Improving Information Use
with JustSystems' Metadata Tools

Analyst: Anne MacFarland

Management Summary

The great benefits provided by easy access to information are often constrained by doubts. These doubts can be about the source of the information, its staleness, or its relevance to the need. Like a chowder with unknown contents, consuming contemporary information can be uncomfortable.

Some of this unease may come from enormous changes in the way we get information. In the pre-Gutenberg era, when books were few and most information-transfer traditions were oral, the information people leveraged was comfortably limited to what they could remember. A second era, which brings us up to the Internet, was the heyday of the manufacture of information for consumption. The costs to get into the business (machinery, etc) and to distribute the product were high and profitability, for the most part, iffy. You could know more than you could remember. Published information was well vetted, for the most part, prior to publication.

The Internet removed most of the costs of publishing and distribution – and a lot of the inherent controls. Search made information findable. Many new kinds of data (click-through statistics, for example) changed the dimensions of what could be known about customers and buying impulses. Thus, the domain of information consumption vastly changed.

The situation in medicine provides a specific and clear-cut analogy. We have gone from a discipline focused on symptoms, organs, and drugs to a much wider focus including genomes, proteins, and specific chemical interactions. This expanded focus has generated a need some new nouns and verbs to articulate the patterns derived from information. It is these patterns that turn once mysterious and intractable diseases into multiple levels of understandable behavior, which in turn, can lead to new treatments and cures.

A global company named JustSystems, with a North American headquarters in New York City, has leveraged the standardization of XML to provide tools to make the new breadth and enormous bulk of contemporary information more consumable. It has come to market with tools to support three approaches to information consumption. Its *XBRL* initiative focuses on the particularity of information – as in the data definitions of financial data, but extensible to other areas. This is like knowing the recipe of the chowder you are being asked to eat. Its support for *interactive documents* assures information currency. This is like knowing the chowder uses fresh ingredients. The *DITA*[1] approach deals with the composing content, which is like being able to design not just a chowder but also a whole meal. In each case, it is XML that provides the value; JustSystems provides the XML discipline to keep the value high. For more about JustSystems' solutions for you information indigestion, please read on.

---

[1] Darwin Information Typing Architecture.

The Nature of the Problem

Technology has endowed information with frameworks, like relational databases, and accelerated its use with workflows. Communications has let us to spread the word more immediately. Feeds and sensors have contributed more kinds of information. Scraping software lets us harvest information from websites – even competitors' websites. Search has helped us reach out to many of these information sources – but quickly we have found out that an undifferentiated, large quantity of information is not always better. In many cases, there is a last mile situation – a gap between the presentation of the information and its usability in analysis and decision-making.

What many people are looking for today is better methods of prequalification[2] of data to know that it is usable. For the congenially oriented, this may seem to point to the rating mechanisms found in social software, but the wisdom of the crowd does not always address the situation of the individual in an organization. The more pragmatic answer is some kind of system that would identify that information is current, valid, and trustworthy.

Operational clarity and organizational knowledge comes not only from shared information, but also from shared data definitions. It comes from use of information that is automatically up-to-date, in a way that the paradigm of paper cannot support. It comes from role-based aggregation of information. The enhancements that make information more usable are the basis of the following approaches.

JustSystems' Information Enhancement

Personalization – Honing Relevance to a Market of One

In its *Document-is-the-Application* approach, Just offers documents that render themselves based on the users' role, permissions, and the job to be done. Unlike portal presentations, this rendering persists as the document is printed, emailed, saved, or put into a records management program.

Users also may enter information into the document, which then is fed back to back-office systems. The document becomes the gateway to

business function – almost a user-side application that supports a business process.

Another, more prosaic but still important use is for essential business tools, such as manuals, that are only useful if they are current. Because dynamic documents do not have to be updated to stay current, a large organizational burden disappears. More importantly, a user will not need to worry about staleness of information, as one does about those leftovers in the back of the refrigerator.

With the expansion the business events now being monitored to include analysis of video surveillance and RFID tag capture, far more kinds of business monitoring and reporting are possible. All this monitoring has much more business value if alerts are propagated promptly to the relevant people and systems. Businesses that depend on a mobile workforce quickly see the virtues of interactive documents that feed current information to a mobile endpoint.

Specification – Using XBRL to Provide Information Particulars

In organizations and ecosystems that are international in scope, the incongruity of reporting methods and of the ways the information is derived muddy the views and analysis about the overall business. If you your concern is the definitions behind your financial data, XBRL[3] (eXtensible Business Reporting Language) – another XML standard – will help you. It can identify inconsistency in data definitions. It can also support automation of routine analysis of business financial operations by freeing information from the tables in which they have traditionally resided. It can also be used for textual information in Annual Reports and in footnote, to clarify what is being discussed.

XBRL is only as good as it is complete. The U.S. has finalized a complete XBRL taxonomy for its generally-accepted accounting principles (GAAP). Other countries are following suit on both taxonomies for their own GAAP versions and on bridges between the different versions. Each value in a financial report contains a mouse-over annotation of how (GAAP, non-GAAP, other) that value was derived.

This should bring a much-needed coherence to accounting of multi-national entities. It will give a much more complete and consistent view of the situation at hand. However, to work well,

---

[2] This is a different dilemma than that of *data integration* (which also involves metadata), which will not be covered in this bulletin.

[3] The Securities and Exchange Commission requires using XBRL for company filings, as of 2009.

like any other technical vocabulary, it must be well managed, and that management is what JustSystems provides with its *XMetaL* and *xfy* products. (See product section, which follows.)

While it is most obviously targeted at organizations, like stock exchanges and government agencies that must amass information from different organizations, many enterprises that have pursued an aggressive acquisition strategy have the same data definition challenges. XBRL may also provide a short-term alternative approach to a Master Data Management approach.[4] Where there are intransigent business unit incompatibilities, metadata tagging will at least keep the records clear while common definitions are being negotiated.

Content Aggregation – Pulling Information Together with DITA

The ideal for most information users is a clear path to the necessary information, at the required specificity or breadth, accessorized with the ability to also access related information. The ideal often falls short because the context of the information is not clear or the pieces needed are incomplete or out of date

DITA (the Darwin Information Typing Architecture) is an XML-based architecture for chunks of content. It has been vetted as a specification by OASIS, an industry standards organization. It allows information to be composed for the specific needs of the reader, just like the interactive documents explored earlier, but in larger chunks that can include both structured and unstructured content.

DITA is expressed most usefully as a *maturity model*. The real organizational value comes in supporting reuse of information that does not change, and updating of information that does. There is an order to the journey of adoption:

- The initial phase, level one, is to organize content into topics, regardless of where it lies. That gives immediate clarity. Then, XML tagging can make the content independent of the application and format in which it was generated.
- In the second phase, level two, the content is reorganized to fit the role-specific contexts in which it will be used. This often requires smal-

ler chunks.[5] DITA supports several technical niceties such as the ability to nest topics within a single XML statement, and other sophistications well beyond the realm of an interactive document.

- At level three, deliverables, such as reports (by which the content is consumed), are defined. They are instantiated as maps, or templates, that separate the consumption requirements (what parts are needed for completeness) from the content and relationship between the parts.
- At level four, content delivery is automated, often by some kind of content management system
- Level five involves semantic enrichment of content with data elements (the interactive document concept explored earlier in this report).
- Level six is the nirvana of a unified content (unstructured) and data (structured) strategy. No organization is at this level at present, but some are getting close.[6]

DITA is significant because it transforms what has often been an integrated, waterfall process into separate processes. This lets the structure of the deliverables be evolved independently of the content that will be aggregated in them. Thus, you can add complexity where it is needed without having to design the capacity for complexity in from the beginning. It is another example of a technology use where a just-in-time approach promotes flexibility. When you add in dynamic content and data, DITA becomes not just another content manager but also a higher value source of enlightenment.

JustSystems' Products

DITA is concerned with content, XBRL is concerned with structured financial information, and interactive documents can leverage either structured or unstructured information. XML is the basis in all cases. An overview of the tools that JustSystems offers will give another perspective on what XML can do for an organization. For more specific capabilities that each module of the tool supports, please visit the JustSystems website at JustSystems.com.

---

[4] For a description of Master Data Management as IBM sees it, see **The Clipper Group Navigator** entitled *Sharing Business Data Properly Requires Master Data Management Discipline*, dated January 17, 2008 and available at http://www.clipper.com/research/TEG2008004.pdf.

[5] If you think this sounds similar to the process of deconstructing massive legacy applications into addressable services, you are right.

[6] If you think this sounds like the *Lean Manufacturing*, which Toyota uses to build cars, you are right again.

XMetaL

XMetaL is authoring and content collaboration software. It supports creation of valid and consistent XML tagging using DITA or other architectures. It can be used standalone or integrated into publishing applications and content management systems, such as *Documentum*. It heavily leverages DITA's emerging capabilities – the latest version of XMetaL Author supports most DITA 1.1 elements as well as several PDF enhancements. XMetaL can be used by people with minimal knowledge of XML. XMetaL list price is $1195/user.

XMAX is an embeddable version of XMetaL that lets business people author XML content from within their applications. It is targeted at those who create information as a by-product. XMAX takes time out of processes such as product documentation. As with any process, the more close to the incident or source the reporting or documentation, the faster (and better) it gets done – but there needs to be some means of reconciling and managing the lifecycle of all the tags. XMetaL does this for structured content, supporting content lifecycle management, global content management, and master content management in the process.

XMetaL Connector lets you leverage content from disparate content management systems.

XMetaL Reviewer accelerates the review process. As a separately-priced server-side Web-based component, that enables both synchronous and asynchronous editing with email notifications with a link to the latest revision. It's pricing starts at $15,000/instance.

XMetaL Developer lets an IT administrator customize *XMetaL Author Enterprise* for the people who will use it – simplifying or enhancing it, as desired.

xfy

JustSystems defines *xfy* as a document-based composite application framework. It is the tool that allows the design and support of dynamic documents. It can be used without programming skills to drag and drop a feed of dynamic information into documents as desired. Business users can assemble information from diverse content and data sources

xfy Server is the foundation for managing xfy applications tracking information sharing and asset reuse, version control and distribution. It authenticates users and enforces restrictions and manages vocabulary manages component deliv-

ery. xfy is priced by server in small installations and by client in large installations. An initial project may run in the high five figures – a full installation, substantially more.

Pulling It All Together

Think of some of the new granular sources of business information, such as RFID data or video from surveillance of routine operations. This kind of information can be best leveraged as part of an aggregate. The provenance of that aggregate defines how it can be used. This calls for some kind of XBRL-like specification. It gains more value when aggregated with other pieces of information, which requires a DITA-like composition. To be usable as well as useful, it must be presented in a role-sensitive way – perhaps as an interactive document that, unlike a dashboard, can be saved periodically for retrospective use.

For this kind of scenario to work well, the XML metadata that facilitates this better use must be managed as a whole. If it is merely a mashup of odds and ends using metadata from the environments that created them, the contexts and definitions may not be congruent. This does not matter with Google maps, but it matters with business operations metrics.

With such centralized management, the use of information management tools of classification and tagging can be taken usefully further. Consider the dynamic nature of today's markets. The information that guides decisions will vary, depending on what is happening. During a market boom, you may be concerned with how to staff up to meet local demand. During a retrenchment, you may be more interested in broader operational analysis to best consolidate and pare down operations. The business activity monitoring also enhanced with XML could trigger presentation of interactive documents tailored for each situation. The pieces are all there. Think of how you could tailor information in ways that are particular to your business situation.

In information consumption, satisfaction is in the eye of the beholder. JustSystems' tools offer a variety of ways to make information more useful, more trustable, and more satisfying to the individuals who use it. Think of how your organization could use these tools to its advantage.

About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

➤ *The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.*

About the Author

Anne MacFarland is Director of Data Strategies and Information Solutions for The Clipper Group. Ms. MacFarland specializes in strategic business solutions offered by enterprise systems, software, and storage vendors, in trends in enterprise systems and networks, and in explaining these trends and the underlying technologies in simple business terms. She joined The Clipper Group after a long career in library systems, business archives, consulting, research, and freelance writing. Ms. MacFarland earned a Bachelor of Arts degree from Cornell University, where she was a College Scholar, and a Masters of Library Science from Southern Connecticut State University.

➤ *Reach Anne MacFarland via e-mail at Anne.MacFarland@clipper.com or at 781-235-0085 Ext. 128. (Please dial "128" when you hear the automated attendant.)*

Regarding Trademarks and Service Marks

The Clipper Group Navigator, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group** Captain's Log, Clipper Notes, and *"clipper.com"* are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and *"teraproductivity"* are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.