DITA maps were a way to standardize collection publishing and information architecture/ outlining models.

In May 2002, IBM added domain specialization to topic specialization and demonstrated these in the Open Toolkit - a reference implementation of DITA publishing - that allows DITA documents to be output in a variety of common formats including PDF,

Current version of DITA 1.2 specification can be downloaded at:
http://docs.oasis-open.org/dita/v1.2/os/spec/DITA1.2-spec.html
Word and HTML. IBM encouraged authoring tool vendors to integrate the Open Toolkit as a means of publishing DITA.

In April 2004, the Organization for the Advancement of Structured Information Standards (OASIS), formed a Technical Committee to explore a DITA Standard. The committee included tool vendors, information and content management consultants and, most importantly, end users.

In February 2005, IBM donated the Open Toolkit (OT), a limited version of its internal Information Developers Workbench, to SourceForge. IBM continues to develop the OT, which is not a part of the OASIS DITA Standard efforts.

For further details about DITA, visit http://dita.xml.org

eXact learning solutions and DITA

eXact learning solutions has enhanced its eXact LCMS, to support DITA XML editing and rendering. This adds technical documentation to its single source multiple channel learning content management architecture.

The eXact DITA Editor was launched at the end of January 2012 at Learning Technologies 2012, in London, and then at Learntec 2012, in Karlsruhe, Germany. It offers a fully-fledged DITA authoring, management and delivery environment. This environment is fully integrated with the SCORM-based online and mobile learning content management and rendering capabilities which are at the core of the eXact LCMS platform - currently installed in over 100 organizations worldwide.

At the eXact DITA launch Fabrizio Cardinali, CEO of eXact learning solutions North America stated: “After DITA’s recent announcement of a new learning and training materials specification, many corporate CIOs and CLOs are now asking for DITA compliancy from suppliers. They want to import DITA documentation from other departments - and they want to convert and embed it into AICC and/or SCORM packages to be used for online training and within their instructor-led and technical training content production processes. For many years, sophisticated organisations have found it hard to share and re-use content between technical and training
An Introduction to the Darwin Information Typing Architecture (DITA)

 departments. DITA is the solution to the problem.”

Implementing a structured content management strategy in eXact LCMS

Native DITA support in eXact LCMS complements the overall strategy of supporting blended learning and publishing through single-source content.

A fully blended approach to the production of training materials and technical documentation is an important success factor in every organization addressing change management at any level. A blended approach guarantees that all stakeholders have access to the right knowledge anytime, anywhere, regardless of their role in the organization, their duties, their working time and location, and the technology they can access.

A fully blended approach has key advantages for those organizations that aim to reduce the costs and time taken for creating and maintaining training materials and technical documents. It is particularly important when the process involves multiple delivery mechanisms and a mixed network of authors, subject matter exports, instructional designers and trainers, while requiring consistency in learning content and technical documents.

Authoring DITA

eXact DITA, the leading solution for Learning Content authoring, management and delivery through multiple platforms and delivery channels, recently improved its ability to support authors in the production of multiple output materials from one set of source documentation and across a single content production process.

eXact DITA Editor allows training contents and technical documents to be developed in a structured and organized way. All information can be structured into self-contained information chunks. A chunk (or “Topic”, in DITA terms) encapsulates a single subject, and is the smallest piece of information that can stand on its own.

The eXact DITA Editor has been optimized for DITA authoring, embedding features such as:

- ‘Word processor-like’ user interface. Authoring DITA documents is extremely easy, thanks to a user interface which mimics word processing tools. Structuring
An Introduction to the Darwin Information Typing Architecture (DITA)

and formatting content according to the DITA schema is similar to using paragraph styles and word formatting in MS Word.

- **Immediate preview.** While What-you-see-is-what-you-get (WYSIWYG) is only approximate in the case of XML authoring, due to the focus on semantic tagging and multiple output options, the final appearance of a DITA document can easily be accessed through instant export to a variety of formats, including SCORM, eBook, RTF and PDF formats.

- **Elements in context.** When adding structural elements, such as a new section in a DITA topic or a step in a DITA task, the authoring tool displays a context-dependent list of valid elements that can be added at the current insertion point in the document.

- **Tags-on view.** Although they often disrupt the WYSIWYG look, optional visual representations of the start and end tags for structural elements can be helpful, especially to validate content tagging and ensure re-usability of document structures at a granular level.

- **Document Structure view.** A hierarchical outline view of the document is always available. Authors can expand and contract elements like an outline tool, move content around and synchronize changes with the main document view.

- **Visual DITA Map editing.** The topic references, arranged in a hierarchy, can be changed without looking at the XML code.

- **Drag and drop structures.** Recursive aggregation of topics in DITA Maps is as easy as dragging and dropping them into the tree-representation of the DITA Maps under editing.

- **Spell check.** Spell check functionalities are available and configurable to match the specific language of each author or document. Dictionaries can be configured to enable customized word lists, ensuring that writers use terminology consistently.

- **Multilingual support.** eXact DITA Editor is integrated with the powerful content localization engine of eXact LCMS. Exporting, translating and re-importing translation memories (TMX) in standard X-
LIFF (XML Language Interchange File Format) is yet another feature made available by eXact LCMS.

- **Open Toolkit (OT) support.** eXact DITA Editor uses the open-source DITA OT to take advantage of plug-ins contributed by the DITA community. Additional publishing engines can be plugged in to generate additional outputs.

**Links and Version Control**

Content linking and version control of structured content has been a feature of eXact LCMS since its first appearance on the market. The eXact DITA Editor leverages the dynamic linking and version control engine to maximize reusability of content and support distributed content reviews and controlled sign-off cycles.

Every individual element inside DITA topics can be automatically assigned separate IDs for use in other DITA documents. Reusable components can be managed, reviewed and versioned in eXact LCMS repository and are available at any time to DITA Authors to maximize reusability of content and information across the organization.

Content authored in eXact DITA Editor can be easily published to a variety of output formats - choosing, for each delivery channel, the appropriate topics and style for the intended platform or device.

Once content is ready for delivery, the LCMS parses all relevant content through a template-based publishing engine (included in the LCMS core system) to produce a variety of different publications out of a single source of information chunks, as shown in the figure below:

**Metadata, Workflow and Content management**

In a distributed organization, different authors work on the same content using a variety of authoring tools.
An Introduction to the Darwin Information Typing Architecture (DITA)

Adopting a single-source-enabled LCMS solution, all of the content can be stored in a single digital repository.

In addition, they will be split by the LCMS into logical chunks recognized by the system. Now we can really say that they *are* actual chunks. By that, we mean that each of them will have its own:

- Metadata
- Attributes, Tags
- Versions
- References
- Workflow Status
- Ownership and tasks

DITA Import/Export

In addition to DITA authoring features, eXact LCMS allows you to import any pre-existing DITA publication and chunk it down into its components to maximize the reusability of the information - starting from high level topics (such as manual sections or book chapters) down to the single paragraph or image file.

Imported DITA documents can be further authored, aggregated and reviewed using the standard collaborative tools provided in the LCMS.

At the same time, any DITA topic or topic aggregation (which in DITA is referred to as a “Map”) can be exported and distributed to external DITA-aware systems, leveraging the high level of interoperability provided by the eXact LCMS Service-Oriented Architecture.

Publishing DITA content to SCORM

Users of eXact DITA Editor have the unique ability to publish their DITA contents to SCORM all within a single authoring and content management environment. Indeed, eXact LCMS supports seamless authoring and sharing of content across different kinds of publications, including the world leading template-based authoring features of eXact Packager and the Online Editor, both of which are part of the eXact LCMS Suite.

Publishing DITA content to SCORM is the easiest way to use and reuse technical documentation in our learning content production flow. This reduces the time and cost of producing interactive learning packages and maximizes the Return on Investment.
Output channels
The eXact LCMS architecture shows how, from single source content and within a single content production process, DITA supports the production of a variety of outputs, including:

- Training Guides - for instructor led training
- Web-based training courses and course components
- Roll-Out Guides
- Support Guides
- Instructor Guides (for trainers)
- Online help systems
- Printed / PDF based documentation
- FAQs
- Tech Support website
- Marketing material (both online, printouts, SMS and eMail notifications)

Return on investment in DITA
Because the single largest cost for training content and technical documentation comes from lifecycle maintenance of the content and ensuring it is always up-to-date, DITA’s features offer a significant Return on Investment for organizations.

These features include:

- Translation Savings - the top ROI factor for global firms
- Single-source - allows a diversity of styles to be applied to the same data for consistency in content styles
- Content Reuse - don’t write the same sentence twice. This is one of the principal ways to reduce total ownership costs.
- Structured Writing - reduces authoring time and increases analysis time
- Task orientation - reduces customer service calls
- Minimalism - improves the end-user experience
- Specialization - customize topics to fit existing formats
- Multiple Output Formats - web (HTML), print (PDF), and online Help
- Multi-channel Delivery - add Mobile with minimum extra development
- Simplified XML - greatly reduces the development cost of Schemas/DTDs
- Metadata - supports semantic-guided search and conditional processing
- Conditional Processing - rapid creation of content variations for special needs
- Modularity - assemble documents from manageable chunks
- Component Content Management and Publishing.