December 11, 2019

The Future of DITA

Keith Schengili-Roberts
Information Architect
at Enbridge

About Keith

Information Architect at Enbridge

Member of OASIS DITA Technical Committee and LwDITA Sub-committee

Chair of OASIS DITA Adoption Committee

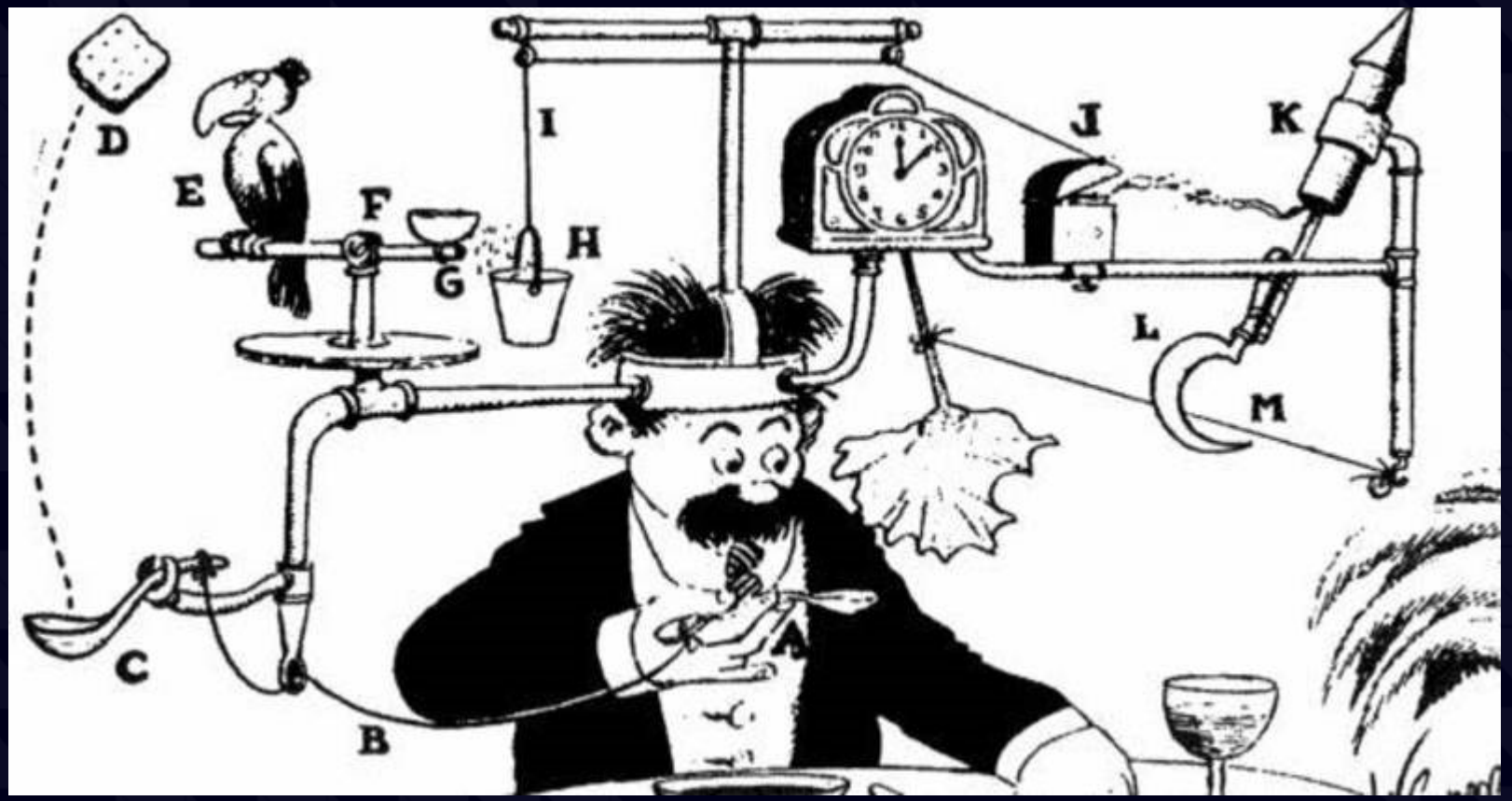
14+ years of experience with DITA XML

Lecturer on Information Architecture at the University of Toronto

Agenda

1. The process of making DITA
2. DITA 2.0
3. Lightweight DITA
4. New multimedia elements
5. The possible futures of DITA and structured content

The process of making DITA



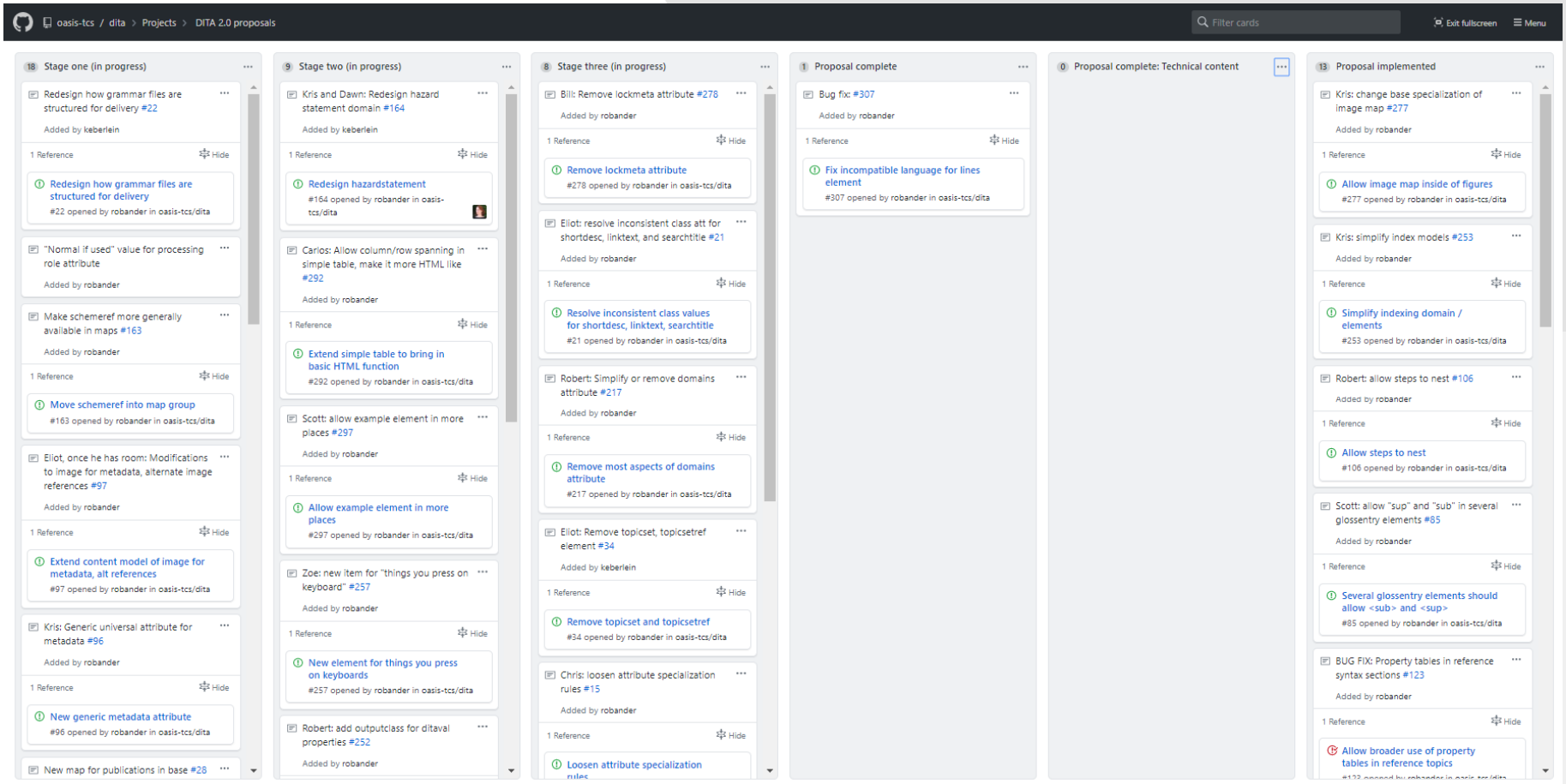
The Role of the DITA Technical Committee

- The chief organizational body overseeing the development of the DITA standard within OASIS
- Chaired by Kris Eberlein
- Holds hour-long weekly meetings every Tuesday
- Members (which includes IXIASOFT) get a say plus vote on new features/ developments
  - Non-members are invited to contribute ideas on the TC's email list

Lightweight DITA: Process

- Currently being defined by a subcommittee (SC) belonging to the main DITA TC; runs every other Monday
- Co-chaired by Professor Carlos Evia (Virginia Tech) and Michael Priestley (IBM)
- Status: Committee Note v1.0 is currently available
  - Committee Note outlines LwDITA elements and attributes
  - Comes with code samples
  - DITAWriter LwDITA code samples also available: github.com/DITAWriter/LwDITA_Code_Samples

DITA 2.0 Triage Process



- https://github.com/oasis-tcs/dita/projects/2?fullscreen=true

The Big Question

- When will DITA 2.0/LwDITA 1.0 be released?

Kris Eberlein has gone on the record as saying:

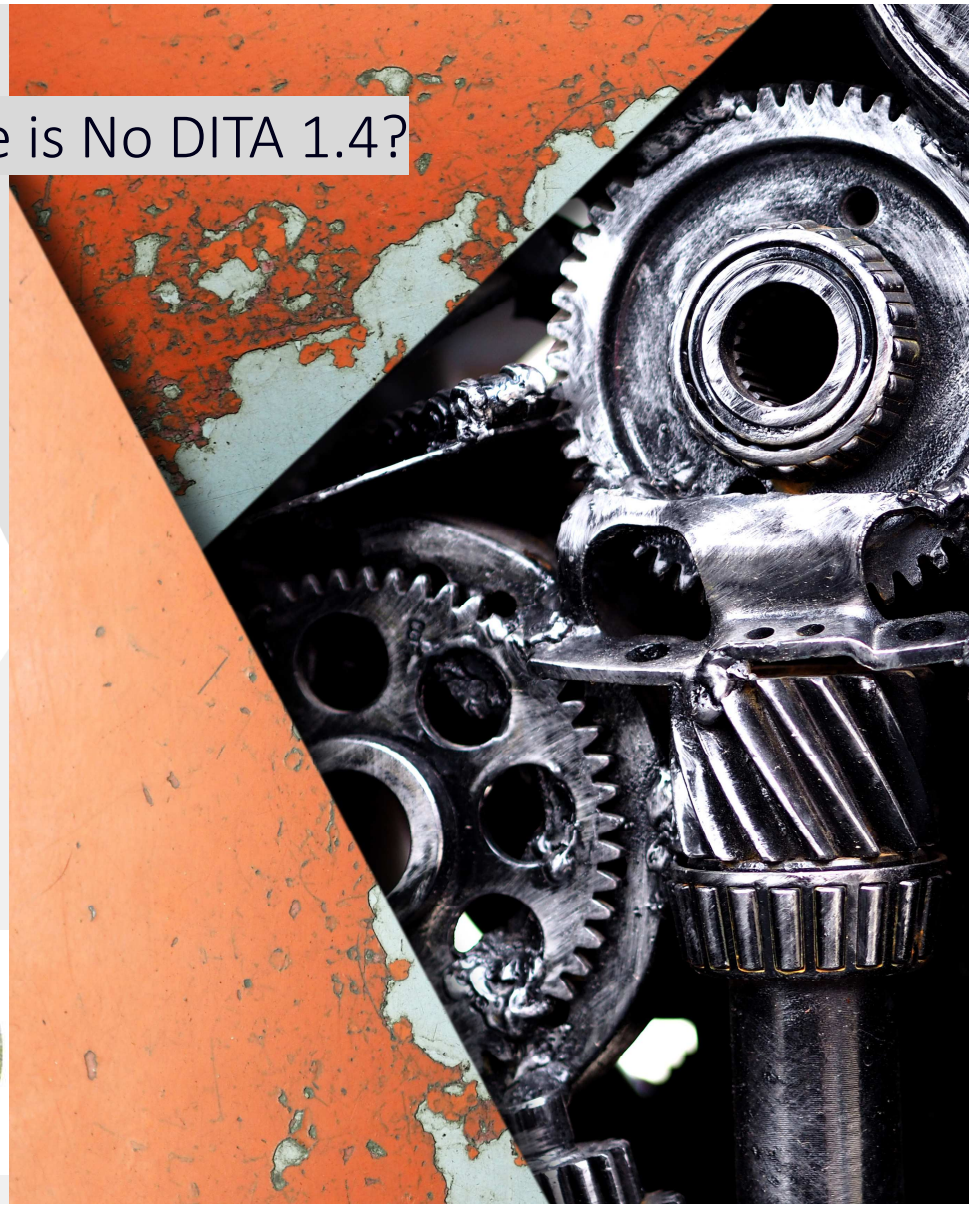
- Late 2020 or early 2021
- LwDITA 1.0 specification will likely be launched in the same timeframe

Another Question: Why is There is No DITA 1.4?

Reason: technical debt

- Some design decisions made in previous versions of DITA did not pan out ("if we knew then what we know now")
- There is a need to strip out and deprecate things that are obsolete that could stop progress in new areas

"When you keep adding cars, it's harder to pull the train. Authors don't know which cars to use, and it's also harder to adapt to new conditions, new tracks"
- Robert Anderson, IBM



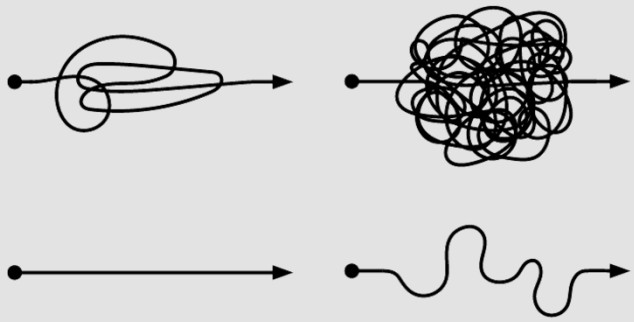
Why is This Taking So Long?

- Remember that this is a volunteer effort; OASIS members contribute their time and effort freely to the cause
- Much of time is spent examining additions/changes; this is still ongoing
- Once a draft specification is created, there is an extensive vetting process, which takes the better part of a year

DITA 2.0

Some of the Goals for DITA 2.0

- Simplification
- Reduce complexity
- Remove unused features
- Redesign hard-to-use features
- Where sensible: only one way to do things



DITA 2.0 Will *Not* Be Backwards Compatible

Includes the following changes:

- @print (replaced by @deliveryTarget)
- @keyref on <navref> removed
- @xtrf, @xtrc removed
- “delayed conref domain” will be removed
- Previously deprecated items from DITA 1.3 will be removed, including:
  - <boolean>, <indextermref>
  - @alt, @longdescref, @navtitle, (replaced with element-based equivalents), @query

Implemented DITA 2.0 Improvements

"Implemented" in this case means that it is already incorporated into the draft code for DITA 2.0

- @outputclass will be a universal attribute
  - Makes it possible to modify just about anything at output
- Nested steps
  - This replaces <substep> with <steps>, makes conref-ing easier
- Chunking has been redefined
  - Easier to understand and to implement
- Allow image map inside of figures
  - Changes specialization base of image maps from <figure> to

Implemented DITA 2.0 Improvements (Cont.)

- Simplify index models
  - Adds @outputclass to DITAVAL properties, for associating formatting specific to the selected value
- New vocabulary element for inclusion of external XML/text markup
  - New <include> element now used as a base for specialization for these types of elements
- Make @audience, @platform, @product, and @otherprops into specializations
  - Makes it easier to provide further specializations of these as needed, such as defining @appserver and @database as specialized variants of @product
- Glossentry elements will allow <sub> and <sup>
  - <ph> could not be used in several elements from glossentry; now it (and its specializations) can

Current DITA 2.0 Proposals Include:

- Redesigning hazard statement domain
- Extend simpletable to add HTML-like features
- Allow <example> in more places
- New element or domain for “things you press on keyboard”
- Add <titlealts> to maps
- Split base and technical content
- Remove <topicset>, <topicsetref>
- Modify bookmap design to allow <ditavalref> before front matter, as well as a <keydefs> container to hold key definitions (“publicationmap”)
- Deprecate note type="fastpath"
- Add and elements, redefine <b> and <i> to more closely match HTML5
- Addition of multimedia domain (more on this later)

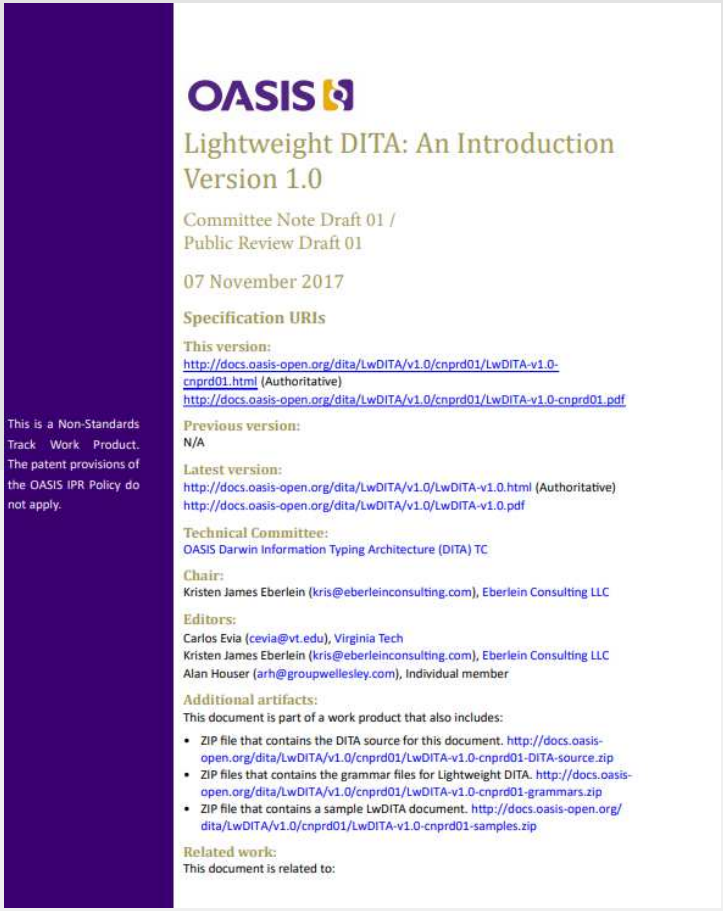
So Where's My Jetpack?

- It's less about revolutionary ideas, and more about evolutionary advancements
- Better to focus on real-world issues, and remove the cruft that has built-up over the years
- There are still some interesting ideas that have not made it to the formal proposal stage yet, such as sub-topic level metadata, so there may be more to come!

Lightweight DITA

Lightweight DITA Highlights

- Fewer tags and attributes than “full” DITA 1.3; designed to be a “simpler DITA experience”
- Designed to be compatible with DITA 1.3; valid LwDITA code is also valid DITA 1.3 code *
- DITA is no longer necessarily bound to XML; three different “flavours” exist:
  - XML-based XDITA
  - HTML5-based HDITA
  - Markdown-based MDITA



This is a Non-Standards Track Work Product. The patent provisions of the OASIS IPR Policy do not apply.

OASIS

Lightweight DITA: An Introduction Version 1.0

Committee Note Draft 01 / Public Review Draft 01

07 November 2017

Specification URIs

This version:
http://docs.oasis-open.org/dita/LwDITA/v1.0/cnprd01/LwDITA-v1.0-cnprd01.html (Authoritative)
http://docs.oasis-open.org/dita/LwDITA/v1.0/cnprd01/LwDITA-v1.0-cnprd01.pdf

Previous version:
N/A

Latest version:
http://docs.oasis-open.org/dita/LwDITA/v1.0/LwDITA-v1.0.html (Authoritative)
http://docs.oasis-open.org/dita/LwDITA/v1.0/LwDITA-v1.0.pdf

Technical Committee:
OASIS Darwin Information Typing Architecture (DITA) TC

Chair:
Kristen James Eberlein (kris@eberleinconsulting.com), Eberlein Consulting LLC

Editors:
Carlos Evia (cevia@vt.edu), Virginia Tech
Kristen James Eberlein (kris@eberleinconsulting.com), Eberlein Consulting LLC
Alan Houser (arh@groupwellesley.com), Individual member

Additional artifacts:
This document is part of a work product that also includes:

- ZIP file that contains the DITA source for this document. http://docs.oasis-open.org/dita/LwDITA/v1.0/cnprd01/LwDITA-v1.0-cnprd01-DITA-source.zip
- ZIP files that contains the grammar files for Lightweight DITA. http://docs.oasis-open.org/dita/LwDITA/v1.0/cnprd01/LwDITA-v1.0-cnprd01-grammars.zip
- ZIP file that contains a sample LwDITA document. http://docs.oasis-open.org/dita/LwDITA/v1.0/cnprd01/LwDITA-v1.0-cnprd01-samples.zip

Related work:
This document is related to:

* Exception for the moment is the multimedia domain; more on this later…

Stripped-down DITA

<alt>
<body>
<b>
<data>
<dd>
<dlentry>
<dt>
<dl>
<desc>
<fig>
<fn>
<image>
<i>
<keydef>
<linktext>

<map>
<note>

<p>
<navtitle>
<ph>
<pre>
<prolog>
<section>
<shortdesc>
<simpletable>
<stentry>
<sthead>
<strow>
<sub>
<sup>
<title>
<topic>
<topicmeta>
<topicref>
<u>

<xref>

Multimedia

<audio>
<media-autoplay>
<media-controls>
<media-loop>
<media-muted>
<media-source>
<media-track>
<video>
<video-poster>

LwDITA vs. “Full” DITA

DITA 1.3 All-inclusive:	• 26 document types, 621 elements
DITA 1.3 Base:	• 4 document types, 189 elements
Lightweight DITA:	• 1 document type, 48 elements

Some General LwDITA Guidance

- LwDITA content is still valid DITA and can be incorporated into “full” DITA
- No *automatic* “round-tripping” between DITA and LwDITA
- Just map, no bookmap
- Mixed content not allowed; all text must be in a <p>
- No CALS table elements (i.e. <table>, <row>, <entry>, etc.)*
- It is really “DI-A”, as content is not typed (result is a generic topic)

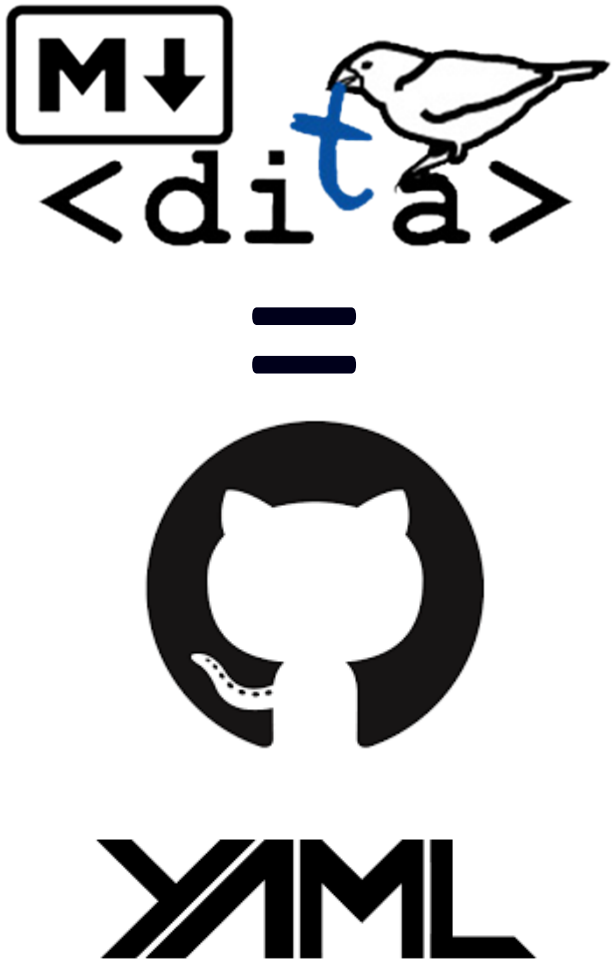
* But that’s not the end of the story

LwDITA Audiences / Scenarios

- XDITA: tech writers wanting reduced/simpler tagset, environments where there is interchange between XML + Markdown / HTML5
- HDITA: technical marketing, software developers, trainers, bloggers
- MDITA: software developers, “individuals authoring content quickly that must be later refactored as structured content”
  - This version is getting the most interest

How MDITA Aligns with Markdown

- There is no single standard for Markdown
- Base MDITA is derived from GitHub-flavored Markdown (GFM) with tables as an option
- “Valid, extended” MDITA includes GFM, YAML headers, and HDITA elements where no equivalent exists in Markdown
  - e.g. `<note>` becomes `<p data-hd-class="note">`



Simple MDITA Topic Example

capabilities_and_advantages.md - MarkdownPad 2

File Edit Insert View Tools Help

capabilities_and_advantages.md

YAML Header

```
---
id: capabilities_and_advantages
shortdesc: A brief overview of the additional features that adding the expansion interface to your TRS-80.
author: Reginald Sutton
---
```

Title

```
# Capabilities and Advantages

The Interface allows you to add the following Radio Shack modules to your system:

1. Screen Printer (26-1151)
2. Line Printer (26-1150)
3. Mini-Disk System (26-1160/26-1161)
4. Cassette Recorder number 2 (14-841)

The Screen Printer and Line Printer allow you to obtain hard copy (printed) information generated by your TRS-80.

The TRS-80 Mini-Disk System is a small version of the floppy disk. It provides vast storage space and much quicker
access time than tape. The number 1 disk contains about 80,000 bytes of free space for files. Each additional disk
has 89,600 bytes of file space. The Disk System has its own set of commands that allow manipulation of files and
expanded abilities in file use. The TRS-80 Mini-Disk System uses sequential or random access. The disks will allow
use of several additional LEVEL II commands.

<p data-hd-class="note">Because of the presence of a Disk Controller in the Expansion Interface, the computer will
try to input the additional commands.</p>

When the Expansion Interface is connected to the computer, it assumes that a Mini-Disk is connected. To use the
Expansion Interface without a Mini-Disk, press the BREAK key on the TRS-80 keyboard. This will override the Mini-
Disk mode and allow normal LEVEL II operation.

The use of two cassettes allows a much more efficient and convenient manner of updating data stored on tape. For
example, if you have payroll data stored on tape, the information can be read, one item at a time, from Cassette
Recorder number 1, then changed or added to and written out on Cassette Recorder number 2. The example cited is a
very simple application; however, very powerful routines can be constructed to allow input and output of data using
two tapes simultaneously.

<p data-hd-class="note">This unit is designed to be used with Level II only. Do not use with level I.</p>

![Image](figure_1.jpg)

FIGURE 1. Expansion Interface.*

| * Catalog Number | Description | RAM |
|---|---|---|
| 26-1140 | TRS-80 Expansion Interface | 0K |
| 26-1141 | TRS-80 Expansion Interface | 16K |
| 26-1142 | TRS-80 Expansion Interface | 32K |
```

Numbered List

Image

Table

Capabilities and Advantages

The Interface allows you to add the following Radio Shack modules to your system:

1. Screen Printer (26-1151)
2. Line Printer (26-1150)
3. Mini-Disk System (26-1160/26-1161)
4. Cassette Recorder number 2 (14-841)

The Screen Printer and Line Printer allow you to obtain hard copy (printed) information generated by your TRS-80.

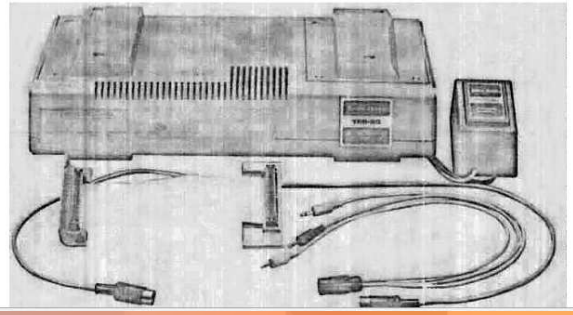
The TRS-80 Mini-Disk System is a small version of the floppy disk. It provides vast storage space and much quicker access time than tape. The number 1 disk contains about 80,000 bytes of free space for files. Each additional disk has 89,600 bytes of file space. The Disk System has its own set of commands that allow manipulation of files and expanded abilities in file use. The TRS-80 Mini-Disk System uses sequential or random access. The disks will allow use of several additional LEVEL II commands.

Because of the presence of a Disk Controller in the Expansion Interface, the computer will try to input the additional commands.

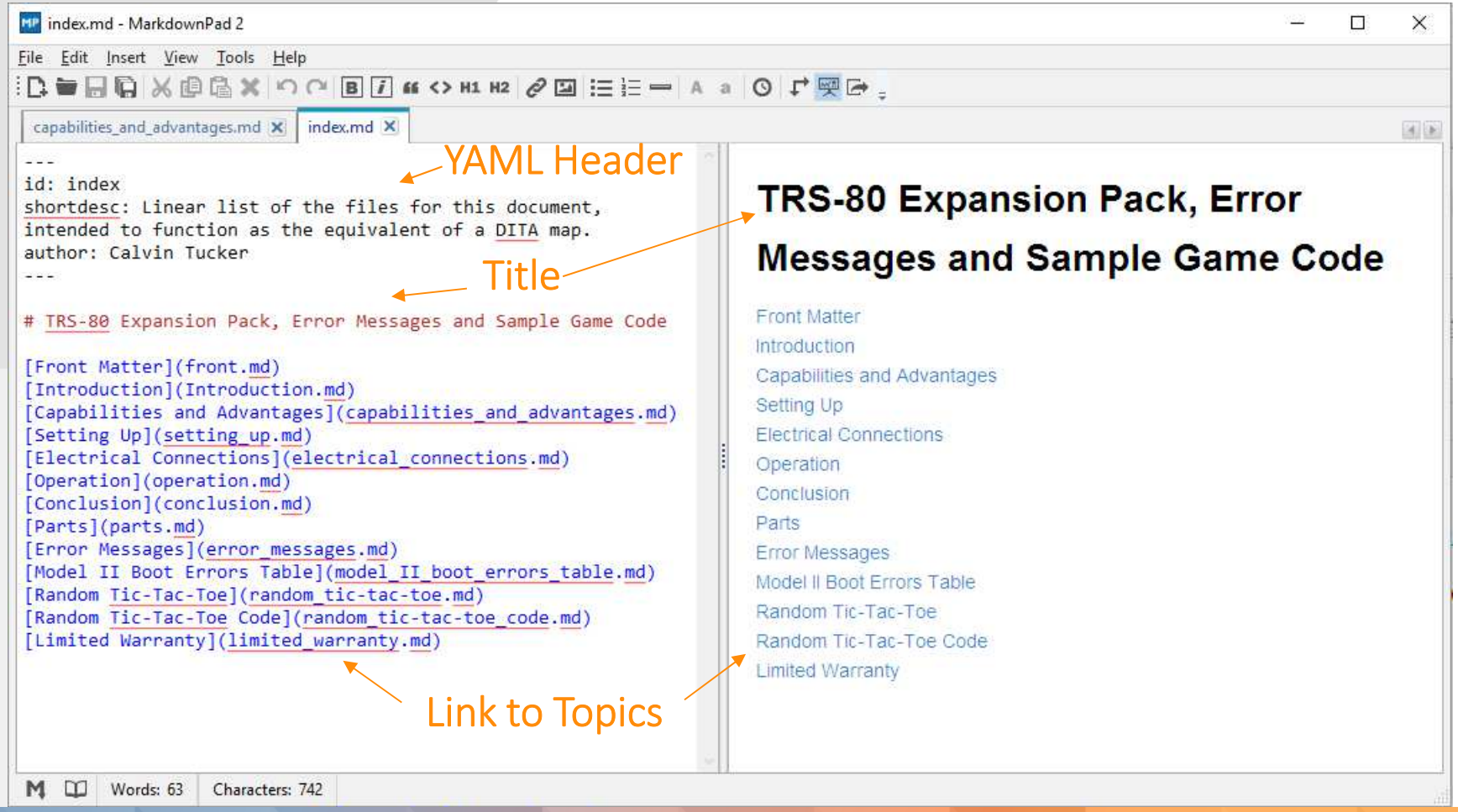
When the Expansion Interface is connected to the computer, it assumes that a Mini-Disk is connected. To use the Expansion Interface without a Mini-Disk, press the BREAK key on the TRS-80 keyboard. This will override the Mini-Disk mode and allow normal LEVEL II operation.

The use of two cassettes allows a much more efficient and convenient manner of updating data stored on tape. For example, if you have payroll data stored on tape, the information can be read, one item at a time, from Cassette Recorder number 1, then changed or added to and written out on Cassette Recorder number 2. The example cited is a very simple application; however, very powerful routines can be constructed to allow input and output of data using two tapes simultaneously.

This unit is designed to be used with Level II only. Do not use with level I.



Simple MDITA Topic Example



The Simpletable or CALS Table Debate

- Early on, LwDITA settled on using <simpletable>, arguing that if people needed CALS table formatting, they should use “full” DITA
- SC told by a few people that LwDITA would be “unusable” without CALS table formatting
- Compromise is to incorporate HTML5 table constructs into LwDITA tables; still being worked on

Lightweight DITA Bonus: Multimedia Controls!

- Arguably was an oversight to not include multimedia elements in DITA 1.3
- Initial push for this came from Lightweight DITA Sub-committee

- Current draft of the Committee Note includes nine additional multimedia elements
- Provides ability to add multimedia sound/video content in line with HTML5
- Original plan was to make this an addendum to DITA 1.3, now instead aiming for DITA 2.0

Multimedia:

```
<audio>
<media-autoplay>
<media-controls>
<media-loop>
<media-muted>
<media-source>
<media-track>
<video>
<video-poster>
```

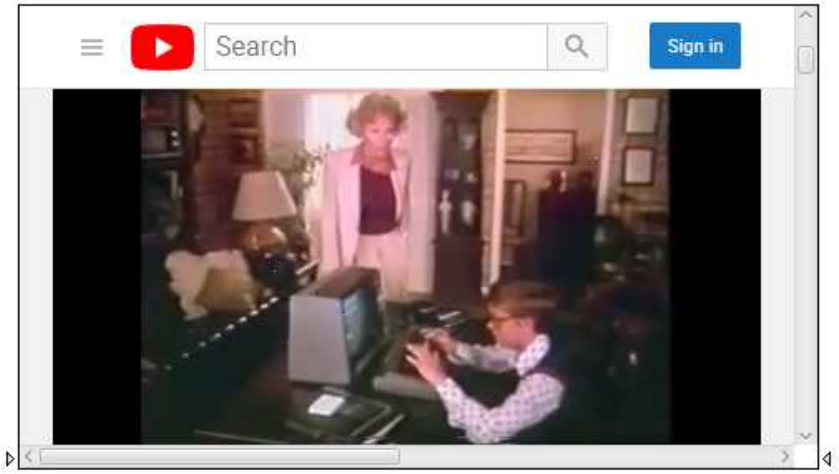
New multimedia elements

What's so new about multimedia and DITA?

- Nothing, and everything
  - <object> has enabled technical writers to add multimedia content, *awkwardly*, since DITA 1.0
- Newly-proposed multimedia elements intended to match those of HTML5
  - Will make it much easier and straightforward for technical writers to add multimedia content

Example of old vs. new ways of doing things

For more information, please see our instructive video about the [computer_model] TRS-80's capabilities:




```
<p>For more information, please see our instructive video about the <ph
      keyref="computer_model"/>'s capabilities: </p>
<p><object data="https://www.youtube.com/watch?v=0xW_4NXU3jI"
      outputclass="iframe"
   />
</p>
```

DITA 1.3 Code, using object element

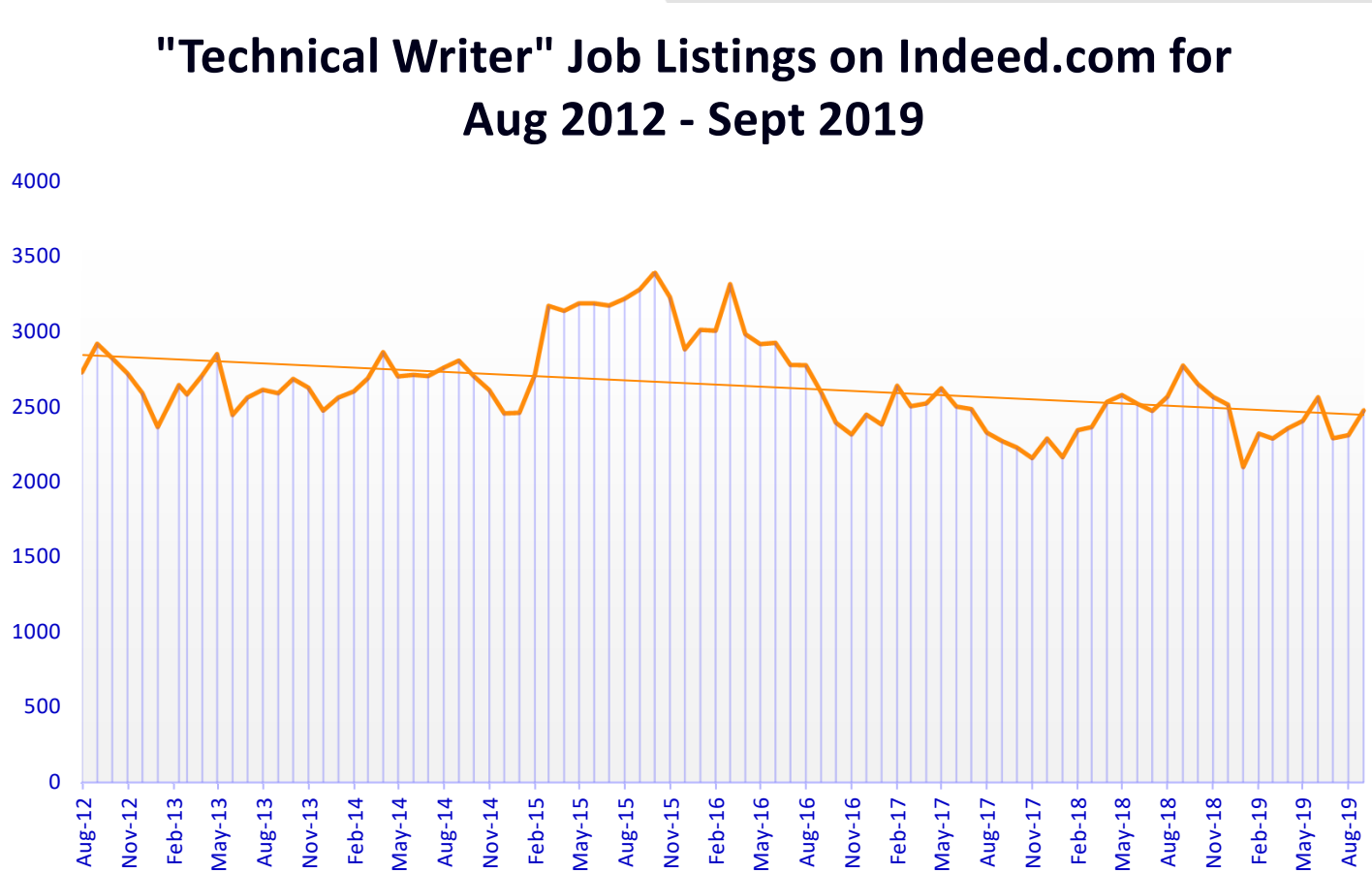
Example of working DITA and LwDITA (XDITA) code, available from github.com/DITAWriter/LwDITA_Code_Samples

```
<p>For more information, please see our instructive video about the <ph
      keyref="computer_model"/>'s capabilities:</p>
<video>
   <video-poster value="images/video_ad_still.png"/>
   <media-controls/>
   <media-source value="https://www.youtube.com/watch?v=0xW_4NXU3jI"/>
   </video>
```

LwDITA (XDITA) Code, using video element

The possible futures for DITA and structured content

The “Technical Writer” Job Landscape is Changing…



One Reason: SMEs are Producing More Upstream Content

- When it comes to API documentation, programmers are expected to provide much of the content. This is often framed and put into context by technical writers.
- The advent of Agile documentation processes in small software development teams means that, in some circumstances, SMEs had to write content.

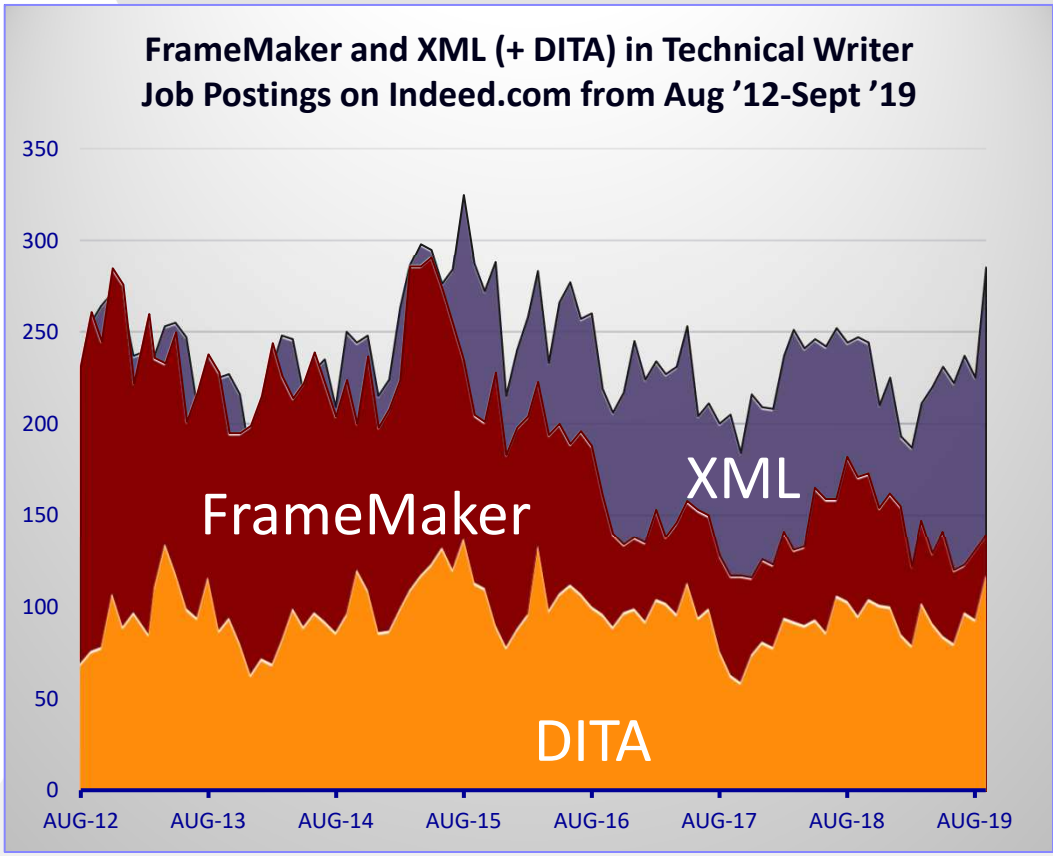
The Other Reason: The Role of the "Technical Writer" is Changing

In a survey I did of 1,500 LinkedIn profiles where people claimed to be using DITA, **66%** were not employed as traditional technical writers; some selected job titles:

- Applications Engineer
- Chief Information Architect, UX Analyst
- Consulting Content Strategist
- Content Architect
- Content Developer
- Content Management Specialist
- Content Strategist
- DITA Architect
- DITA Content Strategist
- DITA Information Architect
- DITA Migration Specialist
- Information Architect
- Information Developer
- Information Experience Manager
- Knowledge Architect
- Lead Information Developer
- Localization Program Manager
- Manager, XML CMS and L10n Systems
- Principal Content Experience Developer
- Principal Information Developer
- Product Architect
- Project Manager and Documentation Engineer
- Senior Content Developer
- Senior Content Strategist
- Senior Documentation Tools Developer
- Staff Information Architect
- Team Leader Technical Documentation
- User Assistance Development Architect
- UX Designer
- XML/DITA Coordinator

We are in the middle of a significant industry change

"Technical writer" jobs are being replaced by positions that use structured content to add more value by focusing on the content experience for users.



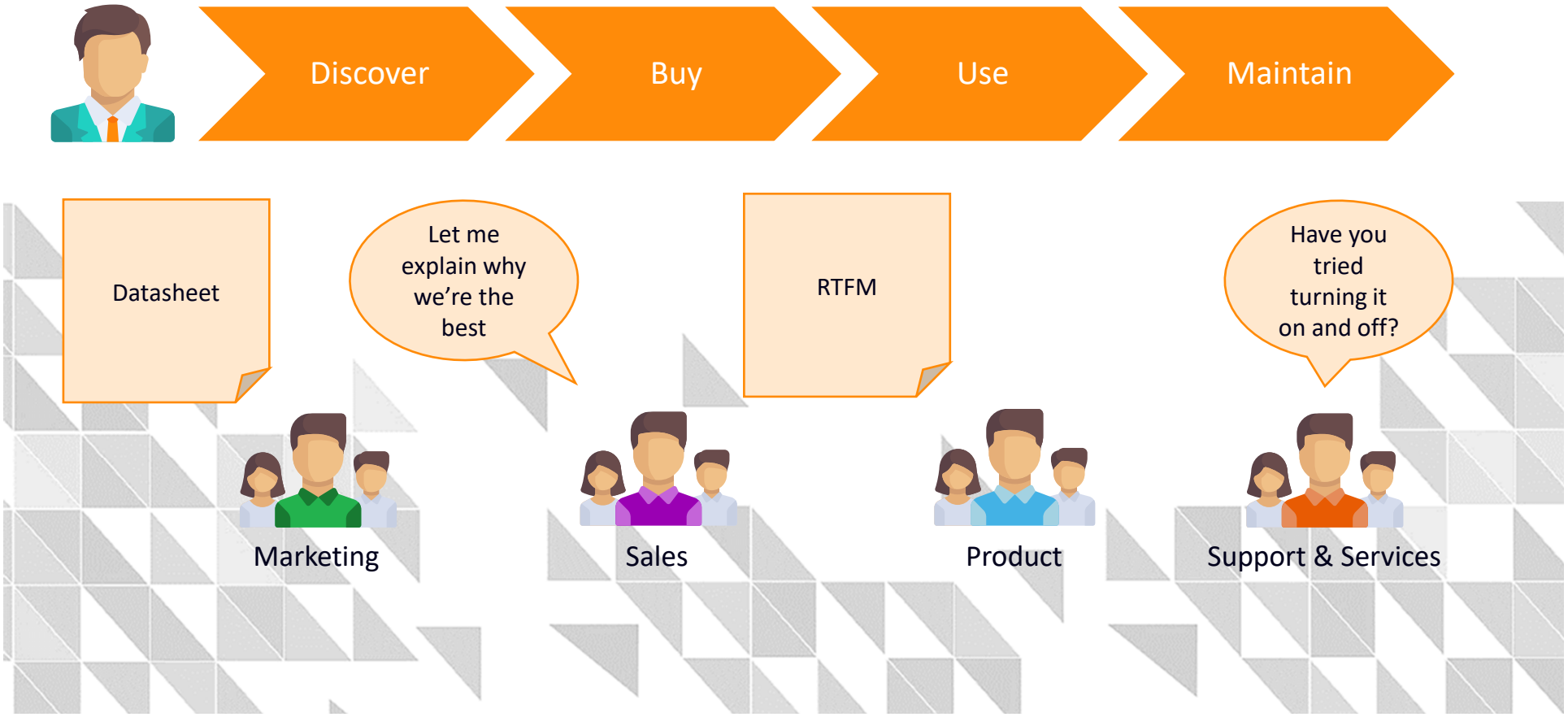
DITA is in a good position right now

In addition to all the things DITA was designed for, when done right, it can also do the following:

- Advance product SEO
- Provide a better ~~user~~ customer experience
- Topic/semantic structure works well with emerging system, like chatbots

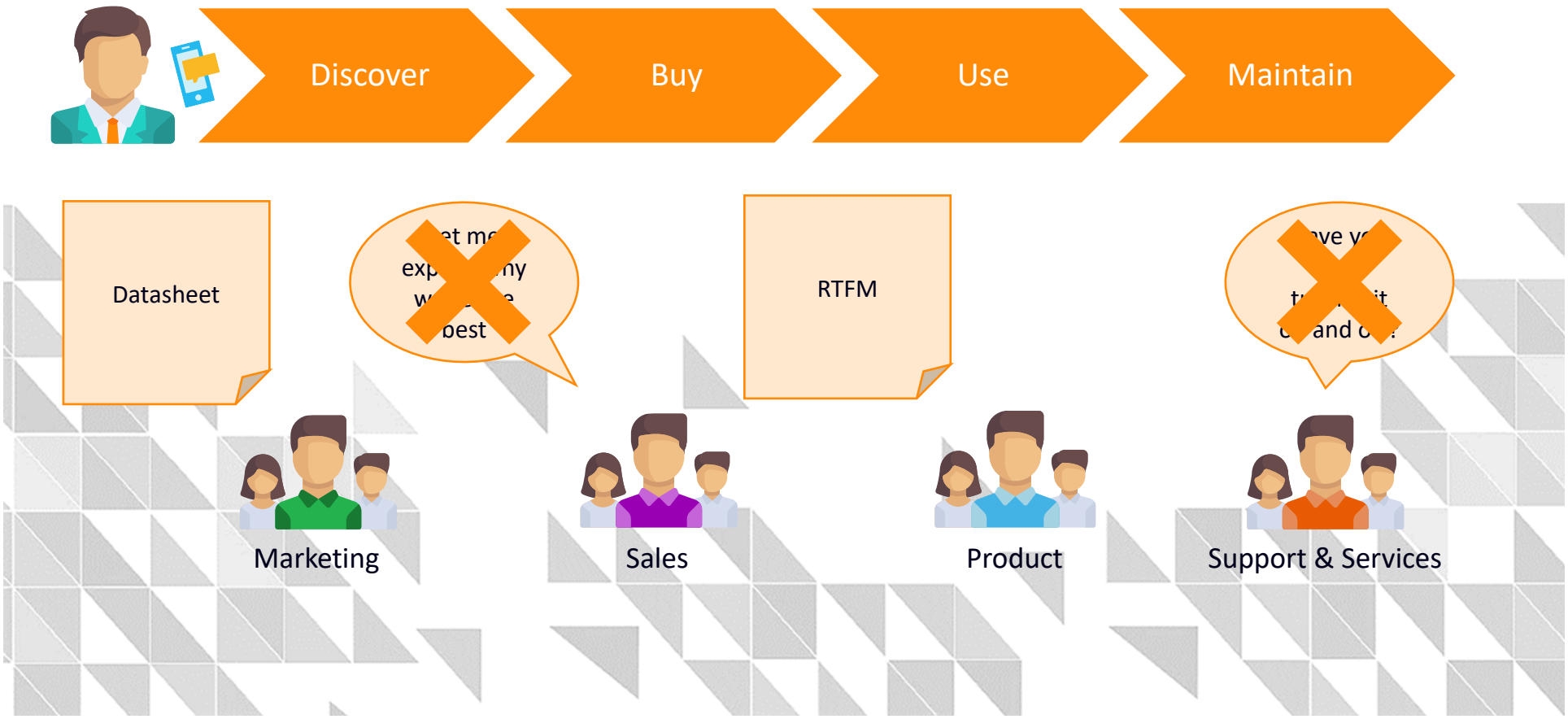
Customer Experience Has Become a Business Imperative

In the old brick-and-mortar days…



Customer Experience Has Become a Business Imperative

In our current mobile-based present

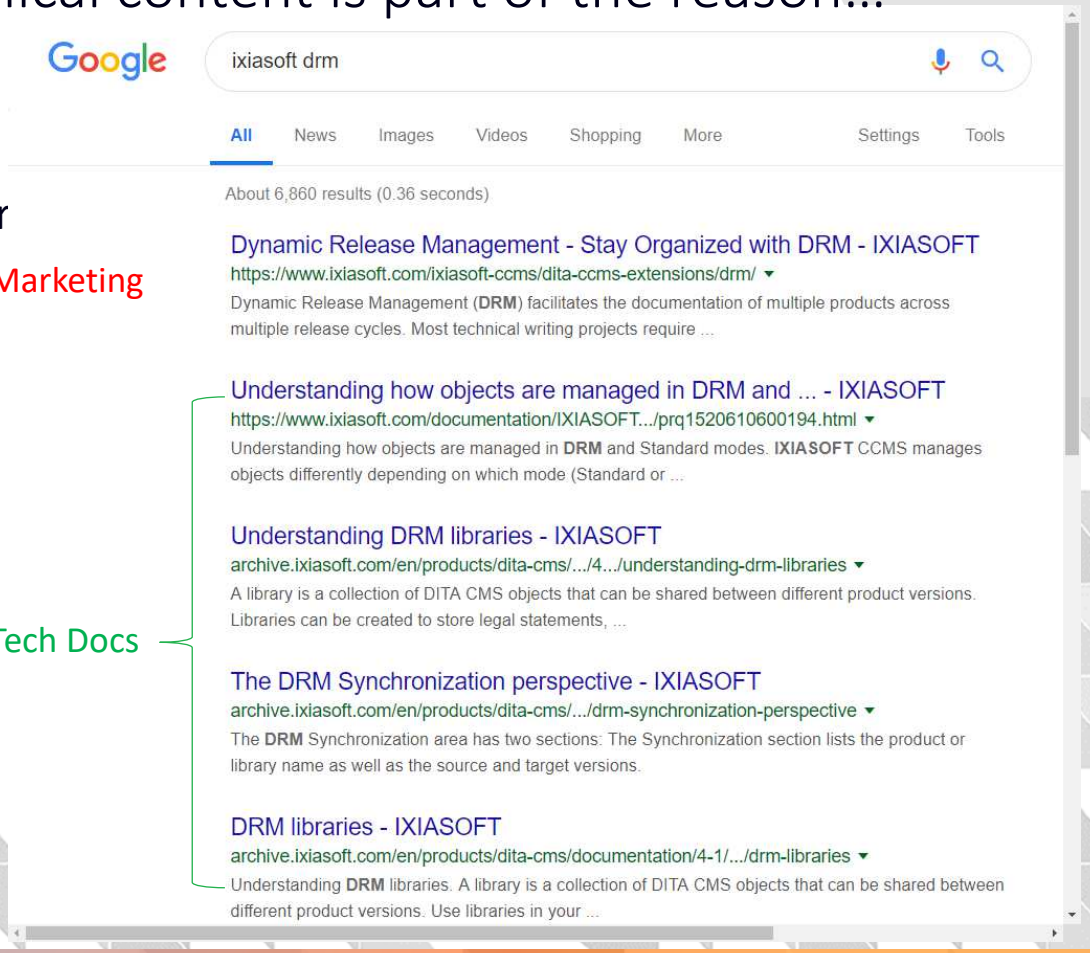


The sheer volume of technical content is part of the reason…

- For example, the IXIASOFT website is comprised of two major portions: marketing content and technical documentation

- Marketing content accounts = ~100 pages
- Webhelp output for technical content = 3,395 pages (for one product release)

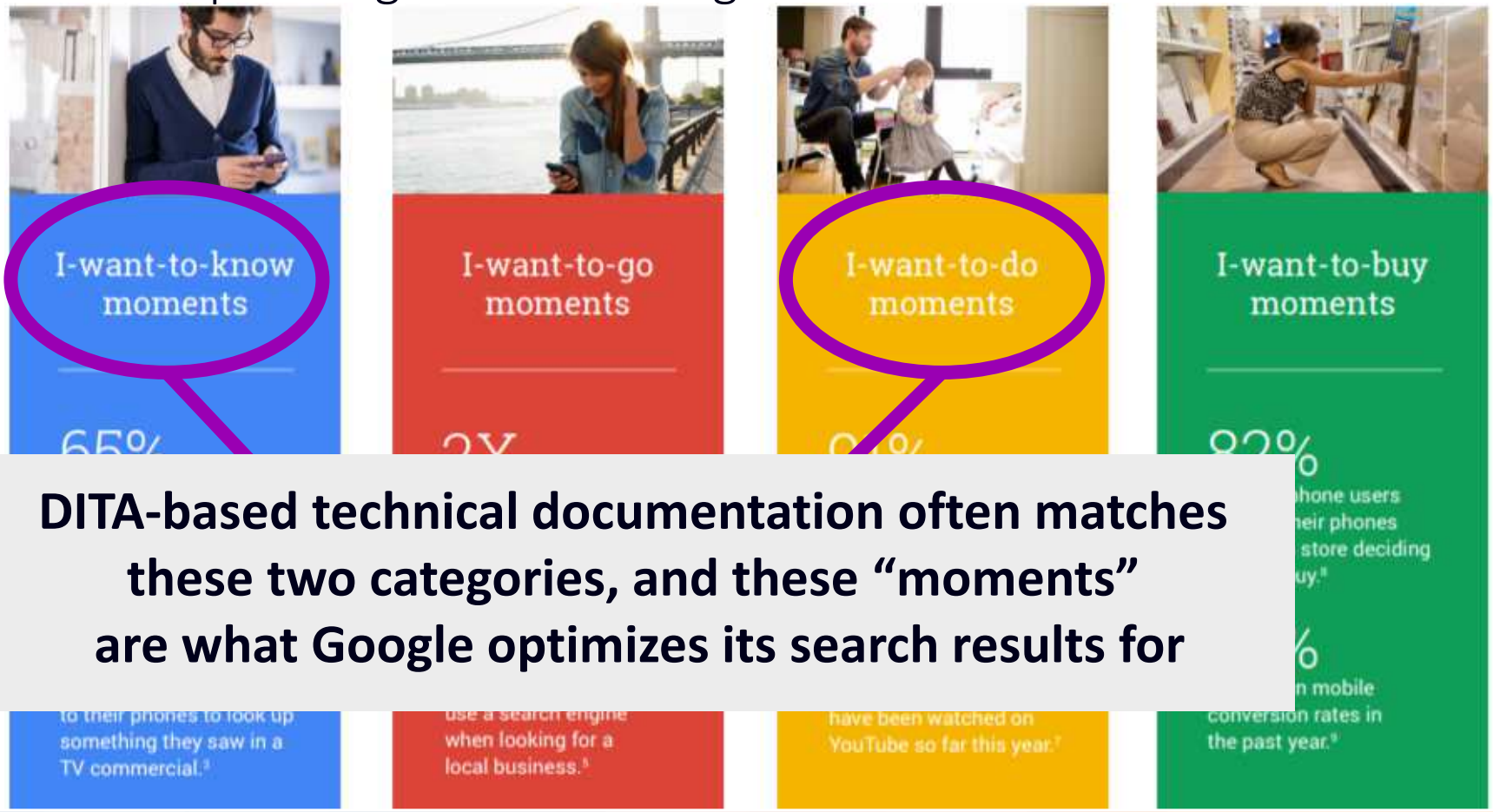
- Given this, is it any wonder that Web searches tend to bring up a lot of technical content?



Marketing

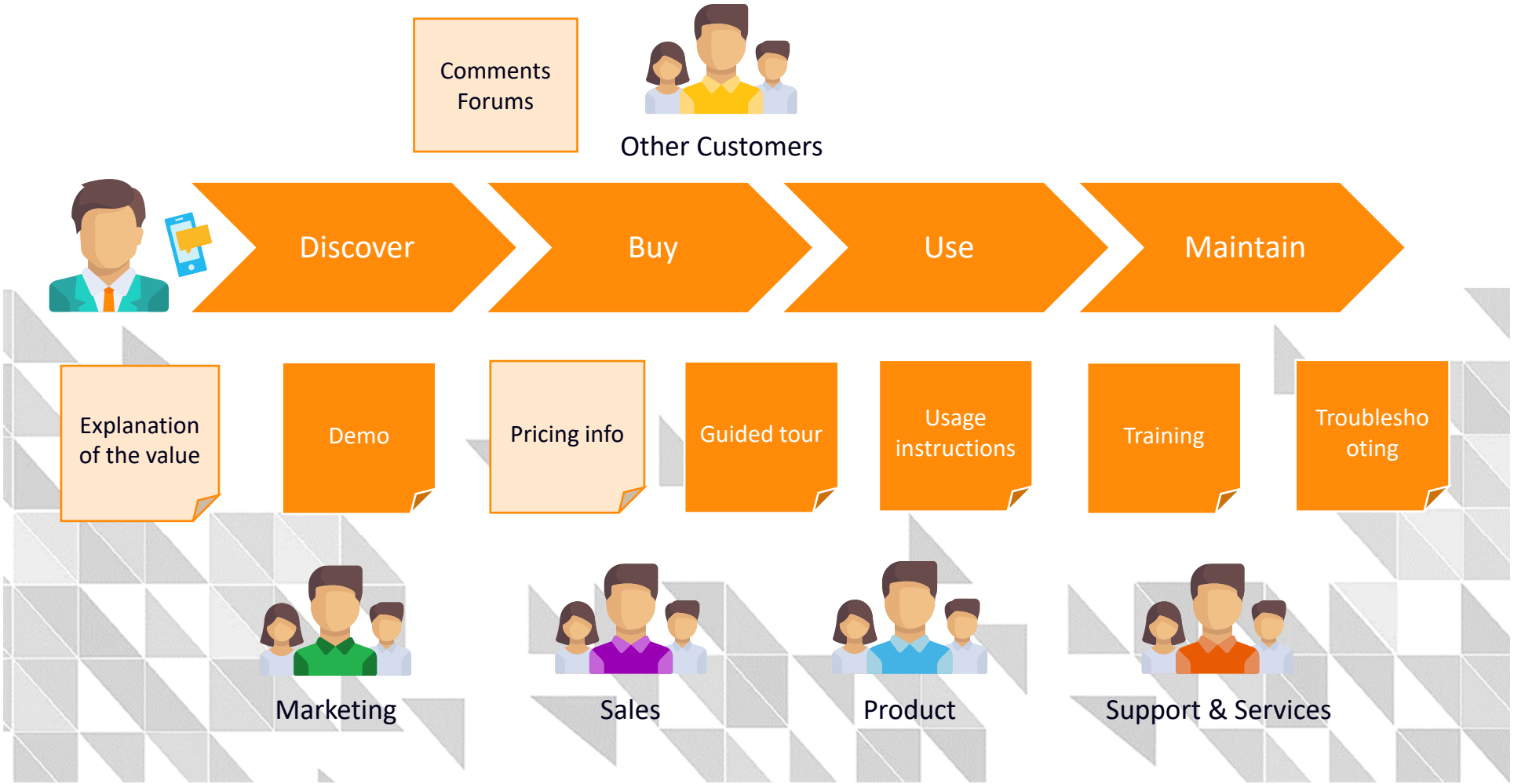
Tech Docs

DITA Topics Target Two of Google's Four User "Moments"



DITA-based technical documentation often matches these two categories, and these "moments" are what Google optimizes its search results for

Technical Content Is a Key Component of Customer Experience



Enter: Chatbots and Artificial Intelligence

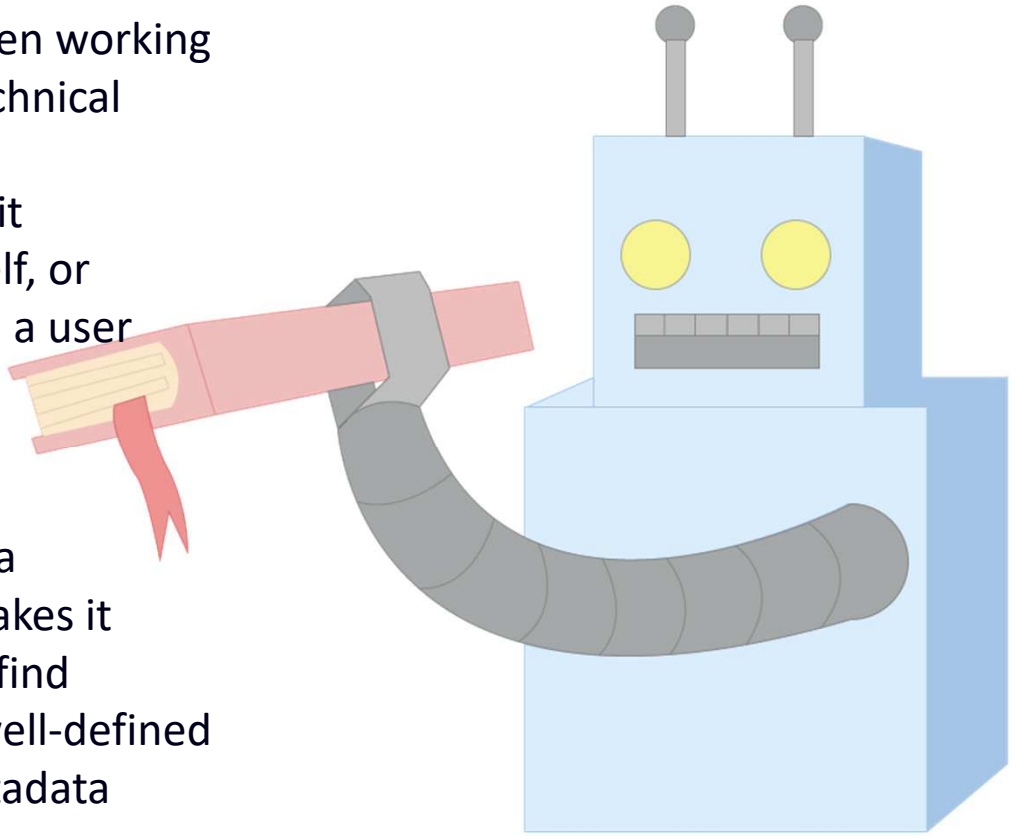
- This subject has been getting a lot of interest in marketing and technical documentation
- We are still in the early days, but it is likely to become another facet in the overall customer experience journey

How Does DITA-based Technical Documentation Fit In?

Chatbots take one of two paths when working with external content (including technical content):

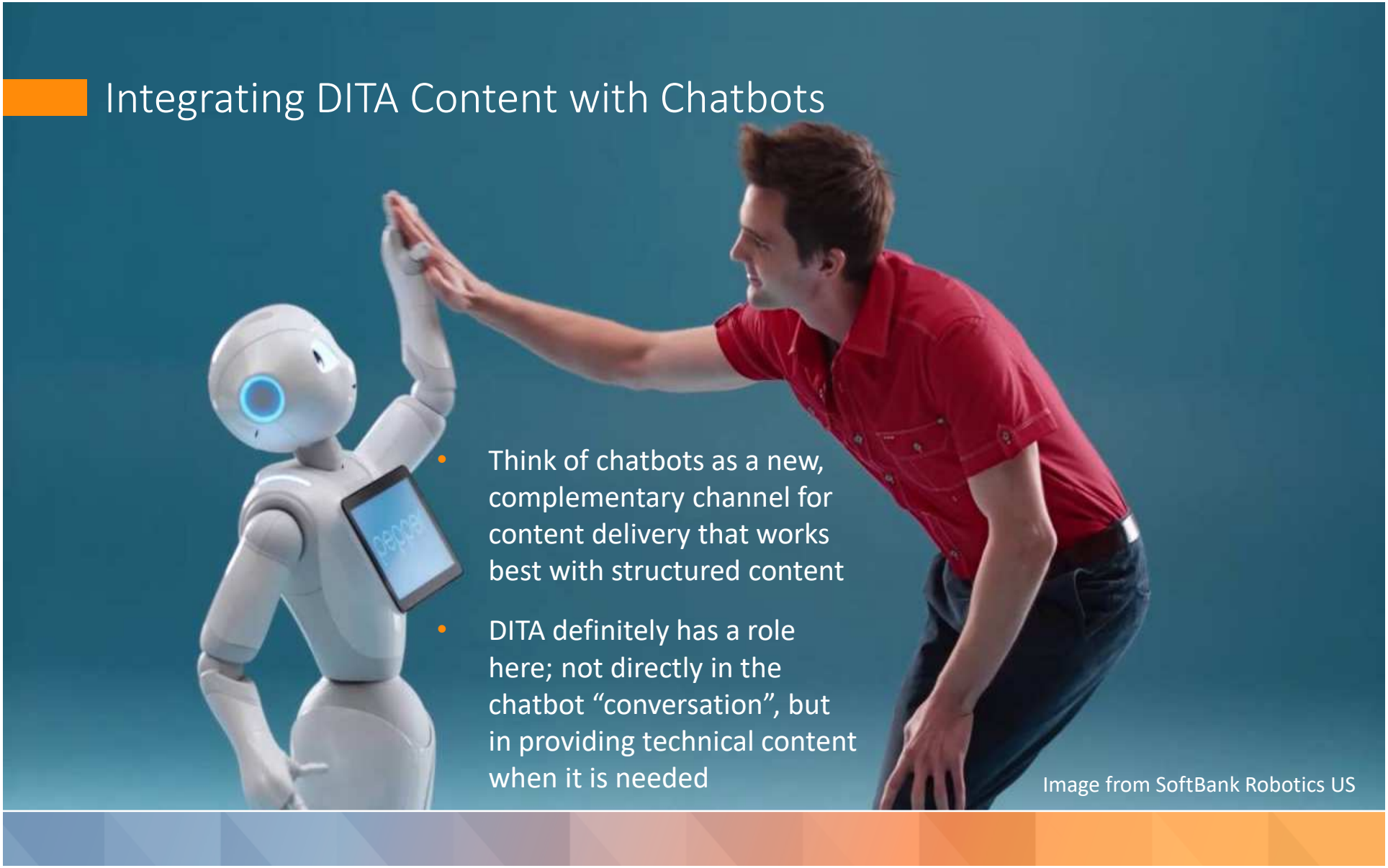
1. Digests existing material which it incorporates directly within itself, or
2. It refers to appropriate material a user wants via metadata

- Either way, having tech docs in a structured format (like DITA) makes it easier for chatbots to digest or find content to reference because well-defined content types + descriptive metadata make it easier for chatbots to use



Robot graphic courtesy of Vanessa Roberts

Integrating DITA Content with Chatbots



- Think of chatbots as a new, complementary channel for content delivery that works best with structured content
- DITA definitely has a role here; not directly in the chatbot “conversation”, but in providing technical content when it is needed

Image from SoftBank Robotics US

In Summary

- DITA 2.0 shaping up to become an evolutionary improvement over DITA 1.3
- Lightweight DITA shows that DITA no longer needs to be tied exclusively to XML
  - Markdown DITA (MDITA) is seeing the most traction from users
- HTML5-like multimedia elements to come with DITA 2.0 + LwDITA
- “Technical Writer” job type is in the middle of sea change
- Focus will be more on enhancing customer experience
- New technologies like chatbots can leverage structured content types like DITA


Questions?
Keith Schengili-Roberts
Senior DITA Content Strategist
keith.schengili-roberts@enbridge.com