Top 10 Best Practices in Content Migration

Open Text Connectivity Solutions Group

Abstract

The ability to easily view information from across the organization in a way that is seamless and transparent has unequivocal business value. However, in today's global economy, an ever-expanding proliferation of data sources—both structured and unstructured, and in various formats and software applications—has brought about a new mindset. No longer is information simply data, but content; content that can be interpreted, integrated, and made more meaningful to corporate end users. This paper discusses the best practices when planning a content migration project.

Contents

Content Migration: Move It and Use It

The ability to easily view information from across the organization in a way that is seamless and transparent has unequivocal business value. However, in today's global economy, an ever-expanding proliferation of data sources—both structured and unstructured, and in various formats and software applications—has brought about a new mindset. No longer is information simply data, but content; content that can be interpreted, integrated, and made more meaningful to corporate end users.

Migrating content from one or more sources to a new environment allows a unified view and unified access to content in order to increase efficiency and maximize investment. Users can essentially tap into corporate content that might otherwise have been inaccessible or too time-consuming to locate. In some instances, content users might never have been aware existed. This enhanced view allows users to leverage all business-critical content from across the organization and use it to extract valuable information on customers, vendors, services, and products.

Single View Benefits

There are many benefits to migrating relevant content to a single repository in your organization. Consider these:

- Allowing fully informed, hence better business decisions
- Decreasing cost of running many different applications
- Mitigating risk by having centralized content
- Improving departmental efforts, i.e., customer service, sales, HR
- Consolidating information for quick and easy access
- Managing permission rights, warranty processes, etc.
- Tracking trends, customer feedback, financial audits
- Using digital media for multiple purposes, such as marketing, opinion surveys

These tangible business benefits can translate into business-critical actions. For instance, being able to decommission applications with integrity means the company will no longer have the associated maintenance costs of running the application back to the vendor. An organization with multiple businesses can minimize its effort to meet compliancy and discovery rules mandated by the government and/or corporate policies through migrating content to a prime application. Sales staff can get a complete customer profile without running a host of separate actions. And the burden on IT will be reduced with only one environment in which to search for problems or manage permission rights.

One of the most compelling benefits of content migration is the “single view of the truth” it offers business users, IT staff, and corporate stakeholders. Take a food manufacturer faced with a product recall. Time is of the essence during such a situation and a single repository of information makes it possible to quickly and easily meet the demands for divergent pieces of information such as a label image, supply chain records, and marketing campaigns. A fast response time can significantly decrease exposure to risk and avoid costly lawsuits.

Structured Versus Unstructured Content

The value of unified access to content is indisputable; however, until recently there existed an inherent organizational challenge in content migration: bridging the gap between divergent systems and two distinct classes of information. The first class, structured content, retains the same context or application as the source. It is information used to organize, analyze, and quantify records. The second class, unstructured or non-relational content includes data in text memos such as email, speech, audio, video, and images. Unstructured content has traditionally been slower to integrate because it is more difficult to analyze than structured data, the algorithms are more complex, and the sheer volume makes it more time-consuming to cleanse and filter.

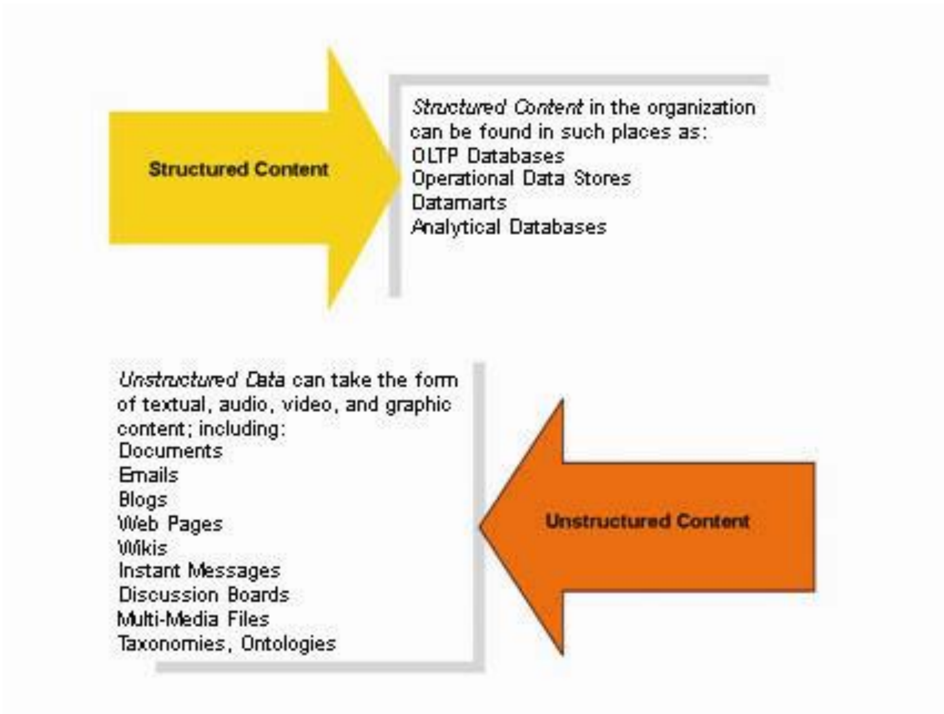

Figure 1
Analysts predict traditional databases of structured content account for 20% of the content in most organizations while sources of unstructured content dominate, and continue to grow

Fortunately, the technology and tools now exist to migrate and marry structured with unstructured content. For an end user, unifying this content puts an end to often futile, multi-layered system searches. It is now possible, for example, to move a large quantity of content into a warehouse and combine it with data from SAP or SalesForce.com. Now you really have an opportunity to be efficient with your information and have both structured and unstructured in a single repository with unified access to all types of information.

The purpose of this whitepaper is to take you through the journey of the 10 best practices around content migration. Adopting these best practices is a way to give your migration project a methodology and focus before, during, and after completion. Together with Open Text's tools and technical know-how, potential roadblocks will be eliminated, more predictable outcomes attained, and unnecessary project risk avoided.

Top 10 Best Practices in Content Migration

1. Define Target Context

A first key step when planning a migration project is to define the functionality of the target context. Knowing what end users want the new target repository to look and perform like on completion is a high-level priority, and from a technical standpoint, a way for IT staff to establish necessary protocols. This important step is not always given due consideration by teams that jump in and focus instead on the source content to be migrated.

Take the time to dig in and draw a clear and complete picture of the new repository or application. Review and refine the target based on the source information to identify gaps that exist or enhancements needed. Careful due diligence at the outset will better leverage target capabilities at completion, in addition to adding business value by improving access to content for better decision-making. Keep in mind that the migrated content will not simply be copied and moved from the old infrastructure into to a new environment. In order to fit the demands of the new system, information will have to undergo change to conform. Without a thorough grasp of what to expect it is a certainty that the project will either encounter overruns of time and/or budget, or fail to meet end users' business requirements altogether.

2. Understand Scope and Scale

It is impossible to set accurate budgets and timelines without first understanding content complexities, relationships, quality, and volume. Spend time making accurate accounts of the time and resources your migration project will take—including expenditures for hardware, software, and personnel. Detailed discussions with IT staff, business analysts, and end-users are essential during this fact-gathering stage to make a full and accurate inventory of the content to be migrated…or to be left behind.

It is recommended that the services of a knowledgeable consultant are engaged to help guide you through this discovery process. An experienced expert will be worth the expenditure as his or her services will mitigate project risk by helping your team be aware of, and plan for, any and all eventualities.

However, even when external consultants are used, migration decisions should always be the responsibility of a business prime. This is a person of authority from the organization who has ownership of the project. He or she must oversee whatever external resources are brought in as well as provide leadership of the content migration team, keeping the group fully cognizant of the techniques, strategies, tools, and mostly importantly, the desired corporate outcomes. The business prime should also ensure that the project has executive sponsorship and buy in.

3. Strategize

A content migration almost never occurs in isolation but as part of a broader application, upgrade or implementation. If the migrated information is going to deliver full operational value and efficiency, then strategic planning to address project-specific requirements must take place early on, and not as a rushed afterthought to the bigger project.

Migration strategy should address high-level business decisions surrounding the original source content being migrated to the new source target repository. Though content migration is in itself a technical process, it is essentially a business issue. First and foremost it is about people and their access to and relationship with information. Collaborative strategy sessions must occur between all business users, stakeholders, and IT staff so that project goals are attained.

A gap analysis should be done to determine if the content to be migrated will match the requirements for content in the new system or if there is any missing information to still be collected. It is also important to identify and analyze taxonomy differences when consolidation of source systems is necessary.

4. Assess Source

Source content is often comprised of a wide ranging and disparate collection of structured and unstructured data. In order for this myriad collection of information to be transformed into content that can eventually be accessed through a single company view, it must first be located, examined, and defined. Data profiling and assessment capabilities available through Open Text will help identify opportunities and complexities from the source system, such as levels of data relationships.

Invite discussion between relevant business users and IT staff; collaboration is key so that no content is overlooked or obtained without permission. A thorough assessment of source content will weed out good from bad content, and point to errors, omission, discrepancies, inconsistencies, and relationships surrounding content. A detailed understanding of the business needs along with the use of appropriate tools here, and throughout the project, will help eliminate duplication and ensure that only relevant, quality data is preserved and transferred.

5. Cleanse Content

A migration project is the perfect opportunity for a site cleanup. Content owners must be encouraged to sift and sort through information, removing outdated or redundant content, thus eliminating the volume of information to be moved. Content cleansing tools can be useful here as they allow information to be brought up to standard and its quality to be measured.

Keep in mind that the purpose of cleansing is not to attain perfection in content quality. In fact, striving to do so can be a lesson in futility. Even if information is one-hundred percent accurate when entered, it can deteriorate almost immediately. In the real world users of the system are often too busy to record changes as they happen, if at all.

The effort put into cleansing content should be dependent on the business impact if the content value is incorrect. Will a new software program run with incomplete or inaccurate data values? How critical is it to have a post office box but no street address for a customer? Questions such as these will help determine the level of accuracy required across multiple areas of the project.

In almost all cases, starting with a clean slate is preferable to sullying the new environment with out-of-date content. There are instances when cleansing could be done post migration, such as when the new environment offers tools that make some kinds of clean-up very easy and robust, or if you're faced with a time crunch and migrating content "as is" is a better alternative than postponing the final migration and go-live.

6. Map Source to Target

Mapping is the vital link that bridges information moving from point A to point B, or in a content migration, from source to target. To illustrate the concept of mapping, imagine that a library wanted to change its classification system from the Dewey Decimal system to a different set of categories defined by the Library of Congress. Mapping scripts are necessary to show how the old cataloguing system's content can be made relevant in the new model.

Mapping should encompass all information, including content metadata and the business and system rules that define the transformations required to convert the information to be compliant with the new system, such as transforming content from the Dewey Decimal system to match the Library of Congress system. A well-defined mapping methodology will focus only on the content to be migrated; extraneous material should be flagged and left behind, as it has no relevance in the new system. Be aware of the importance of having transformations that can be backed out of if part of the migration does not work. Keeping copies of the original data will help make this possibility possible.

7. Determine Migration Method

There are three main strategic options for migrating content: big bang, parallel running and phased. A secondary migration option is referred to as zero-downtime. Each migration method has its own merit and the one chosen should be weighed against specific operational variables. Examining what service level agreements exist in your organization—whether formal or informal—-will point to the performance response time your customers expect. Ultimately, the importance of the application to the organization and its change management practices will help guide your decision.

Big Bang	As the name implies, happens all at once, typically over a weekend. This reduces costs as you don't have two systems running simultaneously. The downside to this method is if delays occur, or if the new system doesn't work as planned, there may be negative impact to your business's top and bottom line.
Parallel Running	Both old and new systems run together. When the content is loaded and thoroughly tested a cutover is made. The advantage here is that there is little likelihood of problems as the systems will run together until the new one is proven. The downside is that synchronization software to keep both systems running in tandem can be costly.
Phased or Incremental	Allows for a gradual transition whereby you turn on only parts of the new system at one time. The advantages are that you can start getting results sooner; it's easier to test smaller, incremental datasets; and you can back out before committing to the new system.
Zero-Downtime	A secondary migration strategy that is implemented with content that is mission critical for 24/7 applications. This approach is implicit in both parallel running and phased migration methods.

8. Analyze Risk & Plan Contingency

A thorough risk analysis should include the building and testing of a prototype system. Representative data will help define performance requirements and serve as a benchmark during testing for software/hardware compatibility, downtime, and budgetary projections. Run the migration program completely against a replication of all the content in a lab environment before running it live. Take all the steps necessary to validate this run and be sure to involve representative groups of end users and internal training services. When the risks have been investigated, results should be signed off on by the migration team.

Have a contingency plan for the first day's support and a plan to keep content live during the migration period. It is critical to have proper documentation. Keep a copy of the original content for testing and training purposes in the event part of the migration does not work and you need to go back. Even then, remember that the migrated content has been restructured for the new system and context will have changed, hence it will be difficult to compare, if not impossible.

9. Migrate Content

Once all the proceeding best practices have been addressed, end users trained on the new system, the migration method agreed on, and the migration date selected, it's time to go ahead and guide the migration process. Again, an experienced consultant will prove an invaluable resource here as content migration is often a one-time event. Having a knowledgeable expert who can leverage past experience will reduce risk of project failures and maximize the value of the software, services, and personnel investments made in the migration efforts.

Before proceeding with the actual migration, it is important to review all the guidance and best practices of the previous steps. Ensuring that the objectives are being met and contingency plans in place, you can feel confident moving forward with the content migration.

10. Validate & Document

The migration team's work is not finished until confirmation that source content has been successfully moved into the new target. Build checklists to make content movement and integrity easier to check and track. Migration expectations against actual results should be documented in a report to validate migration completion and quality. A good audit will involve conversations with end users to ensure everybody is happy with their new single view of information. These are the people with daily hands-on use and knowledge of the new model, and often the true "test" of success is user acceptance.

A final step is to save and archive migration scripts. Content migration is often a one-time exercise; however, with the right tools, the protocols, mappings, and scripts can be reused in future projects within the organization. A documented report of the migration process will serve as a repeatable reference guide.

Summary/Conclusion

The process of content migration need not be a complex, costly, and time-consuming drain on company resources. Some potential pitfalls can be avoided by following the best practices presented in this whitepaper and engaging the services of Open Text, a company with the technology, tools, and consulting know-how to guide your project.

About Open Text

Open Text is a leader in Enterprise Content Management (ECM). With two decades of experience helping organizations overcome the challenges associated with managing and gaining the true value of their business content, Open Text stands unmatched in the market.

Together with our customers and partners, we are truly The Content Experts,™ supporting 46,000 organizations and millions of users in 114 countries around the globe. We know how organizations work. We have a keen understanding of how content flows throughout an enterprise, and of the business challenges that organizations face today.

It is this knowledge that gives us our unique ability to develop the richest array of tailored content management applications and solutions in the industry. Our unique and collaborative approach helps us provide guidance so that our customers can effectively address business challenges and leverage content to drive growth, mitigate risk, increase brand equity, automate processes, manage compliance, and generate competitive advantage. Organizations can trust the management of their vital business content to Open Text, The Content Experts.

http://connectivity.opentext.com

Sales: connsales@opentext.com
NA Toll-free 1 877 359 4866
Worldwide +1 905 762 6400

Support: integrationsupport@opentext.com
NA Tool-free 1 800 486 0095
Worldwide +1 905 762 6400