Ten Strategies for Content Migration

NOVEMBER 2005

By Chiara Fox
with Margaret Mason
chiara@adaptivepath.com
## Contents

Executive Summary ................................................................. 3

Ten Strategies ......................................................................................... 4

1. Choose your leaders early ............................................................. 4

2. Assemble the right team ............................................................... 4

3. Know what you have ................................................................. 5

4. Cut out the ROT ......................................................................... 6

5. Create a database or tracking system ........................................... 6

6. Determine where content will live ............................................... 7

   - File-Naming Conventions .......................................................... 7

7. Migrate in waves ........................................................................ 8

8. Set up a war room ...................................................................... 8

9. Create a Migration Emergency Kit ............................................. 9

   - Migration Emergency Kit Contents ........................................... 9

10. Fix what’s broken .................................................................... 9

The Launch ....................................................................................... 10
Executive Summary

After hundreds of hours of planning, you’re finally ready to launch your redesign. The architecture is genius, the CMS is humming, every last detail is impeccable — all that’s left is to transfer your old content.

What a mess.

The bigger the site, the more arduous the content migration, but even small sites require a great deal of forethought. If you don’t want a single typo, broken link, or missing image on your shiny new pages, then you’ll need a system to manage the transition. Having one in place from the start saves you countless headaches later.

My ten strategies for content migration are drawn from my experience working with PeopleSoft. During my tenure there, I helped manage four content migrations. The last one involved integrating content in ten languages on 23 sites. Oddly, that one was the easiest—because we had a system and a much more experienced team.

This report provides ten strategies that will help you break your content migration into more manageable steps, prepare your team quickly for the work ahead, and still ensure attention to detail:

1. Choose your leaders early, so the project has early direction from a top decision maker and a team manager.

2. Assemble the right team to reduce potential errors.

3. Know what you have by organizing a methodical content inventory.

4. Cut out the ROT—redundant, outdated, or trivial text.

5. Create a database or tracking system so you know where all of your data will settle on the new site.

6. Determine where the content will live by referencing the overall site structure and assigning spots to low-level pages.

7. Migrate in waves by staging rolling content freezes.

8. Set up a war room where the whole team can work together each day.

9. Create a Migration Emergency Kit containing detailed reference materials.

10. Fix what’s broken by retaining part of the team for quality assurance.

Readers will come away with working knowledge of how to negotiate even the most complex content migration.
Ten Strategies

1. Choose your leaders early.

Content migration works best with team leaders in place from the start, so choose a decision maker and a team manager right away.

Your decision maker is the face for C-level management. This person keeps everyone updated and serves as a liaison between management and the team.

Team managers make the content migration happen. Many organizations try to divide this duty amongst several managers. Unfortunately, most managers will be swamped with organizing their portion of the content migration and dealing with day-to-day business issues. Find someone who has a grasp of the big picture, and make this person the single point of contact for any migration questions.

On your team, the decision maker and the team manager might be the same person. This is fine, as long as that one person has a firm grasp of site strategy, plus the time, the ability, and the authority to make decisions in consultation with top executives.

2. Assemble the right team.

A dedicated team ensures that your migration happens efficiently, and reduces potential errors.

Obviously, more complex migrations require more staff. If your team is larger than five or six people, it’s best to organize into sub-teams, or to create separate groups within departments. For instance, a content-based site may have an entire editorial staff that reports to a chief content officer, who then feeds editorial to the technical staff.

For most companies, the following roles will be helpful:

- **Project Managers** keep the team on budget and on time. They take care of the nitty-gritty details of scheduling and work closely with the team manager or decision maker.

  Project managers might also work with the IT group and its project manager. They let the team manager know when work is slipping, and whether it’s necessary to tackle a problem from a different angle.

- **Information Architects** determine the structure of your site, and understand how all of your content links together. IAs develop an intuitive navigation system, and can work with content migrators to help them figure out where each content object belongs. IAs also help with your metadata strategy and ensure consistency in tagging. They work closely with IT, explaining functionality issues.

  When you’re looking for an IA, experience as a project manager or designer is a good sign, as is a degree in a field that emphasizes information handling, such as library and information science, human-computer interaction, design, or even journalism.
• **Content Owners** are subject matter experts. They’re often the ones who originally wrote the current site content, or at least wrote the original documents that became the web content. Content owners look for errors on finished pages. Because their skills are often in demand at a company, waiting for them to approve content can create a bottleneck. Consider working on the assumption that silence equals consent. You can’t hold up the entire project waiting for a content owner to get back from vacation. They can always tweak the content once the site has launched.

• **Migrators** are there to populate the system. They do everything that hasn’t been automated — metadata tagging of content objects, routing content for approvals, and so on. It’s helpful for migrators to have professional writing or editing experience, so they can turn out copy on the spot and edit as they go along.

Depending on your time frame and budget, you may want to hire migrators as contractors. If you have a large web team and an extended time frame, it may be possible to work with the resources you already have.

• **Graphic Designers** create banners, buttons, and any unanticipated graphics that you need on the new pages. You’re looking for folks who can churn out files from Photoshop or Flash.

• **IT Team Members** ensure that servers are running, and build scripts or tweak templates as necessary. Hopefully you’ve finished development on the CMS before you start populating it with content.

3. **Know what you have.**

Next, you’ll need to conduct a content inventory, which is a methodical audit of your website’s content. Through this process, you’ll list every page, and begin to mark page status — frozen, finished, etcetera. Note the identity of the page owner, whether the page needs graphics, copy editing or rewrites, and where the page will live in the new architecture. When you’re finished, you’ll have a definitive map of all the information on your old site, and a solid idea of what should transfer to the redesign.

You’ll find thorough content inventory instructions in the article [“Doing a Content Inventory” by Jeffrey Veen](http://www.adaptivepath.com/publications/essays/archives/000040.php) by Jeffrey Veen, Adaptive Path co-founder and director of product design.
4. Cut out the ROT.

Once you have a good catalog of your legacy content, it’s time to get rid of ROT — redundant, outdated, and trivial content. Examine each piece of content, and see how it measures up to your current business objectives. There’s no point in migrating content that’s dated or inaccurate. Judge the value of each document by labeling it “essential,” “useful,” or “unnecessary.”

This process should happen early, because you might have a lot of items that need updating. It’s better to get rid of useless content and go live with smaller sections. You can always flesh them out later, but publishing useless information on a new site hurts your credibility.

Killing ROT early will also free up your content owners — if their section thins out, they can start making new content for you.

5. Create a database or tracking system.

Now that you’ve catalogued every piece of data on the old site, you need some way to track where it’s going. You can always turn your content inventory spreadsheets into content migration spreadsheets; however, spreadsheets can quickly become unwieldy, and keeping them current becomes its own project.

The best way to track pages is by creating a simple database that catalogues data as it moves through the process. This way, the whole team can see the status, and you don’t have to waste time combining and updating spreadsheets. If you already have a database — even something simple like MS Access or MySQL — that’s ideal.

Your database will note where pages are coming from, how they merge, and where they live in the new architecture. It records who the content owner is and which producer is assigned to the page. It can also include the level of change necessary — does it need to be rewritten, or does it just need editing; does it need graphics?

Give each team member read-only access to the entire database, and the authority to write in their sections. This way, team members can see how everything is going, and can update the database as they progress in their own sections. The team manager, project manager, and possibly IA should all have complete access.
6. Determine where content will live.
At this point, the IA will have a blueprint of the future site, but the plan isn't set in stone. You'll know the overall site structure — the first two or three levels — but more granular, low-level pages need to be positioned now.

Together, the IA and content owners determine where individual content objects will live in the new system, and decide what local and contextual navigation is necessary at the bottom-most level. At this point, either the IA or the content manager communicates file structure and naming conventions to the rest of the team.

If you have content that needs rewriting, you can prepare new text in anticipation of repopulating the site, or write it during the actual process. Now is the time to decide.

Once you've figured out how to produce content, choose where to store it. Presumably, you have a file structure set up to handle this project. It's convenient if that structure can mirror your site structure, but you may have established URLs that need to remain unaltered.

Every time you change a file or directory name, you're breaking a URL. Yes, you can employ technical solutions, like redirects and 301 and 302 server codes, but be careful. It's a lot of work,

---

### File-Naming Conventions

- **Be consistent.** Plan your directory structure and names before you start populating the system, as the directory names form the basis of the URLs.

- **Consider spelling words out.** Balance the use of abbreviations with the need for clarity.

- **Include keywords.** This can help with search engine optimization for external search engines like Google or Yahoo.

- **Remember the basics.** Standard Unix filename rules still apply — no spaces, no special characters. Decide whether you'll use a dash, an underline, or nothing to separate words.

- **Make archiving easier.** Choose filenames that facilitate archiving.

- **Mind your audience.** Your users may or may not know how to hack a URL. It's always better to use words in your directory and filenames than a string of random letters and numbers.
you'll break people’s bookmarks, and depending upon the type of redirect you use, you may sacrifice search engine rankings.

Team members need a printout of where content lives in the directory structure, and where it will end up in the architecture. This helps them determine which templates to use for each chunk of content, and shows them where their section exists on the site. The content migration spreadsheet or database is an excellent place to store this information so everyone has access. More on that later.

7. Migrate in waves.

Be firm in deciding which pages will remain static, or “frozen,” and when. You shouldn’t change frozen pages until the new system is live. Everyone will think their pages can’t be frozen, but to get through everything, you must freeze most of the site early on.

Depending on your site size, you may need to go through two or three sets of freezes. On large sites, making changes in waves helps you set goals and map your progress.

The first content freeze should happen about three months out from launch. Aim to freeze 50 percent or more of the site. The first to go are “evergreen” pages with content that doesn’t change often, like bios and product pages.

Two months before launch, freeze another 20 percent of your content, and another 20 percent one month before. The remaining 10 percent is what I call “slush.”

Slush pages include time-sensitive content, like press releases or event pages. They can also be pages that started out frozen, but had change requests from higher-ups. (This means that you’ll have to change the page in both the old and the new systems, creating more work for the team.)

Aim to constrain the slush by requiring greater approval authority for changes as you approach the launch date. The week before launch is reserved for last-minute IT changes and QA — you should be done with migration by that point.

8. Set up a war room.

Set aside some physical space where the entire content migration team can work together every day. It saves quite a bit of time, because everyone eventually has the same questions: What’s the hex code for ampersands? Which template do I use for a press release? Has IT fixed the display bug with banner images? Write answers to the most frequent requests on a shared white board.

Every night, the entire team should meet in the war room for a status report. People will move through the migration material at different rates, so the project manager can reassign work as you discover which sections will take more time and care. The team manager uses these meetings to keep abreast of potential problems.
9. Create a Migration Emergency Kit.

Each team member should have a Migration Emergency Kit, a binder that contains extremely detailed reference materials, like a recipe book with screenshots.

It tells employees how to make file names, how to navigate in the system, where to save content, and so on. A newly hired contractor should be able to figure out how to create a new page in the system using only your kit. This collection of documents will evolve quickly, so expect to hand out supplements or updated sheets nearly every day.

10. Fix what’s broken.

Even the most meticulous process can’t ensure that you won’t need to tweak anything when the migration is complete. You’ll have a lot of quality assurance to do after launch.

Retain a few team members to fix typos, URLs with errors, and broken links. You can use a bug-tracking database to manage both content and system bugs.

Switch assignments so team members aren’t reviewing the sections they migrated themselves, everyone should have fresh eyes. Translated content may require native speakers to check for spelling errors, but you can still monitor basic functionality like links, images, and encoding problems.

Content owners who have had their content frozen for three months will be itching to update their content, but give yourself time to get everything working before you make further changes.

---

### Migration Emergency Kit Contents

Some suggestions on what to include in your kit:

- Navigation map or blueprint
- Screenshots of interfaces
- An appendix of codes you use regularly
- Naming conventions
- List of contact information for everyone on the team
- Approved color and font lists
- Language style guide with a list of preferred reference materials
- Graphics style guide
- Example pages indicating the ideal for a products page, a press release, and so on
- Step-by-step guides that offer how-tos on metadata, creating a new page, editing established pages, and so forth
Once you’ve established a system, the best thing you can do for your team is to let them know how much you appreciate their hard work. No matter how well you plan, there will be bugs, servers will crash, and code will break. It can be discouraging.

Use your daily meetings to boost morale. Give out small rewards to those who migrate the most pages in a day or who have the fewest errors in their code. A gift card for the local coffee house, lunch on the company, a box of chocolate, these small tokens of appreciation are a big deal to someone who’s been cutting-and-pasting eight hours a day.

During the launch, truck in an endless supply of pizza and sodas, and put on some good music. Once everything is live and working, gather everyone around, pop a cork, and celebrate a job well done.

Chiara Fox is a senior information architect for Adaptive Path, the world’s premier user experience consulting company. Chiara has developed successful information architectures for intranets, informational websites, and e-commerce sites. She’s worked with Fortune 100 and 500 companies such as PeopleSoft, AT&T, Square D, L.L. Bean, and Hewlett-Packard.

Before joining Adaptive Path, Chiara served as the senior information architect for PeopleSoft, an enterprise software company. She implemented a unified information architecture across the public website, customer extranet, and partner extranet in a single CMS. The IA was expanded to include 23 international sites the following year. She also designed the metadata schema for use on all 26 websites. Before joining PeopleSoft, Chiara was an information architect at Argus Associates, a pioneering information architecture design firm.

Chiara is an active member of the American Society of Information Science and Technology and ACM’s local chapter of SIG-CHI, BayCHI. She is also a member of the Asilomar Institute for Information Architecture and serves as co-director of the Institute’s IA Library.

Chiara received her Master of Science degree in Library and Information Science from Simmons College. She spent ten years working in public and scientific libraries as a cataloger and reference librarian before devoting herself to the web full-time. Because of her background as a librarian, Chiara specializes in content analysis, metadata and taxonomy development, and building architectures from the bottom up.