Auditing for Reuse Opportunities

By Katherine (Kit) Brown-Hoekstra

comgenesis, llc
communication for a connected world
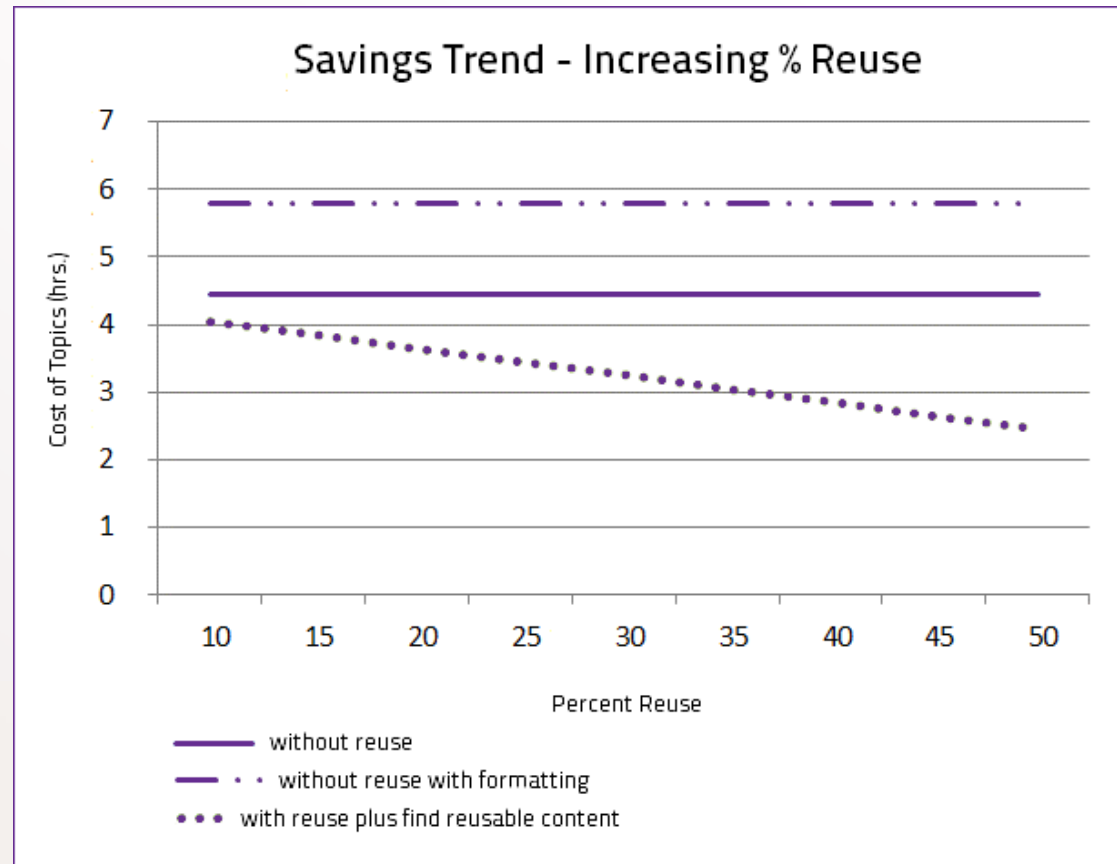
Agenda

- Business Case for Reuse
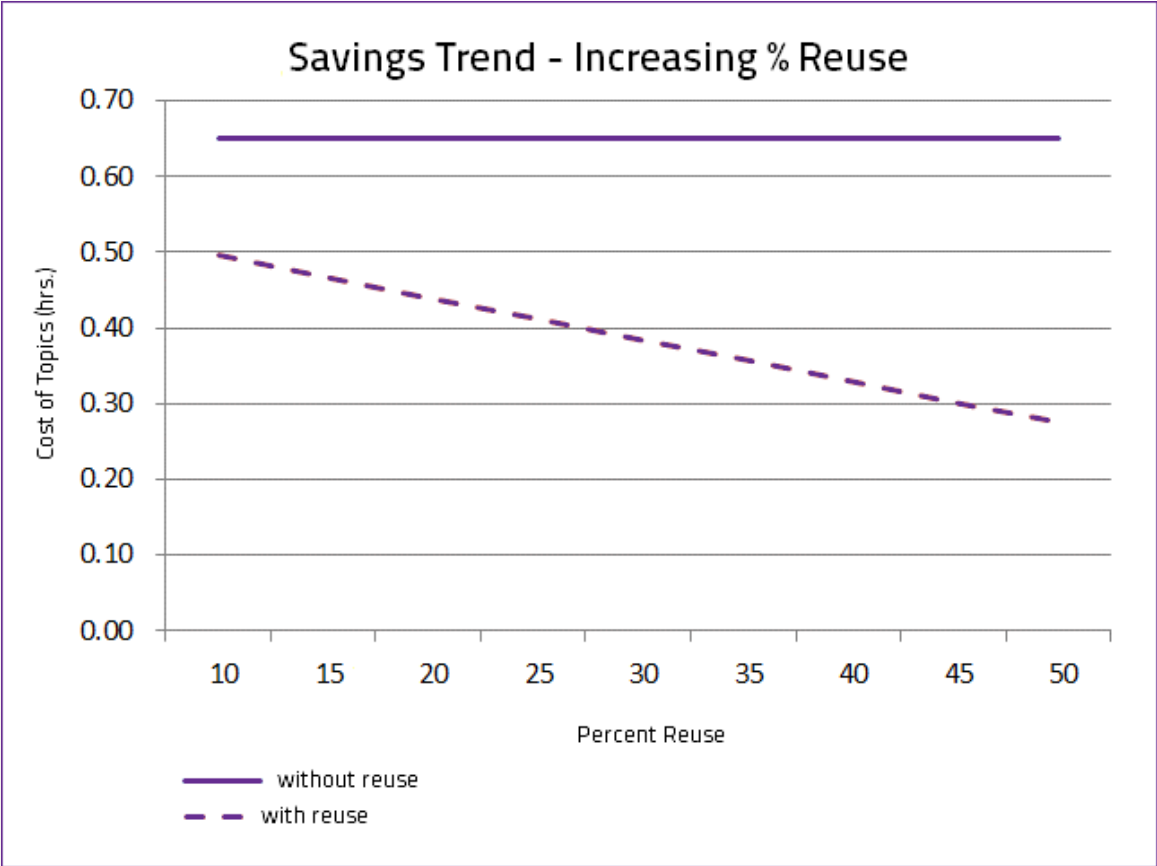- Inventories, Audits, & Analyses
- Pitfalls
- Best Practices

Business Case for Reuse

Cost Savings: Authoring

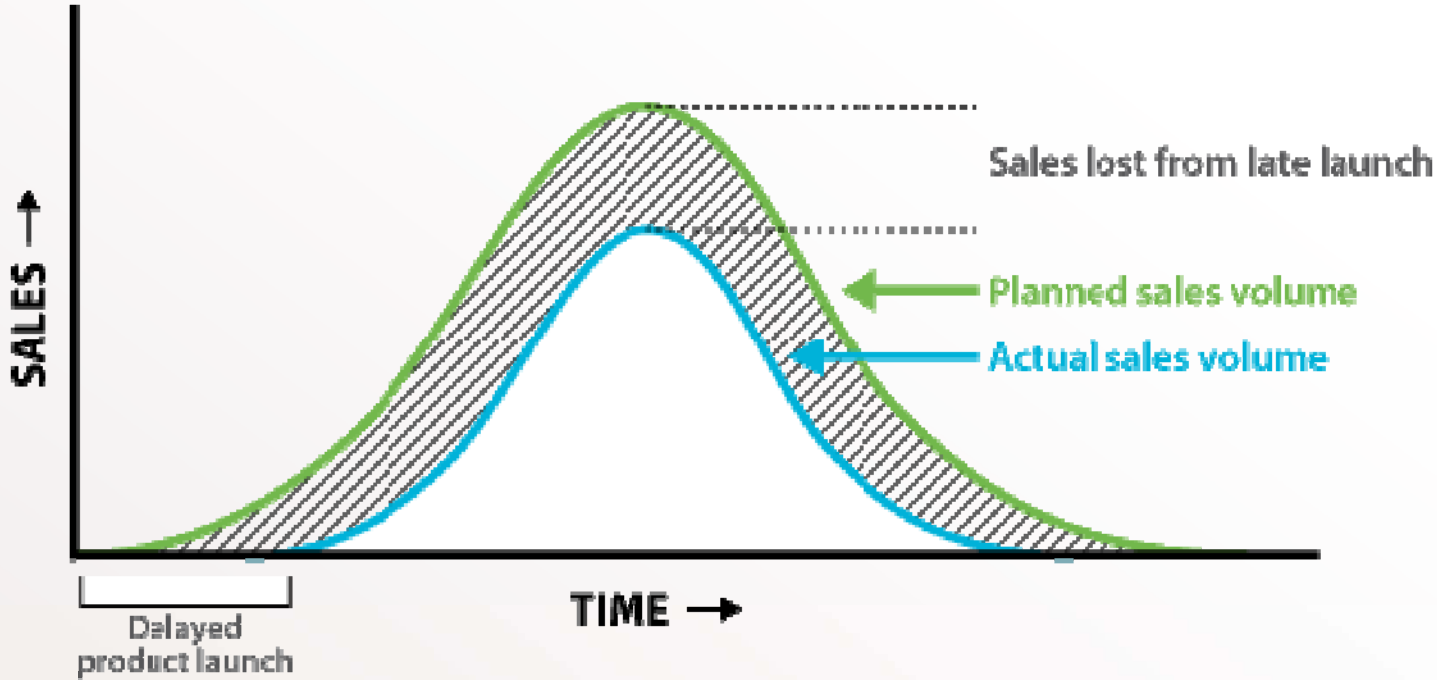
Savings Trend - Increasing % Reuse

Cost of Topics (hrs.) vs Percent Reuse

Legend:
- without reuse
- without reuse with formatting
- with reuse plus find reusable content

From *DITA Metrics 101* by Mark Lewis, pg 78

Cost Savings: Translation
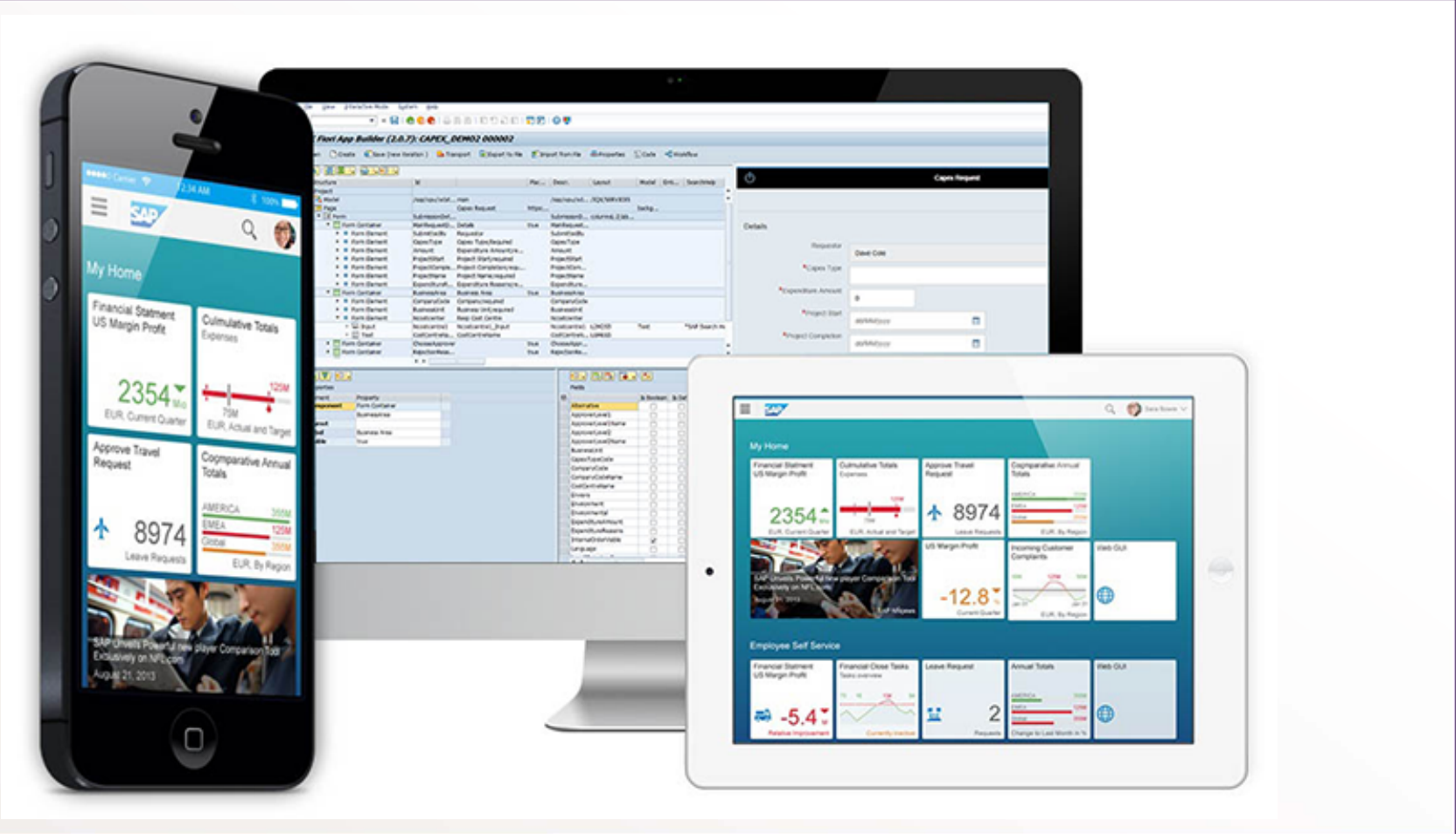

Savings Trend - Increasing % Reuse

From *DITA Metrics 101* by Mark Lewis, pg 101

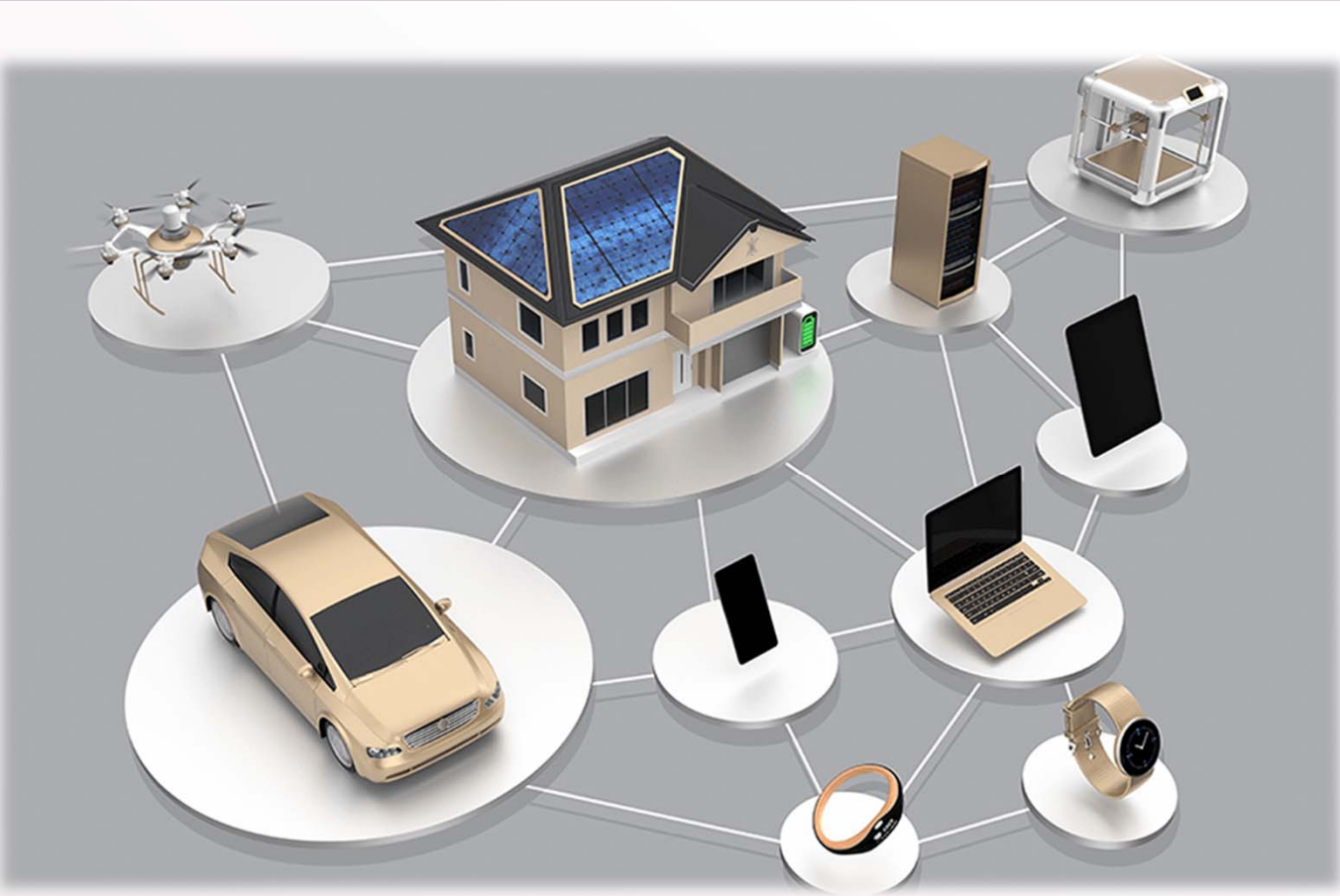# Improved Time to Market

Consistency of User Experience

Improved Automation Opportunities

Credit: chesky/fotolia.com

Decreased Liability

Inventories, Audits, and Analyses

Content Audit Defined

A cross-sectional review of your content to identify quality levels, pain points, internationalization, and so on. Typically, an audit reviews a representative percentage of the content to identify major issues.

Content Inventory Defined

List of all the content in a particular domain, typically including where used information. Most CMSes have tools that can download this information to a spreadsheet.

Content Audit: Purpose

- Identify major issues and pain points before migrating
- Quantify quality so that you can benchmark
- Harmonize content
- Identify training needs
- Reduce localization costs

Content

- Does the doc serve the needs of your user?

- What types of doc do you produce?

- What are the information types?

- Where is the opportunity for reuse?

- Do the templates, style guides, glossaries, etc. support localization needs?

- Is the content well-internationalized?

- What are the top 5 editing issues? Change management? DTP?

- Do graphics contain text in the callouts?

- How many screenshots are used?

- What metadata do you use? How is this tracked/controlled?

Methodology: Audits for Reuse

1. Inventory your content.
   a) Where used
   b) Context
   c) Instances

2. Pull a report on content similarity across ecosystem.
   a) Fuzzy matches of 95% or greater
   b) Database queries
   c) CMS report for duplicates or fuzzy matches

3. Review style guide for reuse issues.

4. Develop a reuse strategy.

5. Model reuse (manual and automated).

6. Prioritize, curate, and clean.

Look for Low-Hanging Fruit

Measurement

- Frequency
- Severity
- Risk
- Impact
- Priority

Forms: Content Audit Example

Forms: Issues/Status Example

Impact: 1= High to 4 = Low									
Cost: 1= less than $500, 2= $501-2000, 3= $2001-4999,4= More than $5000									
Ease of Fixing: 1= Easy to 4=Hard									

Issue	Impact	Cost	Ease of Fixing	Date Logged	Auditor	Comments	Ideas	Action Item Assigned to	Date Resolved
Sample Issue--provide a brief (less than 10 words) name for issue: time zones in tag line	1	0	1	5/16/2009	Kit	(provide info that describes the issue in more detail) As a global organization, knowing what time zones the people we are communicating with are in makes planning meetings and so on easier; it also raises awareness among the membership considering non-local needs	(If possible, provide ideas for solving; these will be discussed, and if agreed, included in the formal recommendations) Include information In News & Notes; Post request to all SIG lists; ensure that Board and Staff have added the information; review website to ensure that time zone information is included in webinars and other meeting announcements, etc	Kit-News and Notes and SIG lists; Fatima/Diana-- Staff; Leah-- Board;	

Content Audit Spreadsheet Example 2

Observation Type	Observation Description	Example Location (URL, Page#, topic)	Doc Type	Priority	Ease of Implementation	Impact	Estimated Cost ($-$$$$)	Recommendation
Re-use	Duplicate content for warning related to risk of injury		Online Help	High	Moderate	Medium	$	Curate all instances, choose one to use throughout, automate, clean TM

Audit Spreadsheet Example 1

<Project Name>

SECTION	CRITERIA1	#	%	CRITERIA 2	#	%	CRITERIA 3	#	%	CRITERIA 4	#	%	CRITERIA 5	#	%	CRITERIA 6	#	%
<Section Name>	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%
	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%
	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%
<Section Name>	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%
	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%
	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%
<Section Name>	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%
	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%
	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%
<Section Name>	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%
	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%
	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%
<Section Name>	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%
	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%
	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%
<Section Name>	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%	1 - NO	1	33%
	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%	2 - SOMEWHAT	1	33%
	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%	3 - YES	1	33%
OVERALL	1 - NO	6	33%	1 - NO	6	33%	1 - NO	6	33%	1 - NO	6	33%	1 - NO	6	33%	1 - NO	6	33%
	2 - SOMEWHAT	6	33%	2 - SOMEWHAT	6	33%	2 - SOMEWHAT	6	33%	2 - SOMEWHAT	6	33%	2 - SOMEWHAT	6	33%	2 - SOMEWHAT	6	33%
	3 - YES	6	33%	3 - YES	6	33%	3 - YES	6	33%	3 - YES	6	33%	3 - YES	6	33%	3 - YES	6	33%

From *The Content Strategy Toolkit* by Meghan Casey

Reporting

- Know your audience
- Stay out of the weeds when talking to management
- Focus on the cost/benefit
- Use examples
- Check level of formality (e.g., slide deck vs technical report or executive summary)
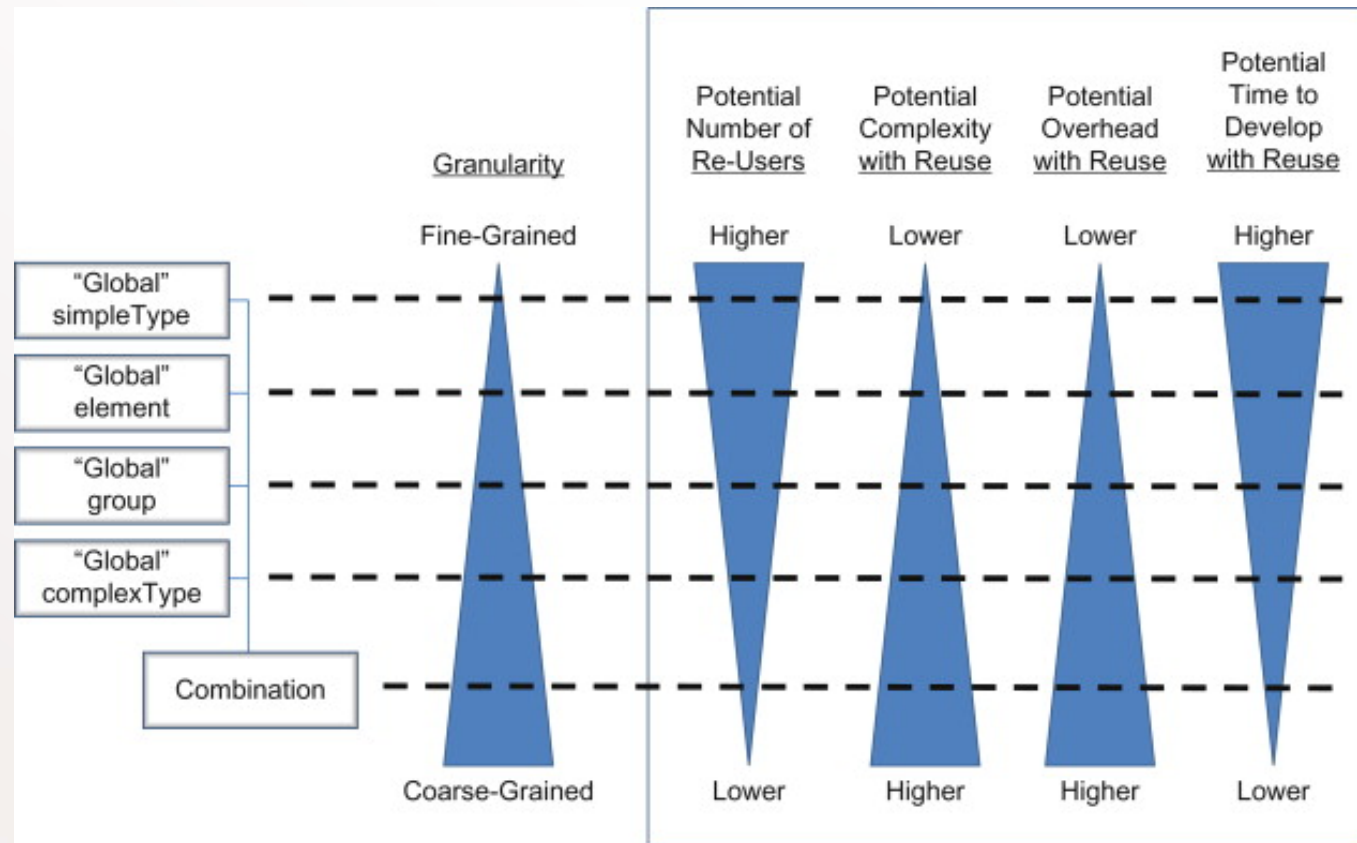- Provide a clear path forward

Pitfalls with Reuse

Wrong Context

Wrong Level of Granularity

From *Relational Database Design and Implementation (Fourth Edition)* by Jan L. Harrington

Wrong Tone/Style

Highly Technical:

"Although monogenic forms of blood pressure dysregulation exist, hypertension mostly arises as a complex quantitative trait that is affected by varying combinations of genetic and environmental factors. Non-pharmacological strategies can reduce blood pressure."

https://www.sciencedirect.com/science/article/abs/pii/S0140673603133028

For Layperson:

Although problems with blood pressure can sometimes be traced to a single gene, most blood pressure problems arise from a combination of multiple genetic and environmental factors. Sometimes, you can control blood pressure without medication.
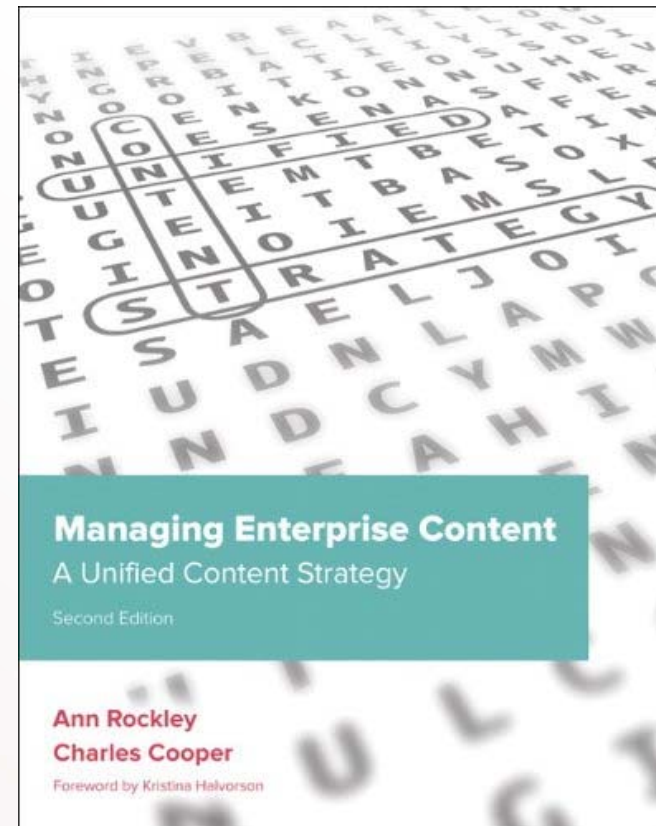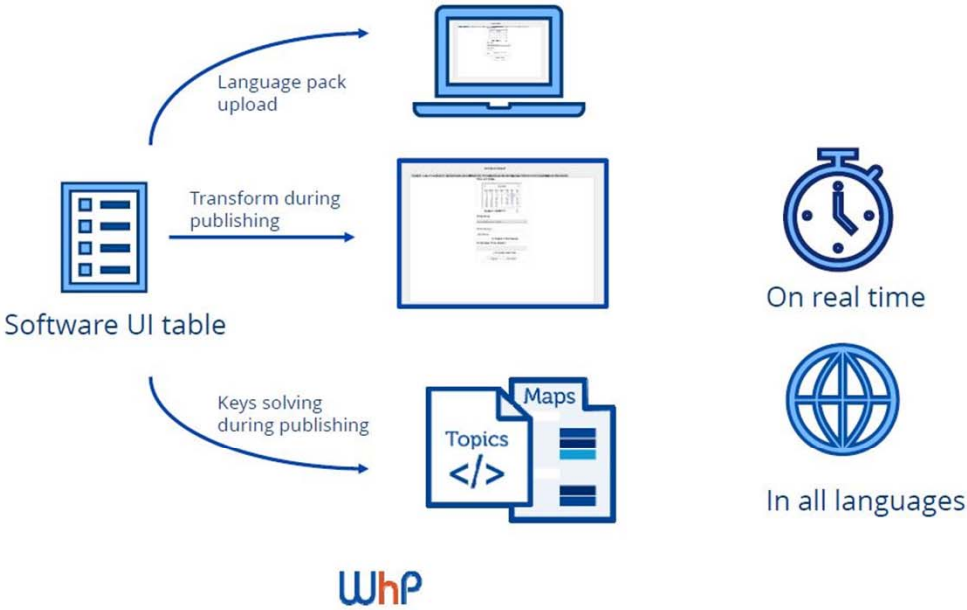
Best Practices for Reuse

Define a Reuse Strategy

Define and Model

- Ways content will be used and reused

- Types of reuse

- Methods for securing content
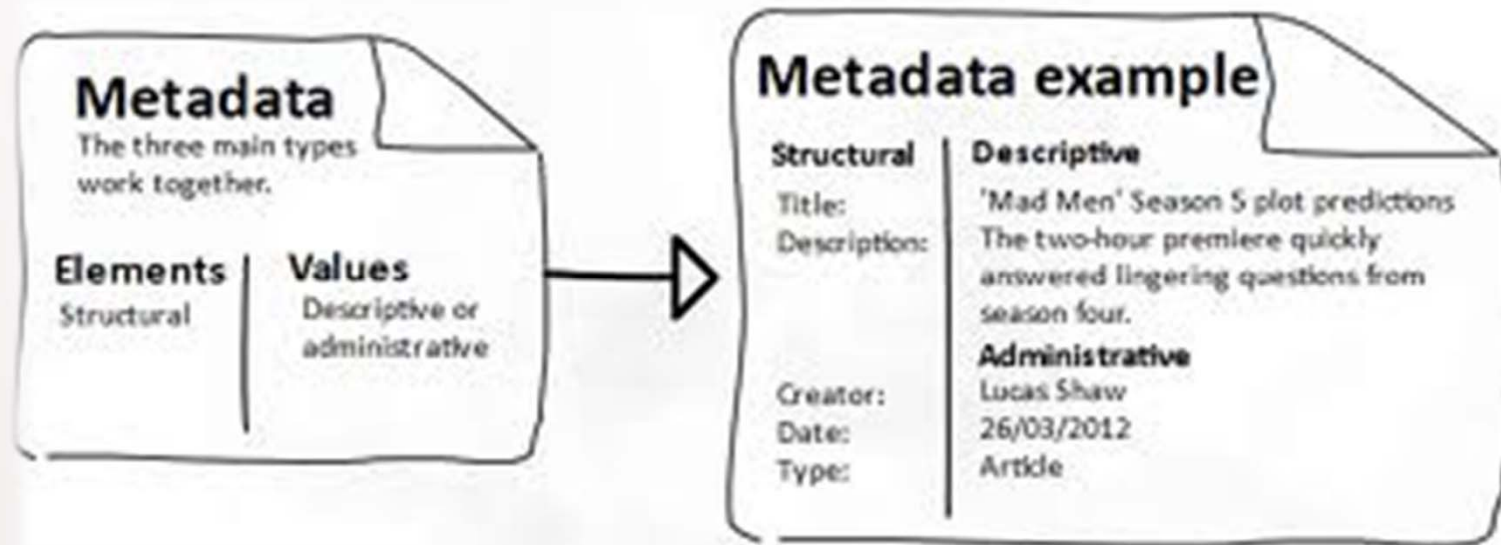
- Level of granularity

- Governance strategy

Managing Enterprise Content
A Unified Content Strategy

Second Edition

Ann Rockley
Charles Cooper

Foreword by Kristina Halvorson

Pg 163

Externalize Your Frequently Used Content
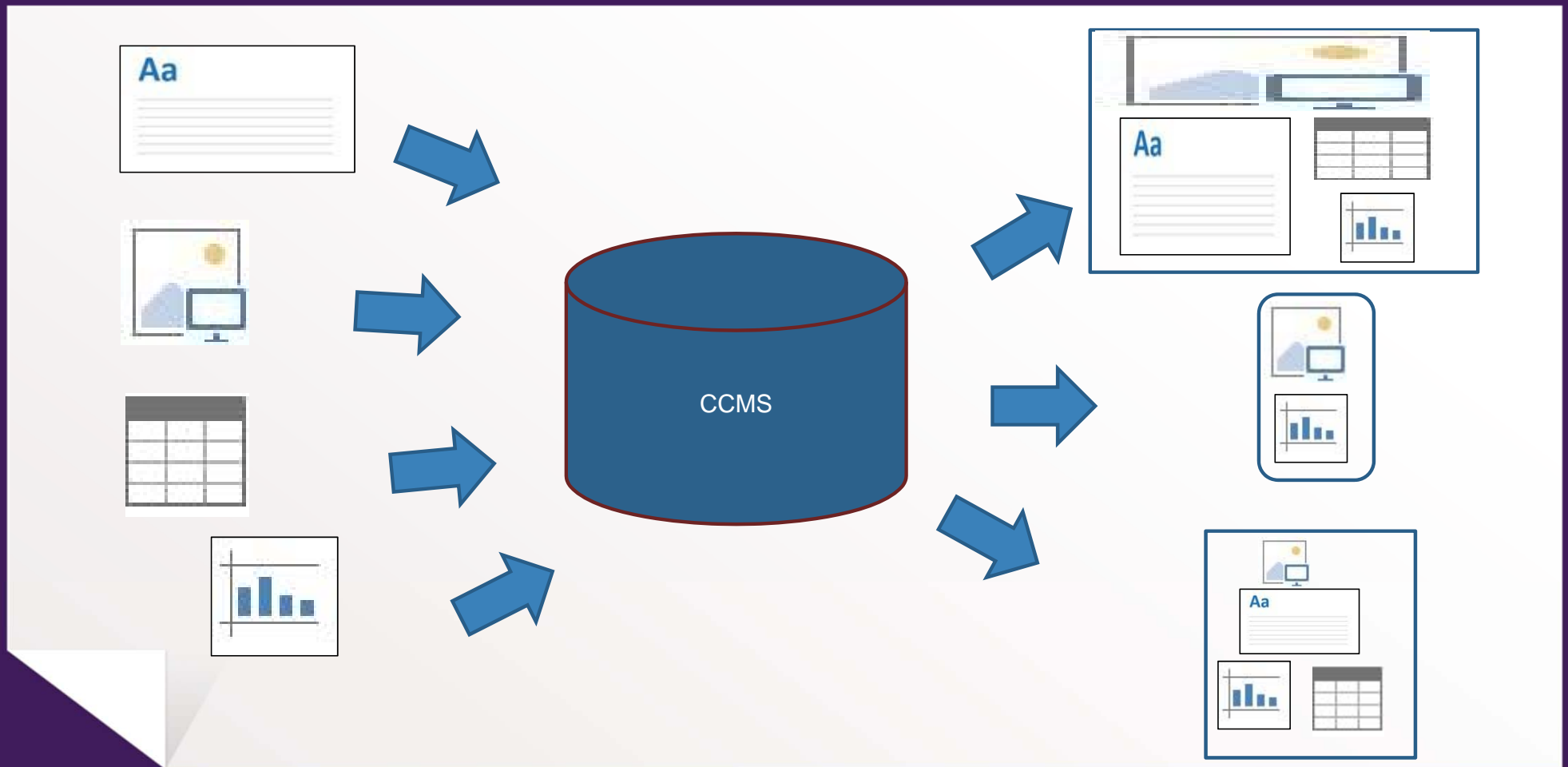
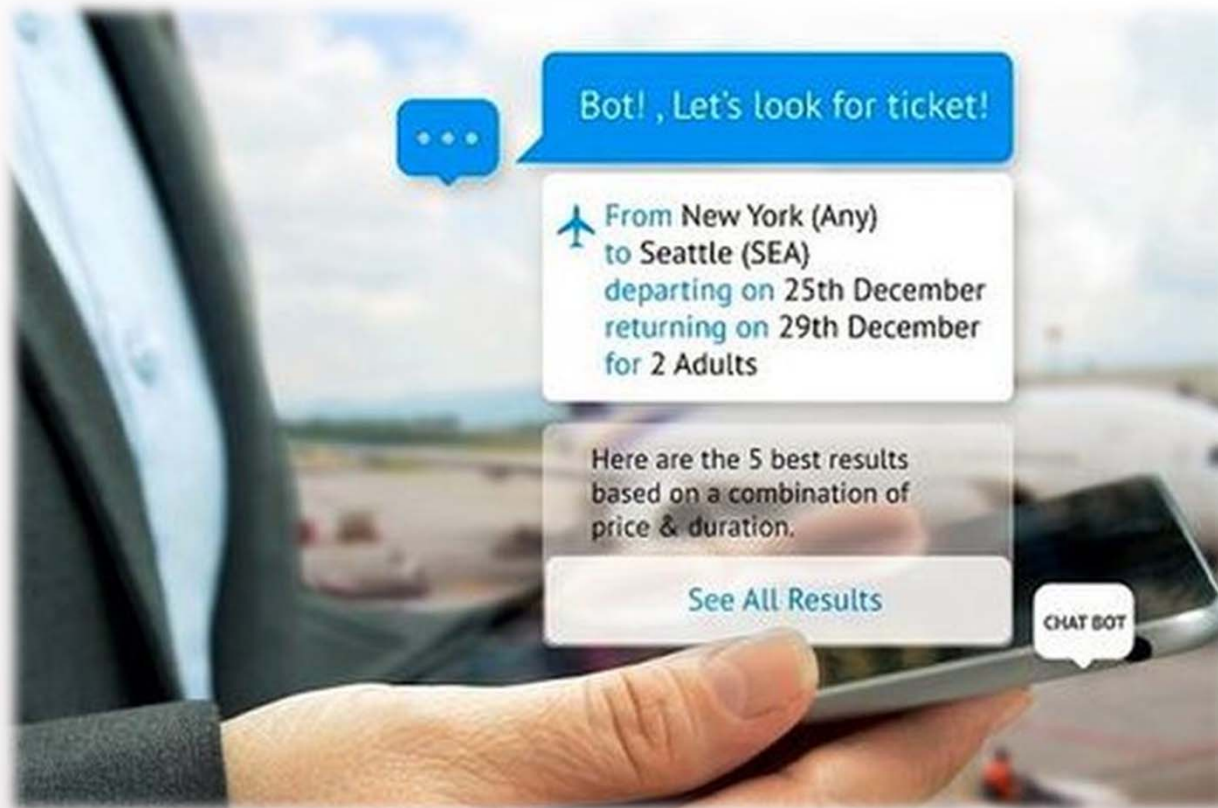# Use Effective Metadata

Use a CCMS

DITA-Specific

- Keyrefs
- Conkeyref
- Conrefs
- Shortdesc
- Conditionalizing

Automate

Thanks!

Any questions?

You can find me at:

@kitcomgenesis

Kit.brown@comgenesis.com

Resources

Resources

- Baker, Mark. *Every Page is Page One.*, XML Press. 2013.
- Casey, Meghan. *The Content Strategy Toolkit.* New Riders Press, 2015.
- Harrington, J. [Relational Database Design and Implementation (Fourth Edition)](#). Science Direct. 2016
- Land, Paula. *Content Audits and Inventories.* XML Press, 2014.
- Lewis, M. *DITA Metrics 101.* The Rockley Group, Inc., 2012.
- Porter, Alan. *The Content Pool.* XML Press. 2012.
- Rockley, A and Cooper, C. *Managing Enterprise Content, 2nd Ed.* New Riders Press, 2012.