

OneSource Information Services

# External and Internal Content Integration

## Opportunities and Challenges

November 2002

By PHIL GARLICK

*Senior Vice President, OneSource Information Services*

### ***The Importance of Content Integration***

---

Any decision is only as good as the information behind it. Companies know that accurate information is their most important asset—capable of shedding vital insight on current customers, prospects, competitors, partners, and suppliers. However, such information is also volatile, highly perishable, often scarce, and costly: professionals who develop client relationships spend 16% of their workweek obtaining, reviewing, and evaluating information. One study of the US business community estimates the total annual cost at \$107 billion (*Outsell, Inc., 2002*).

Aside from the enormous cost of ad hoc research conducted by “information amateurs,” having the best-possible business intelligence available is crucial to a company’s competitive advantage. Educating sales teams on new industries, keeping relationship managers up-to-date on significant customer events, analyzing the financial health of prospects and suppliers, mapping corporate parents and subsidiaries, following changes in executive staffs—all these combine to keep every customer-facing part of the organization on top of its game.

Why, then, do corporations determined to arm their professionals with outstanding business intelligence often fail in this mission? As documented time and again, the vast majority of internal client information applications, such as customer portals and CRM installations, consistently fail to achieve their full potential. Why?

Their failure is not due to a lack of database skills or people or dollars. Instead, failure is usually due to:

- Insufficient in-house content expertise
- Inadequate business information taxonomy
- Poor or no tools

#### ***The Expertise Factor***

Content optimization and taxonomy development are specialized fields. Most internal corporate staff are not experienced enough to create a robust information solution without expert help.

#### ***The Taxonomy Factor***

The classification regime applied to your information can make all the difference.

Most taxonomy solutions being offered on the market have been focused only on *unstructured* (text) documents, like news files. They run deep on business and topic terms but don't perform well for tagging companies with a unique identifier – the key entity in a customer-focused application. They lack global breadth below the top few thousand companies and they almost all lack the hierarchical corporate family information necessary to identify and tag references to subsidiary operations of diverse multi-national corporations.

### ***The Tools Factor***

Tools and services to apply taxonomies have likewise focused on auto-classification of textual information, using advanced techniques that analyze the distribution or presence of key words and phrases. While textual classification approaches have become quite sophisticated, they have a significant weakness: they fall down when it comes to handling *structured* (database-driven) information.

Much critical business information (such as corporate profiles, key contacts, opportunity assessments, purchase history, and much more) resides in databases. In fact, the majority of client data needed to populate a CRM or customer portal is exactly this kind of structured database-driven information.

Integrating structured company data requires maintaining a controlled universe of key corporate data and deploying sophisticated matching algorithms. Such algorithms utilize data points like address, phone, and postal code and don't just rely on company name. And because no company limits its sales to publicly traded companies, common taxonomies like stock ticker don't cover the necessary universe of companies for customer-facing applications.

## ***What to Look for when Considering Content Integration***

---

As a first step, buyers should evaluate the entirety of their integration and classification needs. Due to the specialized skills required, most companies will then want to look for an expert partner who can deliver a complete solution.

Avoid "solutions" that run deep on business and topic terms but don't perform well for tagging companies – as we've said, this is the key entity in a customer-focused application. Evaluate the global breadth—does the solution go below the top few thousand companies? Very importantly, does the solution deliver hierarchical information on corporate families to capture the subsidiary operations of diverse multi-national corporations?

Because much of the critical information on clients and prospects resides in structured databases, probe the solution vendor's experience and sophistication in deploying sophisticated tagging, or matching, algorithms.

A complete content integration solution can handle all types of data, including client internal as well as external commercial data, held in a structured database or as unstructured text. All information on a company should be accessible from a single point in a customer portal or CRM system.

A solution should include:

- A broad, integrated taxonomy that covers companies, topics, industries/lines of business, and geographies equally well. The taxonomy should encompass both metadata for textual documents (e.g. call reports, news stories) and structured company data (e.g. corporate profiles and sales histories).

- Support for company aliases, former names, common identifiers (like ticker and CUSIP), and subsidiary names. The taxonomy should support mapping to a variety of commercial company identifiers, so you have the option to integrate multiple commercial sources.
- Complete integration with industry standards like SIC and ISO codes as well as other proprietary schemes. Any commercial taxonomy that ignores standard codes (perhaps because it was built for application to a particular vendor's data set) is not robust enough.
- Auto-classification tools to apply a taxonomy to unstructured text documents with a high degree of accuracy
- Automated matching tools to map structured company files to a common company taxonomy.
- Editorial review as well as automated matching and classification techniques, because no automated tool can perform matching to the high degree of certainty and completeness required for CRM or customer portals.
- An ongoing maintenance component to keep content up-to-date (e.g. name changes, effects of mergers and acquisitions, etc.)

The bottom line is that whatever content integration solution you choose, it is critical that you first understand your full scope of need before selecting a provider. To do this, ensure that your portal or CRM team includes someone skilled in content integration and knowledgeable about taxonomies and classification solutions.

## ***Content Integration Increases Value***

---

External business information can be a powerful new asset within your internal applications. Business Information services like OneSource aggregate raw content from many specialized sources to create a comprehensive, coherent information base. In fact, OneSource provides *company-linked* business information covering over 1.5 million large and mid-size corporations, including structured data such as corporate family structures, industries, geographies, executives, financials, and analyst reports, as well as newswire stories, trade articles and business press.

Most companies maintain internal data on their customers, prospects, partners and competitors. This information may include customer lists, call reports, purchase histories, financial data, competitor profiles and more. However, it is the combination of internal and external information that can inject real value into applications such as portals and CRM systems.

To take a simple example, internal information on a particular customer, such as location and sales history, can be augmented by a complete assessment including stock price over time, debt ratios, corporate developments, key financials, competitors, and a full industry profile including expert analyst reports. By combining inside information with external data, the value of both is markedly enhanced.

In addition, many companies want to upgrade the quality of their own internally held customer and prospect data before integrating it with external business information. Internal information can be radically improved by normalizing for consistency, filling gaps where possible (e.g. adding SIC or NAICS codes), deleting duplicate records, and applying a standard taxonomy to link relevant items.

The result is a new level of intelligent prospecting and relationship management, where preparation is better, close ratios are higher, risk is mitigated, due diligence is robust, insight is deeper, responses are faster, and confidence at all times is greater.

## ***Where to Begin***

---

Start by qualifying, or assessing the state of, your internal data. Data may reside in multiple repositories and have diverse formats and structures. Quality may be uneven because of incomplete, inaccurate or duplicative records. Your data should be compared with an extensive content database to identify gaps, duplicates and inconsistencies.

Once the current status of the data is known, the normalization phase standardizes analogous information fields across multiple internal databases. This includes mapping analogous fields from different databases to the same template, reconciling executive titles across an established set of job functions, and more.

Depending on the level of quality and consistency within your data, the next step is to fill information gaps where possible and eliminate extraneous information and duplicate records.

Applying an extensive taxonomy to your data creates an organized, interrelated dataset that can be explored and filtered quickly and accurately. Choose a taxonomy that allows you to integrate your data immediately with the external content by assigning unique identifiers to each company. This establishes a seamless link to the external information.

Ensure that you have a process in place to maintain the quality of your data as it changes over time.

When these steps are performed well, data is consistent and relevant across the board. At this point it becomes a straightforward technical project to integrate your internal and external information within the enterprise application.

This is where your information solution provider should offer a robust capability (such as an application programming interface) to easily call up specific information from the external information service for presentation from within the internal application. With proper content optimization and client data preparation, many companies can manage the integration work themselves. Still, make sure that your information service provider can offer all the necessary guidance, or a full-service integration solution, if desired.



For more information about Content Optimization Services, contact your OneSource sales representative or visit [www.onesource.com](http://www.onesource.com).