A PRACTICAL GUIDE TO
CONTENT
MIGRATION
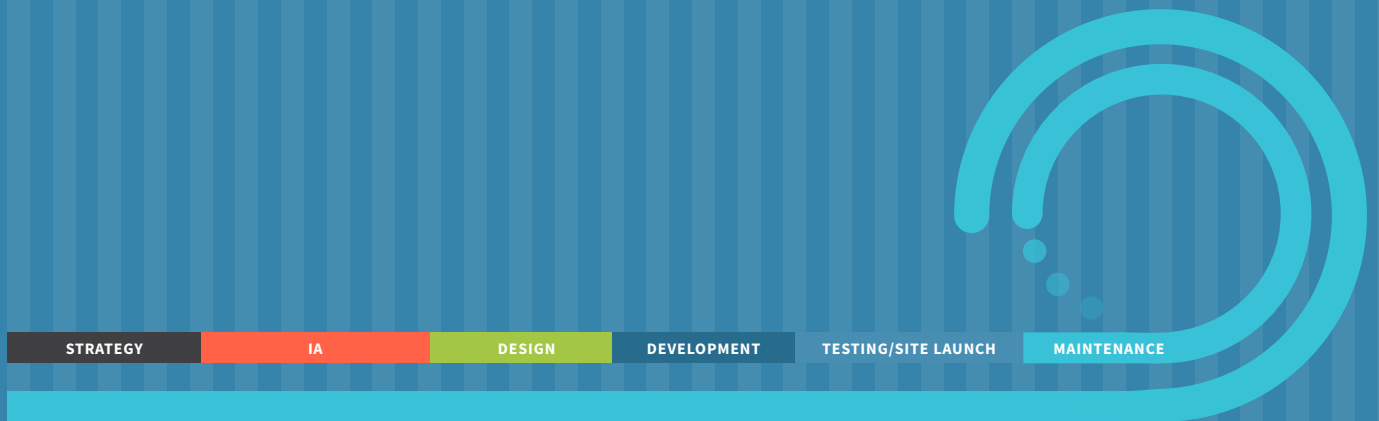
Contents

Introduction

CONTENT SEEMS TO HAVE THE UNIQUE ABILITY TO SABOTAGE DIGITAL PROJECTS LIKE NOTHING ELSE AND SEND TIMELINES COMPLETELY OFF THE RAILS. THE ONLY SOLUTION IS TO PLAN AHEAD AND TAKE A PRACTICAL APPROACH. WHETHER WE'RE CREATING THE CONTENT WITH OUR CLIENTS OR HELPING THEM TO CREATE IT ON THEIR OWN, IT'S IMPORTANT TO USE TRAINED WRITERS, TO WRITE CONTENT WITH REAL HUMANS IN MIND, AND TO DESIGN CONTENT FOR READING ON THE WEB. EACH PAGE OF WEBSITE CONTENT SHOULD BE APPROACHED WITH A CLEAR SET OF GOALS. AND YET, EVEN KNOWING ALL THESE THINGS, WE FACE DOZENS OF COMPLEXITIES WHEN IT'S TIME TO REPURPOSE CONTENT FROM THE RUINS OF AN OLD WEBSITE OR A SHINY NEW SET OF EMPTY WIREFRAMES. THROUGH HUNDREDS OF DIGITAL PROJECTS, WE'VE DEVELOPED SOME BEST PRACTICES TO GUIDE YOU THROUGH A SMOOTH AND STREAMLINED CONTENT MIGRATION PROCESS.

## The content migration process spans the entire project lifecycle.

STRATEGY	IA	DESIGN	DEVELOPMENT	TESTING/SITE LAUNCH	MAINTENANCE

CONTENT PLANNING & MIGRATION

1. Evaluate Your Content

IF IT AIN'T BROKE, DON'T FIX IT, RIGHT?

Before you decide how you're going to fix your content, first find out what's wrong with it. Never assume that you know what you need to do better; use investigative and deep diving tools to get a full handle on what the current state of your content looks like.

| STRATEGY | IA | DESIGN | DEVELOPMENT | TESTING/SITE LAUNCH | MAINTENANCE |
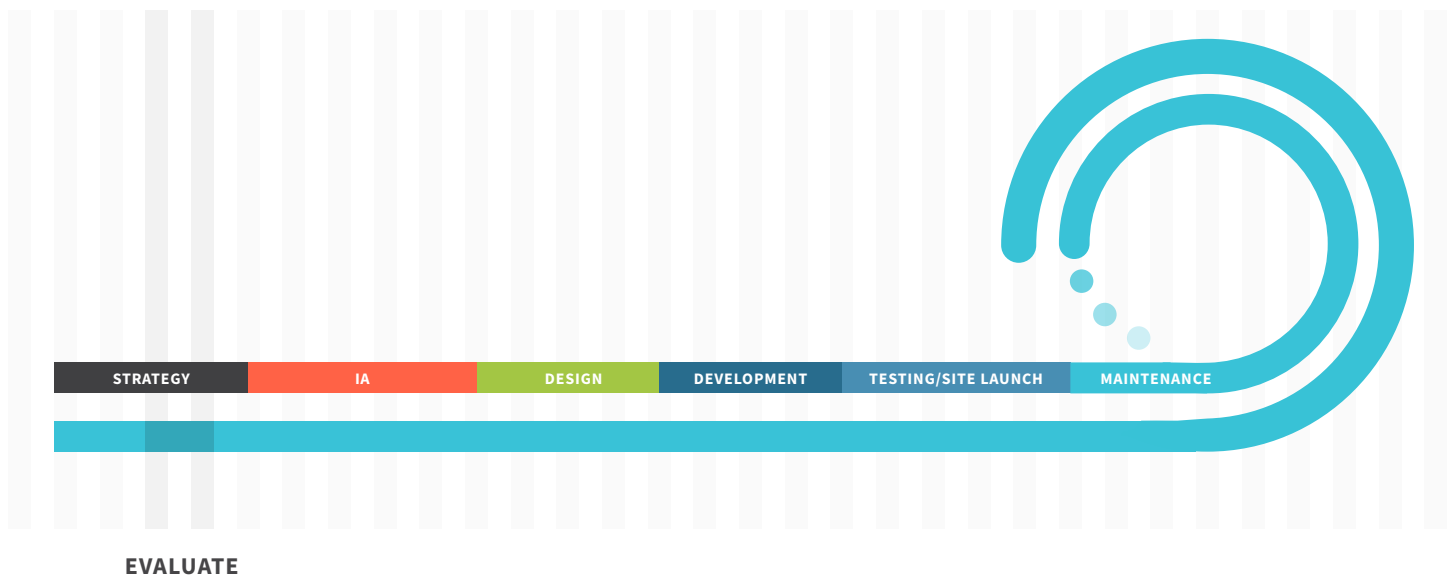
EVALUATE

Evaluate content early-on as part of the overall strategy and planning phase.

Here are a few tools to go about surveying your site's content:

A. CONTENT INVENTORY

Chances are you have a lot more content than you realize. Include existing web, mobile, and print content in your inventory to establish a foundation from which a new site can be built. A thorough inventory can provide an understanding of the breadth of current content, including:

- A breakdown of content formats
- A high-level view of improvements needed to optimize the user experience
- A vocabulary to inform the site taxonomy

B. CONTENT HEURISTIC ASSESSMENT

We use eight heuristics, or rules-of-thumb based on experience, to evaluate existing site content, including:

- Collocation: Items with similar content or items about the same topic are located in one area.

- Differentiation: Dissimilar items or items about different subject areas are placed in different content areas. Use navigation labels for different areas that clearly indicate those differences.

- Completeness: All content mentioned or linked to should exist and be available.

- Information scent: Content labels should be appropriately descriptive so that users know they are on the proper path to finding the information they are looking for. Content labels should therefore also reflect information collocation and differentiation.

- Accessibility: Users should be able to access the content they want through the browsing hierarchy or by using search.

- Multiple access paths: Because users think about content in different ways, they should be able to take multiple paths to get to specific content.

- Consistency: Whenever possible, content structures in similar content areas should be consistent.

- Currency of content: Content should be kept up to date and frequency of updates should be noted somewhere on the website.

C. ROT ANALYSIS

ROT stands for Redundant, Outdated and Trivial. When you're unsure about whether or not content you find is important or noteworthy, this is a quick and easy way to determine if it needs to stay. During analysis for ROT , the goal is to determine what to do with Redundant, Outdated or Trivial content – delete, re-write, or keep as is.

Cleaning up ROT content will result in removing duplicates, archiving out-of-date content, re-writing misleading content, or making it relevant for the user and the organization. It can also help you to fix broken links and avoid having your users reach dead ends on your site.

SAMPLE CONTENT INVENTORY

ID	PAGE OR CONTENT NAME	NOTES	OWNER	EVALUATION	STATUS	LINK
1	About					www.company.com/about
1.1	Intro Text		Marketing	Update Content		
1.2	Our Vision		Marketing	Retain Current Content	Done	
1.3	Our Values		Marketing	Purge Content		
1.4	Leadership Bios		HR	Update Content		
1.5	Overview Briefing	Broken Link				
2	Contact Us					www.company.com/contact-us
2.1	Intro Text	Broken Link	Marketing	Update Content		
2.2	Contact Form		Marketing	Retain Current Content	Done	

"R" is for Redundant

Redundant information also makes your site bigger than it needs to be. User journeys become confusing and long, or reliant on search tools. Content can be redundant when:

• There is no longer a need for it.

• The content is duplicated elsewhere.

"O" is for Outdated

Outdated content can affect credibility and result in confusion and error. Outdated content includes:

• News, events, or instructions that are represented as new when they're not.

• Incorrect contact information.

• Services, benefits, or guidance that are no longer accurate.

"T" is for Trivial

Trivial content doesn't really do or communicate anything. All of your information should be useful. If it's not, users will become disenfranchised and are likely to reduce or discontinue use of your website.

Analytics are a goldmine of insight into user behavior. Look to your analytics to gauge what pages are most popular vs. what pages are overlooked. Here are a few ways to take all the information found in your analytics to find out what your users are doing and how:

• Clickstream Pathways- How are users getting to critical pages? You can find out if users are missing a navigational direction because they are getting to certain pages inadvertently or through more indirect routes.

• Referral Sources- How are users finding you? And where do they land? Make sure you keep or redirect important entry pages.

• Duration- Are users spending a lot of time on one page in particular? Maybe you need to parse out information so it's more digestible, and so that users find what they need more easily.

2. Writing Good Web Content

THE WEB IS NOT PRINT. PEOPLE READ VERY DIFFERENTLY ON THE WEB — THEY OFTEN SCAN THE CONTENT, LOOKING FOR KEY TERMS AND INFORMATION, SO YOU MUST WRITE TO ACCOMMODATE THIS BEHAVIOR.

| STRATEGY | IA | DESIGN | DEVELOPMENT | TESTING/SITE LAUNCH | MAINTENANCE |
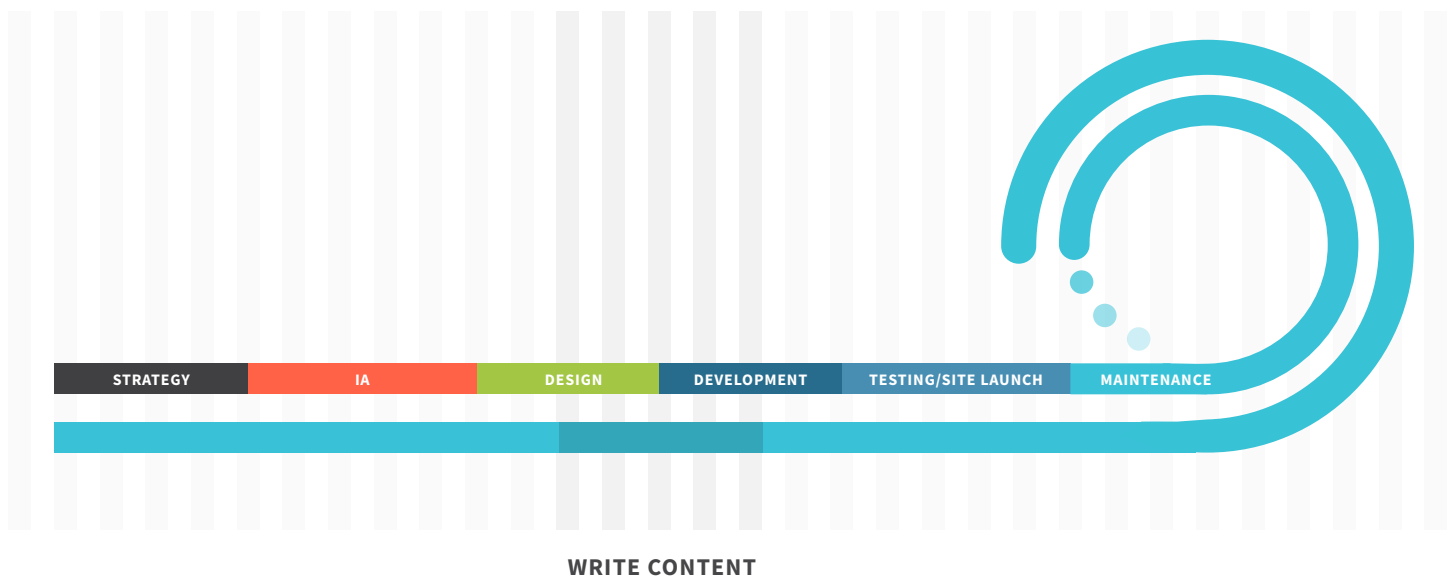
WRITE CONTENT

Design should inform the type of content you will need.
Aim to be finished with content writing before testing begins.

A. WRITE WITH PURPOSE

Make sure each content item has a defined focus that is stated up front. You and your writers should have a clear idea of what you're trying to communicate. Each phrase or paragraph should also have a clear, singular focus. Avoid "marketese" or boilerplate promotional text.

B. USE INVERTED PYRAMID STYLE

1

Open with your most important point first –
don't start with background to build up to a conclusion.

2

Next, provide the most relevant and
compelling supporting information.

3

End with background information
and links to related information
to help create context.

C. KEEP YOUR WRITING CLEAR, INFORMAL, AND EASY TO UNDERSTAND

• Use the active voice.

• Write in a conversational style: Use second person ("you"), contractions and informal construction.

• Show how the product, solution, or information solves a customer's problem or answers a question they have.

• Avoid jargon and abbreviations – write in plain language.

• When you do use an abbreviation, define it on the first use.

• Use examples.

• If there are steps, make them explicit.

• Avoid long sentences. In general, if a sentence contains a dash, separate the sentence into two.

• Cut material that is repetitive or unnecessary. This is more than just a space consideration; it shows that you value readers' time.

• Write for a 9th - 10th grade level, with a Flesch reading ease score of 50 or higher. (This information is available in MS Word – choose File > Options > Proofing > Show readability statistics)

D. FORMAT YOUR COPY FOR EASE OF SCANNING

On the web, people scan content — they don't read. This is especially true with mobile devices. Content usability studies have shown that reducing a web page's word count by half actually doubles the amount of information the reader remembers.

• People read 25 percent more slowly online. Write half as much.

• Use meaningful (not clever) subheadings to help break up long pages.

• Use bulleted or numbered lists to simplify text where appropriate.

• Keep paragraphs short – focus on one idea per paragraph. Users do not respond to heavy blocks of text, or copy that demands excessive scrolling.

• Keep sentences short – ideally, no more than 15 words.

• Use bold text (sparingly) to highlight important words and ideas.

• Think about how visuals such as graphs, charts, photos, infographics, etc., can help readers understand the information.

• If a word or phrase is superfluous (even if it's clever), delete it.

Define the editorial style of written content on your website, including word usage, capitalization, punctuation, terminology, voice, and advice on writing effectively for the web; lay out best practices in quality assurance for scripts used in video and audio content. This documentation will help maintain consistency as your content evolves.

3. Don't Forget Microcopy

MICROCOPY IS THE LABEL, INSTRUCTIONAL TEXT, OR GROUP OF WORDS THAT MAKE YOUR WEBSITE EASIER AND MORE FUN TO USE.
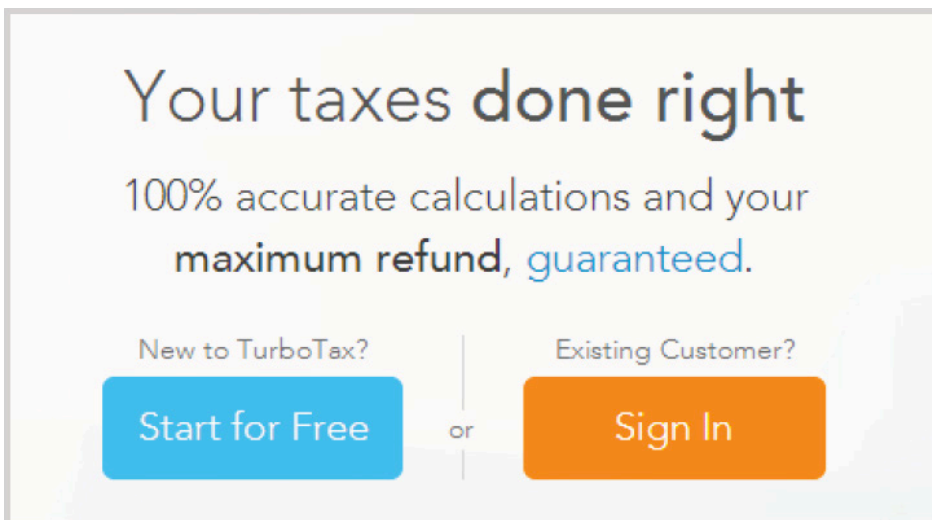

Notice how the text above and on each button leaves no room for confusion.

THIS IS GOOD MICROCOPY.

MICROCOPY IS EVERYWHERE:

• A loading message

• A button label

• The instructions to fill out a form

• The message on a thank you page

• The reassurance that you won't spam under a form

• An explanation of what you will do with the information users submit

• Error messages

Microcopy should be completed before design begins so it can be incorporated into the designs.
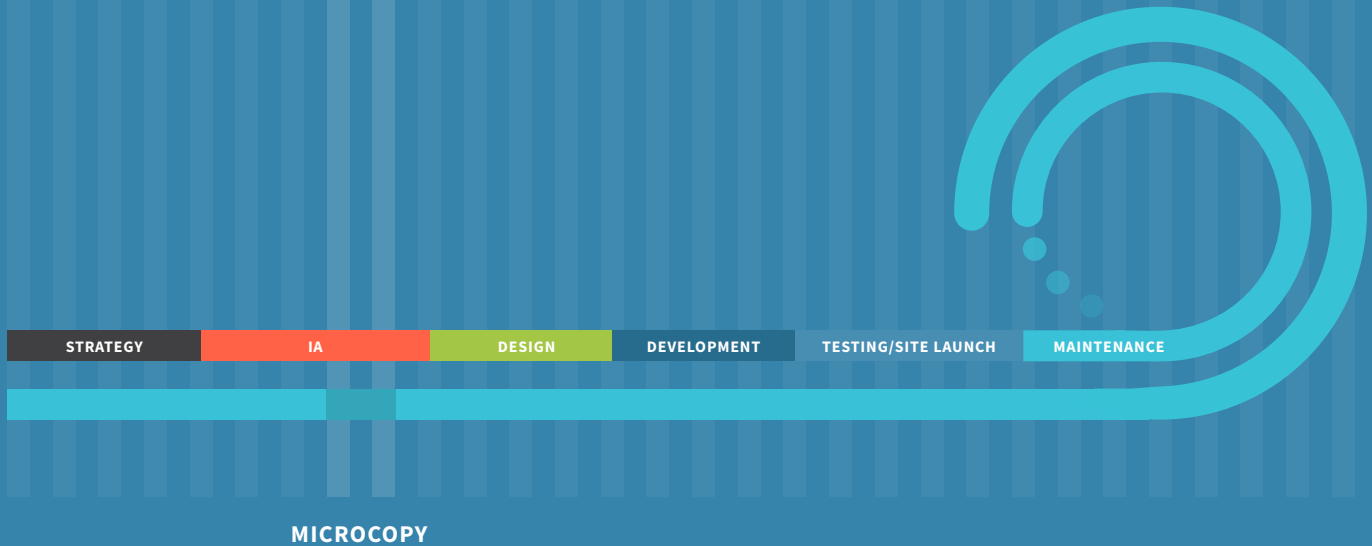
These groups of words are so small, most people don't give them a second thought. But these few words have an enormous impact on the user experience! Microcopy often falls victim to personal bias, internal terminology, poor branding, broken contextual flows, time crunches, and other factors. Any of these can undermine even the most well-designed UX and the copy within.
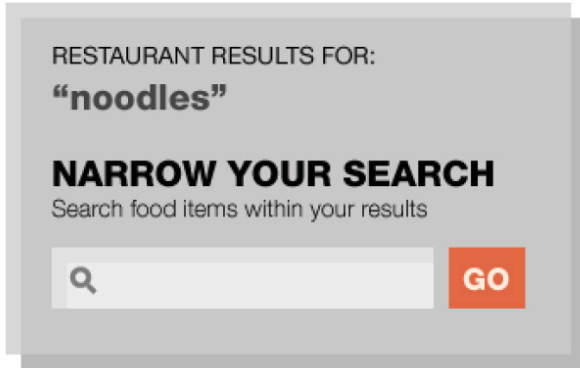
Here are a few tips for writing effective microcopy:

- **Use an appropriate number of words.** There's a fine line between using too many words and being overly sparse. Write microcopy in as few words as possible, but use enough words to ensure the meaning is clear.

EXAMPLE

"Choose the best answer. Then click Submit."
These very brief test instructions are direct and to the point. It's doubtful they could be misinterpreted.

- **Remember its value**. Microcopy is a service to your audience. Through these little phrases, you are helping users get oriented, find their way, or feel assured. Use care.

- **Use the vocabulary of your audience.** What words will trigger a desired behavior?

RESTAURANT RESULTS FOR:
"noodles"

NARROW YOUR SEARCH
Search food items within your results

🔍 [   ] **GO**

EXAMPLE

"get it", "download", or "buy".

- **Run through the interactions in context** – Run through the actual flows your users will follow and experience your microcopy in context.

- **Do it right!** - Unlike your other content which will evolve and change over time, microcopy is often hard-coded. Make sure to address it when you have the chance.

4. Search Engine Optimization

SEARCH ENGINE OPTIMIZATION (SEO) IS AT THE TOP OF MOST MARKETING AND WEB PROFESSIONALS' MINDS. BUT THERE ARE SOME KEY ELEMENTS THAT OFTEN GET OVERLOOKED IN THE THROES OF A REDESIGN.

Let's start with the basics. SEO is the means by which organic (not paid) traffic finds your site via search engines like Google.

SEO is impacted by factors both visible and invisible to the web users. But it is the invisible or less-visible elements that are in danger of being overlooked.

H1 HEADLINE
Chocolate Donuts from Mary's Bakery

IMAGE FILENAME
Chocolate-donuts.jpg
ALT TEXT
Chocolate Donuts

BODY TEXT
chocolate donuts
donuts
chocolate

Keywords – are the terms users search to find you. These should be defined and prioritized early on so they can inform things like navigation, taxonomy, and page titles. When defining terms, it's important to think like a user.

| EXAMPLE	TIP
Will your users search for an industry term like "sofa" or "couch," a more colloquial term?	The Google Keyword Planner is a great tool.

Page Title- Sometimes called "SEO Title," the page title is the main description of your page and one of the most important on-page elements in SEO. It appears in three key places – browsers, search engine results, and external websites like social media sites. Try to keep these below 55 characters.

Optimal Format: Primary Keyword – Secondary keyword | Brand Name

Meta Description – Sometimes called "SEO Description," the meta description is not important to search ranking, but is vital to gaining click-through. These are the descriptions a user sees in search results. They should employ keywords intelligently, but also create a compelling description that a searcher will want to click. The description should be between 150-160 characters.
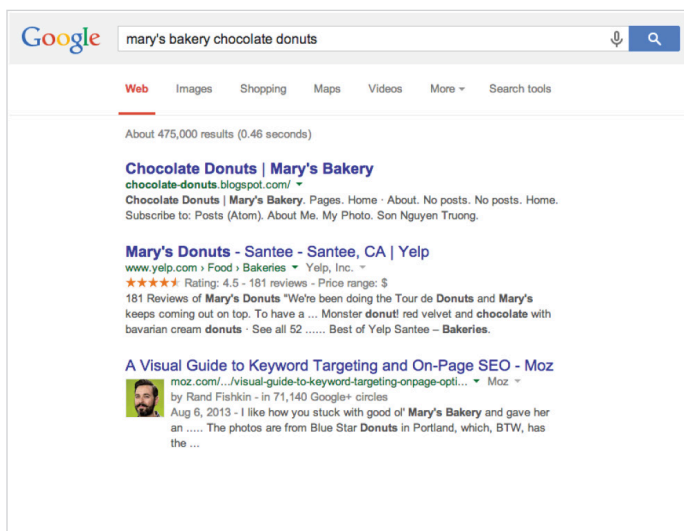
The Page Title and Meta Description should not simply repeat the content of the page. They should be geared toward providing a bird's-eye view to an unfamiliar user.

Unfortunately these are an oft-forgotten content element, leading the content migration team to leave these fields blank or improvise.

Alt Text – is the description of an image on a page. This is designed for visually impaired users, but factors very importantly into search ranking. This text should be short, employ keywords, and describe the image or repeat any text in the image.

EXAMPLE

Our logo Alt Text should be "EPAM: Results Relentlessly" not "Company Logo"

5. Map Content to New Structure

A NEW SITE DESIGN OFTEN MEANS ENTIRELY NEW PAGE TEMPLATES OR MODIFICATIONS TO AN EXISTING ONE. BUT HOW DO YOU TRANSFORM CONTENT FROM AN EXISTING LAYOUT INTO A NEW TEMPLATE?

| STRATEGY | IA | DESIGN | DEVELOPMENT | TESTING/SITE LAUNCH | MAINTENANCE |
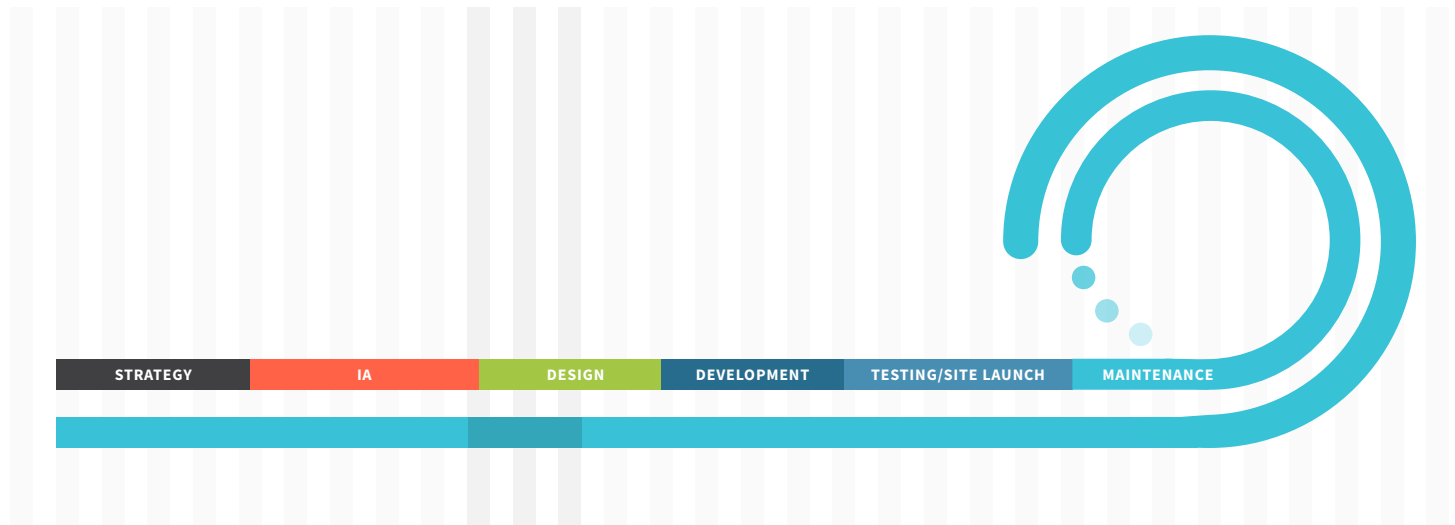
MAP CONTENT

Mapping content can reveal content holes.

One tool that can be useful is the "copy deck." A copy deck is a set of content templates that serve two purposes: it's a paragraph-level companion to your website's wireframes, and it's a simple, effective means of getting useful information from your experts to your writers. (It is not the same thing as an HTML template you feed to your content management system.)

PAGE LOCATION Location on site map

PAGE NAME Use page name on site map

PAGE HEADLINE

PAGE SUB-HEADLINE

PAGE CONTENT
Place in the order you want it to appear on the webpage
Include ALL relevant URL links
Section home pages should NOT have more then 100 words

CALLS TO ACTION Reference content title from the calls-to-action library

| **PRIMARY ACTION** Title, Link | **SECONDARY ACTION** Title, Link | **TERTIARY ACTION** Title, Link |

PAGE DETAILS

MAIN PHOTO Make a suggestion or send photo with copy deck

PAGE DESCRIPTION What is the page or section about? Approx. 20-40 words

KEYWORDS

You might think of content templates as a kind of wizard for content development. Whereas branding, voice, and editorial guidelines are often prosy and stylish, the content templates should be lo-fi, simple, and relentlessly practical, and they should contain at least the following information:

THE PAGE TITLE

• A short description of each chunk of content that will be on the final page, including what each chunk of content must do, and what formats it can be in (paragraph, simple bulleted list, multi-level bulleted list, data table, screenshot, callout box, etc.).

• Examples of each chunk of information, written by actual writers and supplemented by inline guidelines as needed.

A COPY DECK IS USUALLY CREATED BY CONTENT STRATEGISTS, BUT IF YOUR PROJECT DOESN'T HAVE A DEDICATED CONTENT SPECIALIST, THE TEMPLATES CAN BE PRODUCED BY AN INFORMATION ARCHITECT, PROJECT COORDINATOR, OR OTHER PERSON WHO IS IN CHARGE OF YOUR CONTENT.

Content templates can be useful for one-off pages like About Us, but they're particularly effective when you're working with whole classes of pages, like product pages or departmental landing pages, all of which need to contain similar information, presented in a consistent way.

By letting you show your experts exactly what kind of content you need for each page, content templates can help you:

• Collect information more quickly, by giving experts an easy, fill-in-the-blank structure to work with.

• Speed up and simplify the content development process by producing more uniform first drafts that are easier to turn into final web copy.

• Improve the structural consistency of your final content.

• Reveal any gaps between the communication needs of the organization's various divisions and the content structure you thought you needed, while there's still time to fix it.

6. How Do Taxonomy and IA Inform Content?

A TAXONOMY IS A CONTROLLED VOCABULARY OF TERMS AND RELATIONSHIPS WHICH SHOULD BE USED TO CATEGORIZE AND CLASSIFY CONTENT.

The wireframes and sitemap, along with the taxonomy structure, provide guidance to the content editors and creators for creating and organizing content within the content management system.

Smart, well-structured metadata helps content managers to identify, organize, use, and reuse content in ways that are meaningful to key audiences. The most important set of metadata is your taxonomy.

Taxonomy terms are applied to content within the content management system (generally represented as "tagging") and are

not visible on the site. However, parts of the taxonomy may sometimes be exposed directly to end users in the form of faceted search (sometimes also referred to as faceted navigation), personalization, and content aggregation (lists and archives).

The key is to organize content in the single best location, typically identified in the sitemap, and tag it with only the most relevant tags. This will ensure that content can be found in a logical way using your site's primary navigation, contextual navigation, and search.

7. Maintain Good Content

ONCE YOU'VE PUT IN ALL YOUR CONTENT YOU'RE DONE, RIGHT?

Not unless you want your users to be done too. Your website, intranet, or portal must continue to evolve along with your company and your users. Ongoing maintenance consists of three steps: audit, analyze, and edit. These steps should be taken at least quarterly to ensure continuous high quality content.

Audit

It is important to conduct a regular audit of content. This is essentially an inventory of all of the site's content. An audit can often be automated using the content management system or online tools. An audit can help identify orphan pages and content that falls outside the approved site architecture.

Analyze

Review ALL content and answer the following questions:

• Does the content still convey the right message?
• Is the content accurate?
• Is the content still useful to its intended audience?
• Is there newer information that can be included?
• Is the content using the best format/presentation?
• Are the right taxonomy and metadata terms applied?

Look to your analytics. Only web analytics will provide you with hard data about how your content efforts are resonating with users.

Web analytics can help you prioritize which content to focus on first, where your efforts have the greatest impact, and how you can clean up your content to help users find information and complete tasks as quickly and efficiently as possible.

Edit

The final step is to edit content based on measurement and analysis. An editorial calendar can make it easier to plan for more complex changes, but make sure to take the time to address all site content as well.

Conclusion

A content migration plan and content strategy should take place throughout the entire project cycle. It should begin before or during the discovery phase and continue past development. Content is one of the most important parts of your site so give it just as much attention as the design, code, and testing. Content is the real reason your users will come back to your site!

Questions? Contact us at Sales@EPAM.com

FOR MORE INFORMATION, PLEASE VISIT EPAM.COM

Established in 1993, EPAM Systems, Inc. (NYSE: EPAM) is recognized as a leader in software product development by independent research agencies. Headquartered in the United States, EPAM serves clients worldwide utilizing its award-winning global delivery platform and its locations in 19 countries across North America, Europe, Asia and Australia. EPAM was ranked #6 in 2013 America's 25 Fastest-Growing Tech Companies and #3 in 2014 America's Best Small Companies lists by Forbes Magazine.

FOR MORE INFORMATION, PLEASE VISIT EPAM.COM

AUTHORS

Shivani Aneja
Director, UX Design and Content Strategy

Matt White
Director, Client Engagement

SALES CONTACTS

Please contact us for any sales inquiries or general information on EPAM.

41 University Drive, Suite 202
Newtown, PA 18940, USA
P: +1-267-759-9000