

## Dynamic Content Software Strategies Consulting Service

August 29, 2003

---

### Fueling Knowledge Discovery: Stratify Accelerates Towards the Oil and Gas Industry

#### Introduction

The categorization and taxonomy technology segment has been hit hard during the past few years. Although the economy is taking its toll on these vendors, a handful of them are not only surviving but even have plans in place for growth. Almost all vendors have been conducting battles within specific vertical industries that have been spending money on these technologies, most notably publishing and the federal government and to a lesser extent some manufacturers. Nevertheless, we still believe that this is not a niche problem, but is universal to all companies that create, manage, and use information.

At some point down the road, all organizations will be forced to wrestle with the problem of making it easier to find pertinent information in the ever-growing pool of content and documents. That is why we took note when Stratify, one of the vendors rising to the top, identified the chemical, oil, and gas (COG) industry as one where the company has seen success and has identified as a target market.

The chemical, oil, and gas industry generally consists of large, multi-national companies such as Schlumberger and BR Petrobras, two of Stratify's recent sales successes. These companies have large volumes of data representing multiple languages in a wide variety of formats. It is impossible for a layperson to even imagine the information that COG companies produce daily, but it must include e-mails from the field about exploration efforts, technical specifications for wells and other projects, as well as Web and other research. COG companies deal with regulatory requirements and also feel the pressures of a tight economy, with the driving need to better utilize existing knowledge.

An industry that is this complex demands a state-of-the-art knowledge management system that can organize data from around the globe. It must also be able to identify relationships and patterns in this data so employees can actually use it efficiently. Logically, this is a great fit for Stratify. With any luck, this industry is on the verge of tackling the information overload problem.

#### Corporate Overview

Recognized by publications such as Computerworld and Transform Magazine as a technology leader, Stratify brings data mining technology to unstructured data. Stratify's goal is to become a major player, rather than just a boutique provider, in the three verticals it targets (Media and News Aggregation, Government and Intelligence, and Oil & Gas).

Mountain View, California-based Stratify would seem to have its work cut out for

CAP Ventures is a strategic consulting firm for providers and users of business communication technologies and services. We deliver key research, analysis, forecasting, benchmarking, and strategy recommendations to make a competitive difference in our clients' businesses. Additional information about CAP Ventures is available on the Web at [www.capv.com](http://www.capv.com).

#### Headquarters

CAP Ventures, Inc.  
600 Cordwainer Drive  
Norwell, MA 02061  
Phone: 781 871 9000  
Fax: 781 871 3861  
E-mail: [info@capv.com](mailto:info@capv.com)  
[www.capv.com](http://www.capv.com)

#### Europe

**CAP Ventures, Ltd.**  
3<sup>rd</sup> Floor, Sceptre House  
7-9 Castle Street  
Luton, Bedfordshire,  
United Kingdom LU1 3AJ  
Phone: +44 1582 400120  
Fax: +44 1582 411001  
E-mail: [euro.info@capv.com](mailto:euro.info@capv.com)

#### Japan

**G.S.M. Corporation**  
Hiroo Office Building  
1-3-18 Hiroo  
Shibuya-ku  
Tokyo 150-0012 Japan  
Phone: +81 3 5475 2663  
Fax: +81 3 5475 2710  
E-mail: [info@gsm.to](mailto:info@gsm.to)

This Material is prepared specifically for clients of CAP Ventures. The opinions expressed represent our interpretation and analysis of information generally available to the public or released by responsible individuals in the subject companies. We believe that the sources of information on which our material is based are reliable and we have applied our best professional judgment to the data obtained.

it. Founded in September 1999, the company is privately held and has backing from many organizations including In-Q-Tel, the venture capital arm of the CIA. Founder Ramana Ventaka has years of experience in data-mining technologies dating back to his days at Intel's Microcomputer Lab. Stratify's earnings data is unavailable, but the company claims that it is a growing firm with 50 employees.

Stratify's flagship product is the Stratify Discovery System, for which Version 3.0 was released in July 2003. Stratify's Classification Server and its other lines augment the total Stratify solution.

**Table 1: Stratify Quick Facts**

Company:	Stratify
HQ:	Mountain View, CA
URL:	<a href="http://www.stratify.com">www.stratify.com</a>
Stock Symbol:	Privately Held
Tag Line:	"Discover More"
President & CEO:	Ramana Venkata
2002 Revenue:	\$4 Million (CAP Ventures Estimate)
Employees:	50
Customers:	10+
Products:	Stratify Discovery System, Stratify Classification Server, Stratify Analytics

## Products

Stratify's products tackle the difficult task of organizing unstructured information. This is a particularly daunting task within large companies. Once the information is organized, the Stratify system discovers new and emerging relationships as well as patterns in the data. This allows a business user to more efficiently use existing corporate information, from documents to e-mails to Web research.

Stratify Discovery System 3.0 can collect documents and information through a variety of sources, including e-mail repositories (Notes and Exchange), shared file systems, and the Internet and intranets through its crawler. Once collected, the Taxonomy Builder uses its patent pending clustering and pattern matching algorithm technology to organize these unstructured documents in an organized taxonomy. The user is also allowed to import existing, company standard taxonomies and definitions to leverage any existing or customized structure.

Through the Taxonomy Manager, users can have complete editorial control of the taxonomy building and classification process. The J2EE-compliant Data Hub works with the Metadata Server to deliver organized metadata to Stratify user interfaces such as Taxonomy Manager and Stratify Web Access, as well as many third-party systems. Put together, these allow for end-to-end taxonomy lifecycle management.

Through these features and technologies, the Stratify Discovery System provides the user with many options for use. The Discovery Wizard can automatically create personal taxonomies on the fly. For example, the system categorizes data on an information-intensive site like [www.capv.com](http://www.capv.com), organizing it into categories for easy perusal and efficient use. The system can do this for data in multiple languages and present them in the same browser-based display.

Discovery System takes this browsing one step further by providing customized data views, based on taxonomy facets or mapping criteria. Analogous to database views, these taxonomy views speed the access to relevant data and simplify product development.

Business users can take advantage of Stratify Analytics to select taxonomy items or entities and get a Quick Report. For example, if the user-selected taxonomy topic is "Boston Red Sox," a Quick Report can

list the most referenced people in documents about the Red Sox. A Quick Report for most popular topics on the same taxonomy would present the topics that appear in the greatest number of documents. For this example taxonomy, "Fenway Park" and "Nomar Garciaparra" would no doubt be high on the list.

Stratify Analytics also provides features such as a Heat Map and Network Graphs. A Heat Map presents taxonomy topics and entities (such as People, Locations, and Organizations) in a highly visual form. The relative size and location of topics on the screen relates information about those subjects that the user can then drill down on. The Network Graph feature shows, in diagram form, the connections between topics and entities. This is especially useful to uncovering hidden relationships between entities that involve "multiple degrees of separation." The system can automatically show the documents responsible for relationships between entities. Using the Red Sox example from above, a user could discover the shortest path between "Boston Red Sox" and "Dominican Republic" to be "Pedro Martinez," as well as alternative relationships connecting the two entities.

## CAP Ventures' Perspective

Stratify and some of the other leading vendors in the content analysis and enrichment segment have smartly identified that their selling proposition should not be to create taxonomies that are better than the other solutions in the market. Although this is a core requirement, and certainly a selling point for some of the targeted industries with solid cost savings related to it, larger value comes from helping users locate and use content. Facilitating information discovery and usage is the strongest motivator for deploying these types of technologies in knowledge-intensive industries.

The challenge for the vendors in this space, of course, is that all too often search technologies are perceived as the answer to the information discovery problem. Anyone who has ever used the Web, let alone dealt with trying to find corporate documents or contents, knows that searching is not an ideal solution in and of itself (especially when any significant volume of content is involved). Stratify does not provide searching, but something rather complementary to it (in fact, Stratify teams with FAST Search & Transfer when appropriate). The company leverages its categorization and taxonomy capabilities to provide context within visual aids to pinpointing appropriate content. It is, in a nutshell, key to any knowledge management effort.

In the Oil and Gas industry, Stratify has been able to sell its technology due to certain industry conditions that contribute to a knowledge management challenge, such as:

- The large size of companies and their business conditions lead to large amounts of unstructured data.
- The companies tend to be multi-national, creating the need for multi-lingual solutions.
- These companies tend to invest in infrastructure and see the need for a broad-based solution.
- The companies face regulatory compliance issues, and must respond with appropriate documentation when requested.

Not surprisingly, Stratify's message has resonated with at least a couple of firms in an industry that is facing this challenge. In both of those cases, projects carrying a knowledge management banner were underway. Stratify was selected along with document and content management solutions, enterprise information portals, and search technologies to form complete solutions. The fact that multiple technologies are integrated as part of the solution hints that this is not a small project for those organizations. Will others in this industry follow suit with similar projects? Our belief is that most large firms will eventually be forced to solve this problem. We would not be surprised to see this industry become one of the first to invest in a better way to drill into their content.

*Mike Maziarka, Jon Franke*