

Why Is DITA So Hard?

Or

Understanding why people who are new to DITA are often overwhelmed when it's not really DITA's fault

Eliot Kimber
Contrext

DITA Europe 2015



About the Author



- Independent consultant focusing on DITA analysis, design, and implementation
- Doing SGML and XML for cough 30 years cough
- Founding member of the DITA Technical Committee
- Founding member of the XML Working Group
- Co-editor of HyTime standard (ISO/IEC 10744)
- Primary developer and founder of the DITA for Publishers project
- Author of *DITA for Practitioners*, Vol 1 (XML Press)



Agenda

- What is DITA?
- What people complain about and exploration of their complaints
- Why DITA is so challenging, with a focus on the hyperlink nature of DITA content
- DITA is Aikido
- Discussion of hyperdocument management requirements and strategies



WHAT IS DITA?



DITA

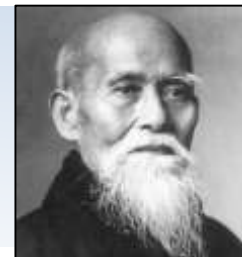


A small digression about Aikido

Please indulge me



Aikido



- A defensive martial art based on blending with an attacker's energy, capturing their balance, and redirecting their energy in order to return them to harmony
- Goal of Aikido is ultimately universal peace and harmony
- There is no one true way to do Aikido
 - Aikidoka are expected to develop their own expression and interpretation of Aikido as they develop their skills
- It's all about connection





DITA



- A standard XML application architecture for human-consumed documents
- Optimized for interchange and interoperation of content, processing, and DITA-specific knowledge
- Distinguishing architectural features:
 - Specialization: enables controlled extension from base DITA markup vocabulary
 - Use-by reference: Content components can be used in multiple contexts (DITA maps, content reference)
 - Indirect addressing: keys and key references
 - Designed to work entirely from a file system
- DITA is all about connection



DITA IS TOO HARD



Common Complaints About DITA

- DITA is too hard...
 - To author
 - To process
 - To manage



Is it DITA's Fault?

No



Then Whose Fault Is It?

Technical documentation itself



Fundamentally A Mismatch of Expectations



- DITA exposes and makes explicit the inherent complexity of sophisticated technical documentation
- Desktop publishing and non-DITA XML systems do not fully support this sophistication
- Therefore, authors aren't asked to do the most challenging things



DITA Exposes Complexity

- DITA is not just another way to tag technical documents
- DITA is a fundamentally different way of structuring and managing technical documents—not just books anymore
- What DITA does is complex, no question
- But it is complex because technical documents are complex



DITA Reflects Requirements

- The DITA standard reflects requirements organizations have:
 - Separation of content and formatting
 - Separation of publication structure (maps) from publication content (topics)
 - Fine-grained use-by-reference (conref)
 - Indirect addressing (keys)
 - Conditional content (filtering and flagging)
 - Detailed semantic markup
 - Interchange in the face of different tag vocabularies



Nothing Else Does What DITA Does



- No desktop publishing system does what DITA does
- No other standard XML language does what DITA does
- Few, if any, custom XML applications do what DITA does



Writers Naturally Are Not Prepared

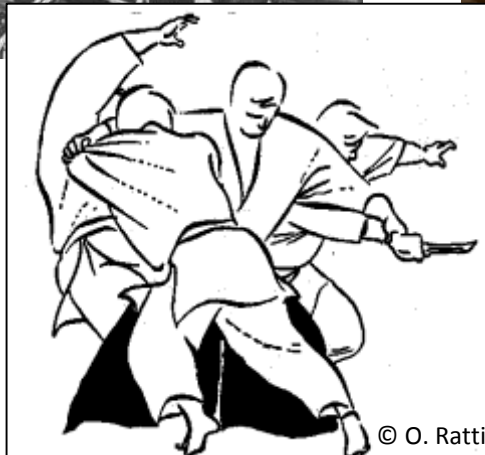
- Writers coming from non-DITA systems have expectations set by what those systems could do
- The sophisticated things DITA enables are going to be new ideas and tasks to most writers
- Should be no surprise that there is a shock
- Can't fault writers for putting the blame on the technology (DITA) and not the inherent requirements of sophisticated technical documentation



THE HARD PART: HYPERDOCUMENT MANAGEMENT



Link Management and Configuration Management



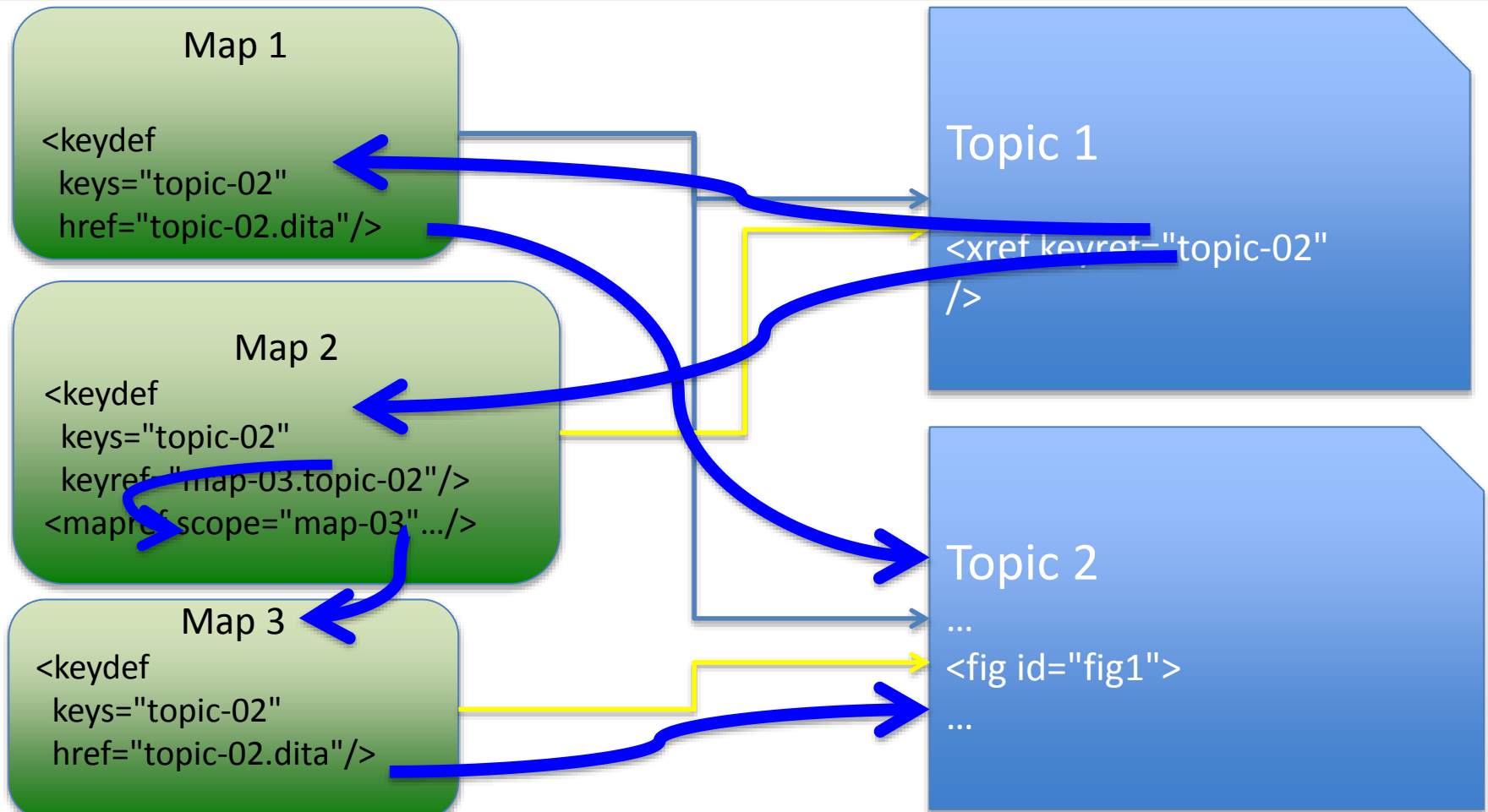
Map/Topic Separation



- DITA's fundamental reuse architecture
- Adds complexity to authoring...
- ...turns all publications into sophisticated hyperdocuments
- Sophisticated hyperdocuments are challenging in a number of dimensions

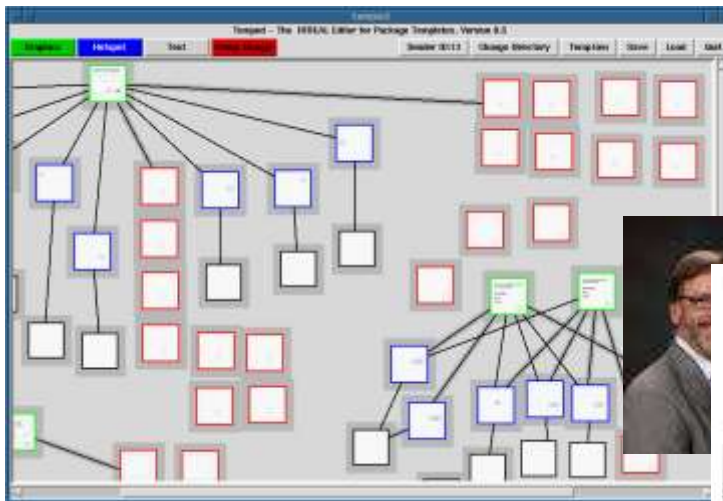


Maps and Topics and Links

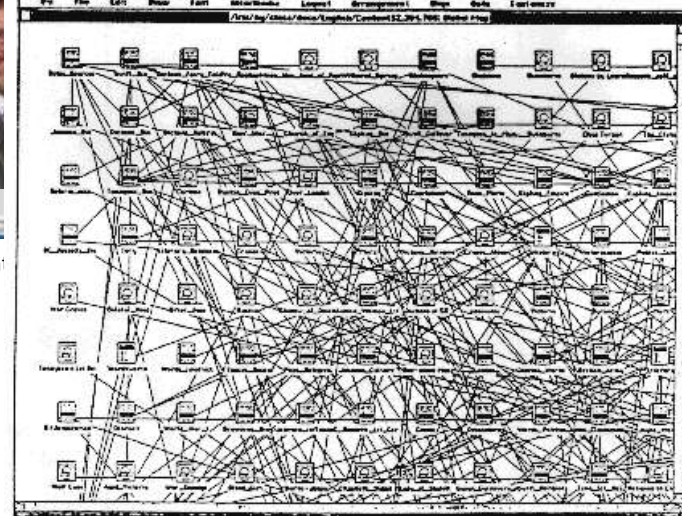


Realistic Hyperdocument Pictures

Google: <https://www.google.de/search?q=hyperdocument+images>



<http://www.cip.ifi.lmu.de/~butz/publications/papers/buildings/node3.htm>



<http://www.ickn.org/elements/hyper/cyb56.htm>



Indirect Addressing Is Required by Reuse



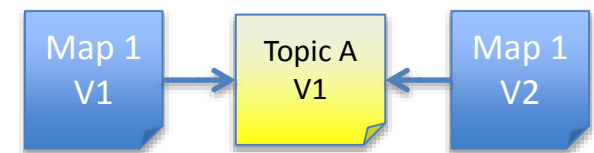
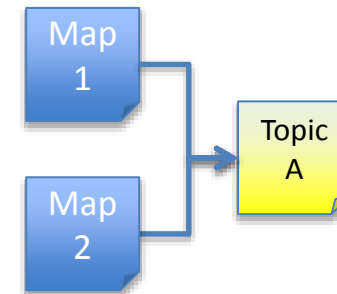
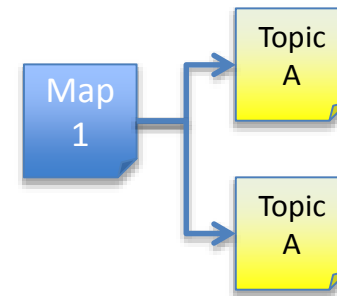
- Must be able to have same reference resolve to different topics in a given use context
- Thus, direct addresses (direct URLs) won't work
- DITA added key/keyref feature in 1.2 to satisfy this requirement



Different Use Contexts



- Same component used multiple times in the same hyperdocument
- Same component used in different hyperdocuments
- Same component used in different versions in time of a given hyperdocument



Direct vs. Indirect Addressing



Direct addressing



- Quick, effective, fragile.
- Relatively easy to learn and execute
- Predetermined response to a given attack
- Death results

Indirect addressing



- Blend and redirect to appropriate target
- Harder to learn and execute but more effective
- Many options at time of action
- Death does not result





HYPERDOCUMENT MANAGEMENT FUN



The Problems



- As an author: What can I link to and how do I address it?
- As an authoring tool: What does this indirect address point to?
- As a deliverable producer: What is the set of resources I require in order to produce a deliverable from the input publication source?
- As a manager: What is the version-specific configuration of this publication in a specific repository access context?



The Essential Issue



- Given a collection of source components with links among them and managed through asynchronous revision processes, what is the time-specific configuration of those components at any moment in time as viewed by a given agent for a specific purpose?
- In DITA terms:
 - **When I process a map in a specific access context, what *do* I see and what *can* I see?**



Hyperdocument Management Requirements



- Must understand the hyperdocument nature of the content:
 - What links to what
 - Answer the where-used question
 - Manage addresses: what keys are defined? What IDs are available
- Version-aware link management
 - Ability to resolve links to specific versions of resources
 - Fundamentally a configuration management concern
 - Requires branching
- Link and address information services for use from authoring and processing tools



WHAT TO DO?



Apply DITA Aikido



Aikido For Writers

- Attacked from all directions:
 - Topic-based authoring!
 - Maps!
 - Keys!
 - Content references!
- Relax
- Maintain your center
- Breath
- Pick your enemy and engage



Managers Must Be Prepared

- Understand that the authoring and production environment is fundamentally different
- Realize that the added value of DITA may appear to require additional work, especially at first
- Understand that there may be new roles in the organization
 - Information Architect
 - Document configuration manager
- Understand that DITA is not a product



Writers Must Be Prepared

- Prepare them for the move from "simple" publication structures to hyperdocuments
 - Help them understand that what they are doing is more sophisticated than before
- Specific training on DITA details
- Discussion and development of new practices and processes
 - Hyperdocument authoring requires communication and coordination among team members



Provision Appropriate Tooling

- Document types appropriately configured and constrained
- Authoring tool configured and adapted
- Component content management system to support linking and addressing
- Deliverable generation to produce good-quality results
 - Don't give authors a reason to blame DITA for ugly output





To Summarize

- Sophisticated technical documentation is inherently challenging
- DITA reveals that complexity, drawing the focus of frustrated writers
- Writers often do not understand the sophistication of what they are being asked to do
- Must set expectations and provide appropriate support to avoid frustration
- Connection and engagement are ki



Questions?

