

The Discourse of Component Content Management on Technical Communication Practitioner Blogs

Yvonne Cleary
University of Limerick
yvonne.cleary@ul.ie

Abstract - This study explores discursive blog posts about component content management by technical communication practitioners. It identifies blogs and posts that explore component content management practices over a one-year period (2018). The study offers an overview of current approaches to, and perspectives on, CCM by synthesizing this community's blog contributions. The complexity of some approaches are causes for concern. While many bloggers are committed to their current content development methodology, others propose possible alternative, and less complex, solutions. The findings indicate that bloggers are aware of developments and trends, and are strategizing for the future.

Index Terms – content development, DITA, blogs, discourse.

INTRODUCTION

Component content management (CCM) refers to the process of developing, storing, and managing pieces (or components) of content, rather than long-form documents. CCM is based on the concept of single sourcing, writing content once, maintaining it in a unique location, and reusing it multiple times. This approach to content development involves structured authoring, and is widely used in industry. Its adoption has gained “critical mass” [1, p. 248]. CCM has many advantages over traditional content development methods, including reusability and personalization for different output media, audiences, and user needs.

Notwithstanding the prevalence of this paradigm in practice, many academic programs do not include CCM in their courses, and academic articles about teaching CCM have only recently begun to issue. Because practitioners have more advanced knowledge in this area, academics can learn much about this aspect of practice, its advantages and disadvantages, and potential alternatives, by examining experts’ insights.

Although there are many structured authoring methods, DITA (Darwin Information Typing Architecture) is a common approach.

This paper analyses the professional discourse about CCM among bloggers in technical communication and related fields over a one-year period (January to December 2018). The blog sample has been selected by filtering listings of relevant blogs in [2] and [3] to identify bloggers who write about these subjects. Analysis focuses on content in the blog posts. The findings provide insights about current CCM approaches, as well as possible alternatives for content development. These insights facilitate the development of technical communication curricula that prepare students for the workplace.

The paper begins by exploring relevant literature on component content management, discourse in blogs, and technical communication blog studies. It then outlines the methodological approach, including the blog and post sample selection and the analytical procedure. Finally, it presents and analyses key findings, draws conclusions, and offers recommendations for teachers and researchers.

LITERATURE REVIEW

I. Previous studies of component content management

Component content management has become a dominant content development approach. However, until recently, few academic studies examined CCM. In the past five years, systematic literature reviews (see, for example, [1], [4], and [5]) have examined current CCM practices and implementations, including for multilingual delivery and in terms of roles and skills required. Other studies have outlined ways in which CCM can be incorporated into academic curricula (see, for example, [6] and [7]).

CCM has several advantages for large organizations that produce a lot of documentation. It results in content that is reusable, current, consistent, targeted to users’ specific contexts, and linked to other relevant information [8]. Content that is dynamic responds to users’ needs, devices, and tasks [9]. Implementing CCM increases consistency and accuracy [1]. Advantages for authors,

include increased productivity and less complex reviews [8].

Nevertheless, it is acknowledged that CCM, and DITA implementations in particular, can be problematic because of the complexity of the architecture [1], [10].

Alternatives to DITA, including Lightweight DITA and other lightweight markup languages, are emerging [10], [11].

Lanier [12] surveyed technical communication practitioners to determine important technical communication developments in the previous five years. DITA and related aspects of CCM were mentioned by 157 respondents (23%), indicating that this is a live issue for practitioners.

II. Discourse in blogs

Myers argues that blogs do not have specific linguistic features, but “some styles come to be favored for some uses” [13, p. 19]. Luzón conducted a study of blogs that communicate new scientific research, and found that the blogs were “heterogeneous in terms of producers, purposes, types of content, [...] and audiences” [14, p. 434]. Blogs do have features that distinguish them from other social media. For example, they are “time bound” (i.e. date stamped and in reverse chronological order [13, p. 75]. Unlike Twitter where tweets have character limits and statements can be curt, blogs permit deeper examinations of subjects—though of course not all blog posts are profound.

Luzón [14, p. 447] listed strategies bloggers use to engage readers, including self-disclosure, conversational tone, inclusive pronouns, references to the reader, personal opinions, humor, and questions.

Unlike Twitter, where fast-paced conversations coalesce around hashtags [15], individual blogs and posts can seem to exist in a vacuum. The comments function facilitates conversation, but posts are disconnected, and emerging patterns are difficult to discern.

This study aims to identify and highlight emerging patterns in blog posts about CCM in particular, and proposed alternative content development trends more generally.

III. Relevant blog studies

Studies of technical communication blogs are limited. In other professional fields, studies of blogs help to explain practice and “develop the kinds of shared understanding that is necessary for communicating well with members within a diverse community of practice” [16, p. 329].

Clery [17] examined technical communication practitioner blogs about professionalization. Her findings demonstrated that practitioners actively discuss their profession and are invested in promoting and developing the field. The bloggers in her study were strategizing about the future of technical communication, and she

recommended that academics should find ways to leverage the knowledge of bloggers, including by studying and participating in the discourse community.

Hannah and Lam [16] explored technical communication bloggers’ knowledge-sharing strategies. Their study suggested ways in which practitioners and academics could use blogs to connect. They recommended further studies of blogs using qualitative methods. Hannah and Lam’s study touched on bloggers’ reactions to aspects of CCM, including DITA, but no previous study has systematically examined bloggers’ reflections about component content management or related content development practices.

Because practitioners have more advanced knowledge in this area, academics can learn much about this aspect of practice, and potentially about what to teach and how to teach it, by examining experts’ insights.

METHODOLOGY

This section describes the methodological approach adopted for this study.

I. Research questions

Based on the foregoing discussion, the questions this study sets out to explore are:

- 1) What patterns can be discerned about structured authoring and CCM discourse from blog posts?
- 2) What alternative content development paradigms, if any, emerge from the blogs?

II. Selecting the blog and post sample

In order to respond to the research questions, I selected an initial blog sample by filtering listings of relevant blogs, in [2] and [3], to identify prominent and relevant bloggers in technical communication.

Technical communication blogs are primarily authored by:

- 1) Individual practitioners. These are personal blogs characterized by discursive posts.
- 2) Corporate-affiliated bloggers (service or software providers). These blogs are sometimes used to promote the organization. Marketing posts or posts about tools the company develops or promotes are characteristic.
- 3) Individual academics. Less than 2% of the blog posts Hannah and Lam [16] identified were authored by academics.

I selected 15 blogs for analysis based on the following criteria:

- Each blog was maintained by an individual practitioner, and not affiliated to a corporation.
- Each blog was active during the period studied (January to December 2018).

- Each blog included posts on subjects relevant to structured authoring and CCM. I reviewed blog tags and categories to identify content themes.

Within the 15 blogs, I reviewed post titles and synopses (or first paragraphs) to identify blog posts related to the research questions. Hannah and Lam [16] and Cleary [17] identified common subjects that technical communication bloggers discuss, including communication, technology, collaboration, and professionalization. The posts in my study may include several of these subjects. Furthermore, Hannah and Lam developed a typology of post types, including process, argument, how-to, news, and research. I filtered the posts to include only those in the ‘argument’, ‘news’, and ‘research’ classes.

III. Limitations

The sampling strategy resulted in a somewhat homogeneous selection of bloggers: primarily (though not exclusively) white men based in North America.

The sampling strategy likely excluded some bloggers who write about CCM in technical communication. Filtering the posts by reviewing only the post title and first lines may have excluded some posts about CCM.

The final post sample arrived at is quite small. It lends itself to qualitative analysis, but cannot be said to be generalizable.

IV. Data analysis

This study adopts a discourse analysis approach. Discourse analysis is a suitable approach for this study because it enables pattern searching within a disconnected corpus of blogs. Discourse analysis is appropriate when working with large data volumes, and combining quantitative and qualitative analysis.

Apart from numbers of posts and comments, data are qualitative, and analysis entailed close readings of blogs and blog posts to identify relevant posts, further to explore the content of those posts.

FINDINGS

This section presents the relevant blogs and posts. In the Discussion section, these blogs and posts are analyzed qualitatively.

From the 15 blogs I initially reviewed, just six included relevant posts that could be classified as ‘argument’, ‘news’ or ‘research’ (see Table 1). An additional categorization not included in Hannah and Lam’s list, but relevant to this study, was ‘podcast transcript.’

TABLE 1. BLOG SELECTION.

Blog title	Number of relevant posts
Every page is page one	4
Scriptorium	4
Cherryleaf	4
Leading technical communication	2
Beyond the bleeding edge	2
The content philosopher	1

The total number of posts I analyzed in detail, therefore, was reduced to 17. Within these posts, the following discussions were common:

- Definitions of component content management, in general, and DITA, in particular (5 posts)
- Advantages of a structured authoring approach (4 posts)
- Problems (10 posts) with DITA (5 posts)
- Solutions/advice (7 posts)
- Potential alternatives (5 posts).

Table 2 shows the posts titles, categories, and number of comments.

TABLE 2. RELEVANT BLOG POSTS.

Blog	Post details		
	Title	Category	Comment
Every page is page one	Experts read more than novices	Definitions	4
	Time to move to multi-sourcing	Problems Solutions Alternatives	18
	Is personalized content unethical?	Problems Solutions	11
	Is single sourcing dead?	Problems Alternatives	1
Scriptorium	Single-sourcing is dead. Long live shared pipes!	Definitions Problems Alternatives	10
	High return on high-design in DITA	Problems Solutions	0
	Managing DITA projects: Five keys to success	Advantages Advice	0
	The DITA	Advantages	0

Blog	Post details		
	Title	Category	Comment
	business case: Maximizing content value		
Cherryleaf	Transcript of our podcast episode Topic-based writing: what is it, and why should I care?	Definitions Advantages	0
	Transcript from What's the deal about structured content? Part 1/2	Definitions Problems	0
	Transcript from What's the deal about structured content? Part 2/2	Problems	0
	Transcript of Trends in Technical Communication in 2018 and beyond	Alternatives	0
Leading technical communication	Lightweight DITA: I've seen the light	Advantages Problems Solutions	3
	Carrying the earth on our shoulders	Problems Alternatives	3
Beyond the bleeding edge	Is Single-Sourcing Dead?	Problems Solutions	18
	Is Single-Sourcing Dead or Alive – the Debate Continues	Problems Solutions	34
The content philosopher	The Marriage of Structure and Semantics	Definitions Problems	2

DISCUSSION

I. What patterns can be discerned about structured authoring and CCM discourse from blog posts?

Bloggers expressed commitment to structured authoring, and acknowledged the need to develop reusable content, as expressed in several explanations of the advantages of this approach:

Each DITA element represents what a thing is (hence the term semantic) rather than how it looks. Just think: you can take a big document and generate a list of all the command names, or all the screen names. You can't do that when you're just tagging things as boldface and italics. *Leading technical communication, March 2018*

A Scriptorium post explained the advantages of a DITA implementation for localization:

The most immediate benefit was with localization. Content and string reuse reduced localization overhead. In addition, the PDF transform handles formatting automatically, eliminating the lengthy hand-layout portion of the localization workflow. This reduced the cost per language enough that the client was able to localize into all of their target languages. [...] With fewer workflow bottlenecks, productivity has increased. *Scriptorium, November 2018*

Considering that the audience of these blogs is likely to comprise technical communicators, basic definitions should not be needed. Nevertheless, the number of posts that included definitions of basic principles suggests that the bloggers did not expect readers to fully understand the subject matter:

The single-sourcing concept – write once, re-use many times in many ways and many places – has some problems. But the basic concept is so useful that I see nothing that can replace it yet. *Beyond the Bleeding Edge, September 2018*

Back in the 1990s, single-sourcing meant creating content in one format and then converting that content into another format. *Scriptorium, April 2018*

In DITA, a piece of content isn't boldface or italics. It's a command name. *Leading technical communication, March 2018*

The need for explanations of a mature architecture that has been evolving in technical communication for two decades underscores a common concern among many of the bloggers, that structured authoring, and DITA especially, is complex.

We have a complex landscape with lots of different places to create, store, manage content. Different software vendors are promoting different ways of doing things. *Scriptorium, April 2018*

Single-sourcing requires many tasks beyond just writing the content. Authors have to decide which output is primary in order to decide which features to use because some won't work well or at all on different outputs. That means understanding those features. Authors have to create and assign conditions

to control which content to use for which output. Define re-usable chunks of content. Create style sheets that behave differently depending on the output. Perhaps define microcontent. And more. And this all must be documented somewhere for reference by the current authors and later ones. The result? The increasing power of our tools and increasing customer demands are leading to increasingly complex projects that that can easily go out of control. *Beyond the Bleeding Edge, September 2018*

A DITA project has many moving parts: content modeling, tool evaluation, customizing delivery options, conversion, and more. *Scriptorium, November 2018*

In addition to complexity, other concerns were about restrictions on author creativity, and the need to plan for legacy content:

The posts with the most engagement (see Table 2) envisaged alternatives to DITA. Conversations began and unfolded among bloggers, who commented on one another's posts, and wrote posts in response to one another, indicating a lively discourse community.

II. What alternative content development paradigms, if any, emerge from the blogs?

Bloggers expressed a need for a return to more straightforward processes and less variety in tools.

Lightweight DITA was variously explained, championed, and rejected in blogs. This quote is from a blog post presenting a transcript of an interview with Carlos Evia and Michael Priestly, co-chairs of the Lightweight DITA (LwDITA) sub-committee.

One of the drivers for creating Lightweight DITA was really looking at where DITA as an existing standard was hitting resistance. DITA, the full DITA standard is working great for the group that have adopted it and it's still getting adopted by new groups today. Where it was encountering resistance, there were two main reasons that we thought we could maybe address. One was, there was a perception that even if you started simple in DITA, the simple starting point was not simple enough. Could we make that starting point even simpler? The second one was when you have groups who would consider DITA and they like the capabilities it brings, but they are really tied to an existing authoring format, and an existing tool stack around that format. *Scriptorium, August 2018*

Larry Kunz was positive about the potential of LwDITA: I like it. Its simplicity and flexibility won't be ideal for every situation, but much of the time they'll be just right. *Leading technical communication, March 2018*

However, comments on his post (by blogger Mark Baker) were skeptical about whether LwDITA could deliver the simplicity and robustness required:

While [DITA] is riding a wave of undeniable enthusiasm at the moment, I think it stands at a fork in the road. On the one hand, we are seeing alternative less structured systems for topic-based reuse developed. On the other, there is definitely a growing demand for highly semantic content which DITA may not be best structured to deliver, and for which its own rhetoric of standardization may work against it. LwDITA seems to be an attempt to let it take both roads, but it is not clear that it is the strongest candidate for either road.

Baker is developing an alternative lightweight markup language, which may explain his lack of enthusiasm for LwDITA. In September 2018, in a post entitled "Is single sourcing dead?", he argued that one missing piece in the current array of tools, skills, applications, and standards, is "a lightweight semantic markup language. That is why I am working on the SAM project."

In a post entitled "Time to move to multi-sourcing", Baker argued that "single sourcing has been the watchword of technical communication for the last several decades" but "we have never fully made it work" This post was in response to "seminal posts by prominent members of the community" (including one by Sarah O'Keefe writing for the Scriptorium blog) and advocated a move towards "multi-sourcing".

The O'Keefe post that Baker was responding to was entitled "Single sourcing is dead. Long live shared pipes!" This post proposed "a shared infrastructure for terminology, information architecture, and localization."

Prominent bloggers whose blogs were among the initial sample I reviewed, advocated "docs as code" or related approaches to content development. According to the Write the Docs community [18], "docs as code" refers to "writing documentation with the same tools as code" including plain text markup, code reviews, version control, and issue trackers. Anne Gentle, whose blog JustWriteClick is identified as an influencer in [3], writes about Docs as Code. Tom Johnson, one of the most prominent bloggers in technical communication has adopted Jekyll for content development. He writes about this approach extensively in his blog.

CONCLUSIONS

Blogs have scope to be disruptive spaces. The findings from this analysis suggest that practitioners are committed to structured authoring, and recognize the benefits of single sourcing, but are reflexively exploring less complex alternatives to DITA as a development approach.

Hannah and Lam have noted that academics can learn from practitioners and this study underscores the value of

analyzing practitioner discourse. Vibrant and deep discussions and conversations are taking place about the future of content development. It is important to catalogue, analyze, and respond to this practitioner discourse, and academics have a role to play in that process.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the financial support of the Faculty of Arts, Humanities, and Social Sciences at the University of Limerick.

REFERENCES

- [1] R. Andersen and T. Batova, "The current state of component content management: An integrative literature review," *IEEE Trans. Prof. Commun.*, vol. 58, no. 3, pp. 247–270, 2015.
- [2] Society for Technical Communication. "Other Blogs." <https://www.stc.org/notebook/links/> (accessed April 01, 2019).
- [3] Mindtouch. "Announcing 2017's TOP25 Content Experience Influencers and TOP200 Strategists." <https://mindtouch.com/resources/2017-top25-content-experience-influencers-top200-strategists> (accessed April 01, 2019).
- [4] T. Batova, "Component content management and quality of information products for global audiences: An integrative literature review," *IEEE Trans. Prof. Commun.*, vol. 57, no. 4, pp. 325–339, 2014.
- [5] T. Batova and R. Andersen, "A systematic literature review of changes in roles/skills in component content management environments and implications for education," *Tech. Commun. Quart.*, vol. 26, no. 2, pp. 173–200, 2017.
- [6] T. Batova, "Global technical communication in 7.5 weeks online: Combining industry and academic perspectives," *IEEE Trans. Prof. Commun.*, vol. 61, no. 3, pp. 311–329, 2018.
- [7] C. Evia *et al.* "Teaching structured authoring and DITA through rhetorical and computational thinking," *IEEE Trans. Prof. Commun.*, vol. 58, no. 3, pp. 328–343, 2015.
- [8] J. T. Hackos, *Content Management for Dynamic Web Delivery*. New York: Wiley, 2002.
- [9] A. Rockley and J. Gollner, "An intelligent content strategy for the enterprise," *Bulletin of the American Society for Information Science and Technology*, vol. 37, no. 2, pp. 33–39, 2011.
- [10] C. Evia, *Creating Intelligent Content with Lightweight DITA*. New York: Routledge, 2019.
- [11] M. Baker, *Structured Writing: Rhetoric and Process*. Laguna Hills, CA: XML Press, 2018.
- [12] C. Lanier, "Toward understanding important workplace issues for technical communicators," *Tech. Commun.*, vol. 65, no. 1, pp. 66–84, 2018.
- [13] G. Myers, *The Discourse of Blogs and Wikis*. London, UK: Continuum, 2011.
- [14] M. J. Luzón, "Public communication of science in blogs: Recontextualizing scientific discourse for a diversified audience," *Written Communication*, vol. 30, no. 4, pp. 428–457, 2013.
- [15] R. Page, *Narratives Online: Shared Stories in Social Media*. Cambridge, UK: Cambridge University Press, 2018.
- [16] M. Hannah and C. Lam, "Patterns of dissemination: Examining and documenting practitioner knowledge sharing practices on blogs," *Tech. Commun.*, vol. 63, no. 4, pp. 328–345, 2016.
- [17] Y. Cleary, "Discussions about the technical communication profession: Perspectives from the blogosphere," *Tech. Commun.*, vol. 59, no. 1, pp. 8–28, 2012.
- [18] Write the Docs. "Docs as Code." <https://www.writethedocs.org/guide/docs-as-code/> (accessed April 01, 2019)

ABOUT THE AUTHOR

Yvonne Cleary is a lecturer in Technical Communication and Instructional Design, and Program Director for the MA in Technical Communication and E-Learning at the University of Limerick, Ireland. She holds an MA and a PhD in technical communication. Her research interests include professional issues in technical communication, technical communication pedagogy, virtual teams, and international technical communication. She has presented her work at conferences in Europe and the United States and has published in leading journals in the technical communication field.