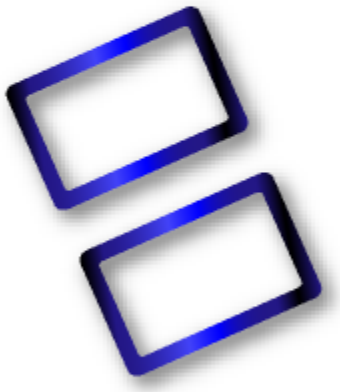
EXAMINING

Veraterm Technology



A Technology White Paper

www.veraterm.com

Veraterm Technology

The information contained in this document represent the current view of Data Management Associates, Inc on the topics discussed and on the date of publication.

Printed: June 2004 Mount Laurel, NJ USA

Company Information:

Data Management Associates, Inc.
Twin Ponds Executive Campus
503 Birchfield Drive
Mount Laurel, NJ 08054-4009
USA
www.veraterm.com

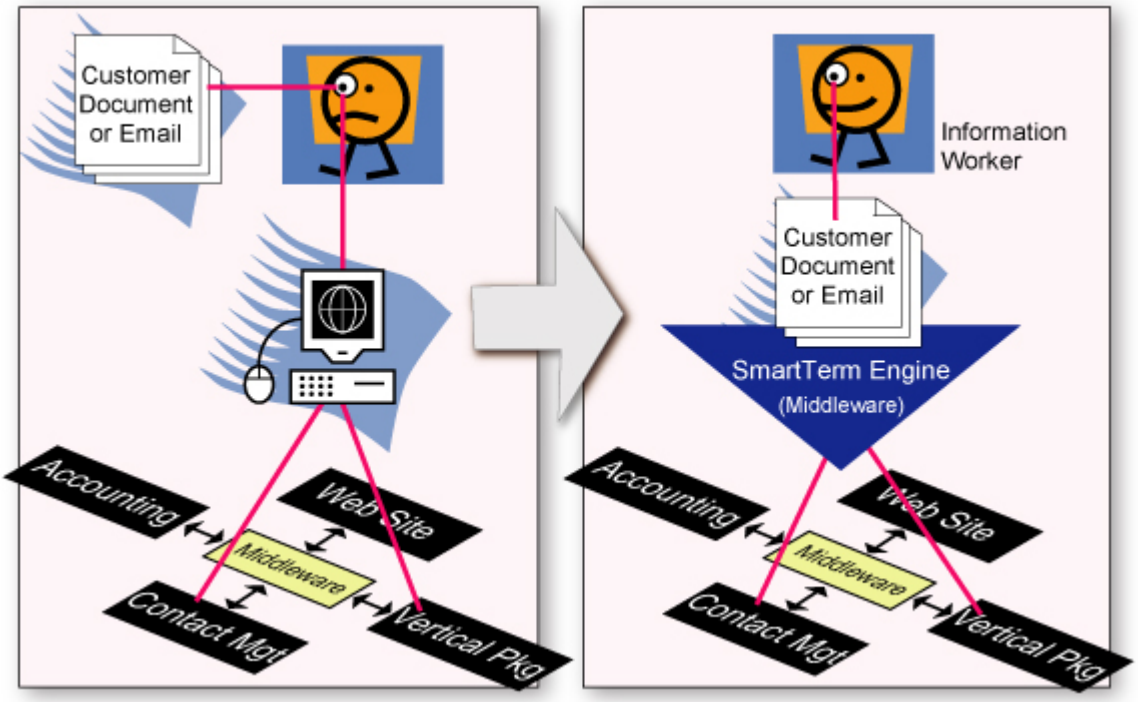
Veraterm Technology

Introducing Veraterm *Intelligent Integration at Your Fingertips!*

Veraterm "InfoTerms" are **intelligent**, context-sensitive links between text and database systems. InfoTerms are created and maintained by DMA's Veraterm software which evaluates text as it is entered in "real time" and **integrates** relevant data to the document workspace.

Veraterm is used to improve workflow efficiency for **information workers** who process and/or produce documents (emails, word-processing documents and spreadsheets) on a regular basis. Veraterm allows people to focus on their work (their document) rather than on the data-access procedures that accompany their work.

The diagram, below, illustrates the shift in focus that Veraterm provides. The mechanics of database access are performed *automatically* by the Veraterm products and the results presented to the user *inside their document!*



Why Veraterm?

Here are some of the **benefits** of using Veraterm solutions.

Flexible Database Connectivity	Veraterm can be configured to connect to **different** database systems. Information from **multiple** systems can be combined in a single view.
Two-Way Information Flow	Not only is information retrievable, but records (such as customers, notes and accounts) can be **added** and **edited** from a pop up window on top of the document workspace
Automatically Identify New Contacts or Customers	When Veraterm detects a person's name, the name is automatically searched in the specified system(s). Different options are presented to the user depending on whether the name exists in the data system or not.
Automatic Document Creation	Use Veraterm to detect **specific content** inside incoming emails (such as requests for information). Reply automatically by having Veraterm create documents or email using relevant information.
Save Mouse-Clicks, Time and Money	Information workers experience a more natural and streamlined pathway to information access while maintaining focus on their workspace. Return-On-Investment (ROI) is tangible and immediate when workflow automation saves processing time.

The Veraterm Technology Engine

The Veraterm Technology Engine lies at the heart of the Veraterm software infrastructure. It is included will all Veraterm products. The Veraterm Technology Engine is published in three different editions to accommodate different industry needs.

Personal Edition

For individual users and small organizations. In-document database access without a need for database integration. The Personal Edition allows action item customization. Configured with a single connector to a specific software application, such as ACT! or QuickBooks.

Professional Edition

For individual users and organizations with the need for database integration and customization Accepts multiple connectors to a various software applications and is extensible at every point.

Enterprise Edition

For organization with the need for enterprise-level features such as application-server integration, custom term-recognizer creation and group licensing.

This discussion examines Veraterm Technology as related to the Veraterm product line.

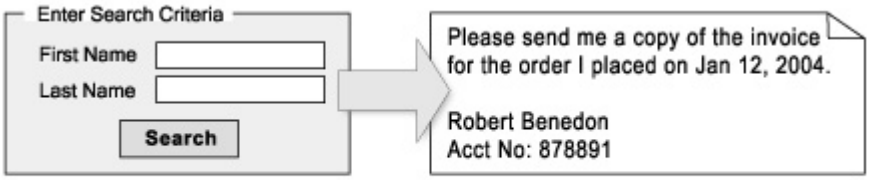
The Veraterm Concept	Using the document workspace as a user interface
Smart Tags	Microsoft Smart Tags are a window on InfoTerms
System Overview	The Veraterm Technology Engine is described
Veraterm Processes	Veraterm functionality described
System Architecture	Key components of Veraterm Technology

1. The Veraterm Concept

The Veraterm Concept is about using **document workspaces as user interfaces**, significantly improving access to information.

Why use documents as user interfaces? Because software developers and computers can not read peoples' minds!

The diagram, below, shows a simple database search window. Software waiting for a request cannot possibly know which customer is on the mind of the user until the Search button is clicked. However, software which is "reading" text does not need to read minds because it is seeing the same information that the user is seeing.



Typical database interface with searches initiated by the user

With the document as a user interface the Veraterm System initiates intelligent searches

Software should be able to pick out the name and account number from such text just as a user would do. What's more, software should be able to retrieve relevant information from a database before the user has even finished reading the document!

How might relevant information be presented to the user? The user may want to view demographic or account information, may want to edit the account or may want to embed customer information into a reply email. The user's typical choices depend on the context of the document. Since the Veraterm System uses the document as a user-interface, choices should be presented to the use **inside the document** so the user never leaves the document! The Veraterm Technology Engine (VTE) makes this possible. Choices are expressed to the user through the use of Smart Tags in Microsoft Office documents. Smart Tags provide users with the ability to associate text and data with actions.

Veraterm assists information workers by automating common and repetitive data access tasks. Veraterm combines intelligence with data integration technology. The VTE contains the infrastructure that is needed to build configurable and customizable time-saving solutions.

Business Perspective

- The Veraterm System adds value to existing business software. The Veraterm Technology Engine is not a replacement for current systems.
- Veraterm creates workflow automation, which provides quantifiable benefits for information workers.
- Veraterm integrates information from multiple places making the document a portal for enterprise information.

Technology Perspective

- The Veraterm System is configurable by users and technology partners, and extensible by developers and solution partners.
- The Veraterm System is a development platform for building solutions that benefit users directly.

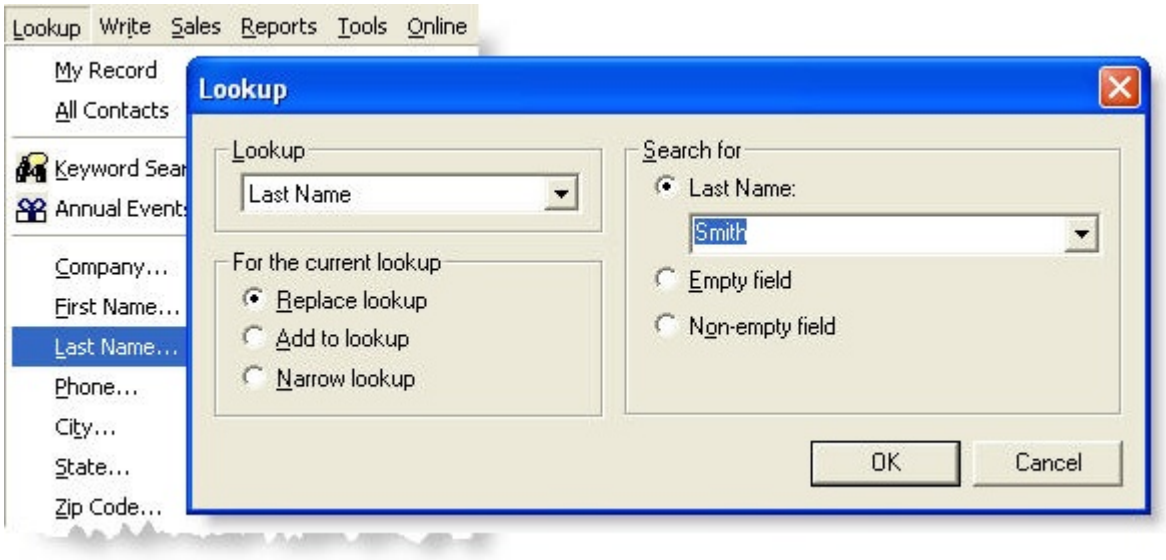
- Veraterm packages the infrastructure technology needed to build enterprise-level Smart Tag solutions, so developers can focus on the specific application.
- The Veraterm Technology Engine includes extensible components and functionality, such as:

  - Data Connectors
  - XML Mapping
  - Data Synchronization
  - Text Mining
  - Entity Containers
  - Term Lists
  - Smart Tag Interfacing
  - Context Intelligence
  - Connection Pooling
  - Logins and Security
  - Parsers and Preparsers
  - Term Recognizers
  - Expression Editors
  - Local Caching and Buffering
  - Personalized Views
  - 3rd Party Interfaces

Data Recognition

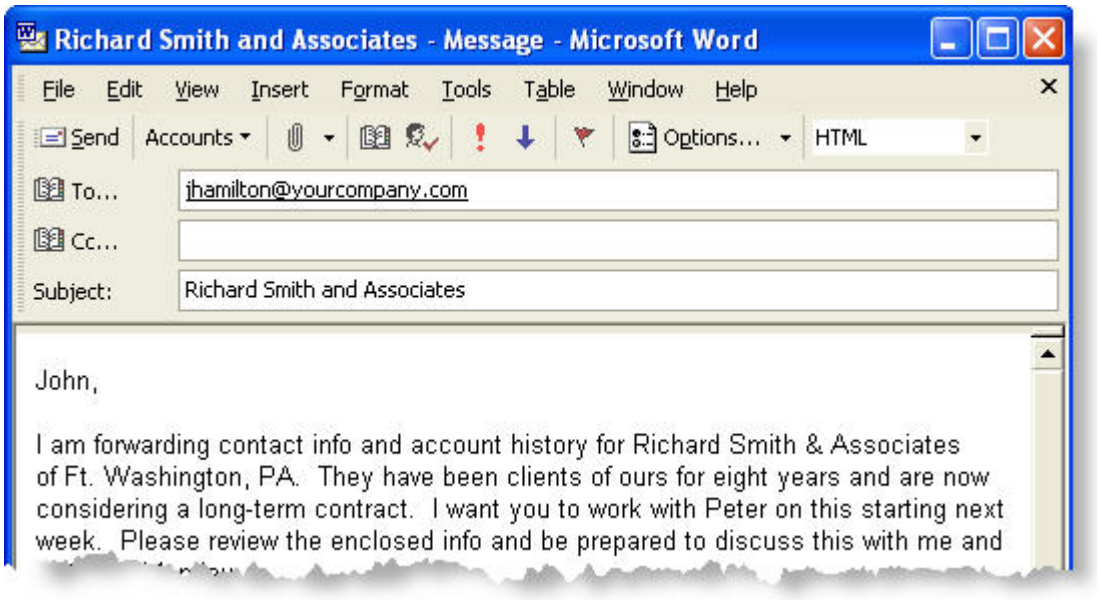
Recognition is the general term used to describe the automatic access of information. Below are examples which illustrate common data access starting points. In the next section the **action** side (the use of the information) is discussed.

Below is a typical database search window. It offers flexible search options and is a necessary tool for navigating through a large database application. Veraterm does not eliminate the need for such windows, but provides a document-based alternative. Where Veraterm can be used, it eliminates the need to leave the document, open search windows and manually initiate a lookup.



Outgoing Documents

Consider the case of a short communication which requires that additional information be inserted...

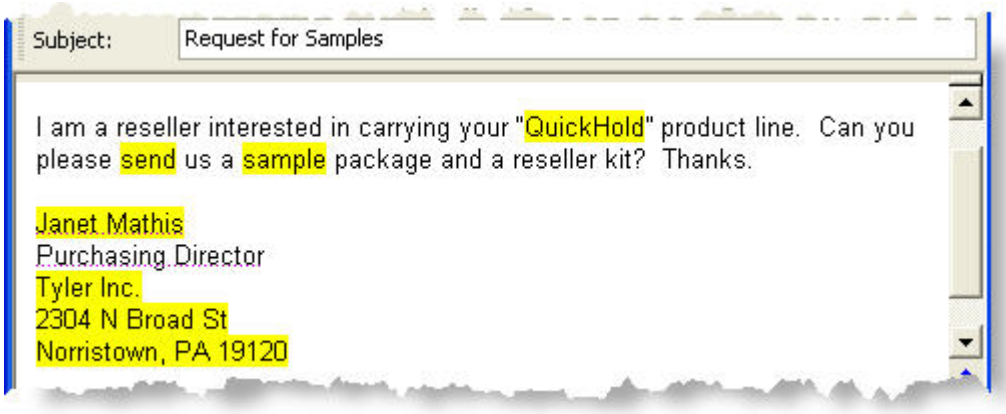


For this example, suppose the recipient (John) works off site or does not have direct access to client information. Or suppose that the information the author wants to send is from the note/history records inside his personal contact management system. There are many reasons why database information such as this gets embedded into documents.

The Veraterm System can automatically associate the name that it detects ("Richard Smith" or "Richard Smith & Associates") with a database entry. After doing so it presents the user with useful options which can include the insertion of relevant information into the current document. By giving the author a **direct link** to the client record, the author is spared the need to manually search for, copy and paste the client's information. The Veraterm System operates automatically while the author is typing and provides these benefits in "real-time."

Incoming Documents

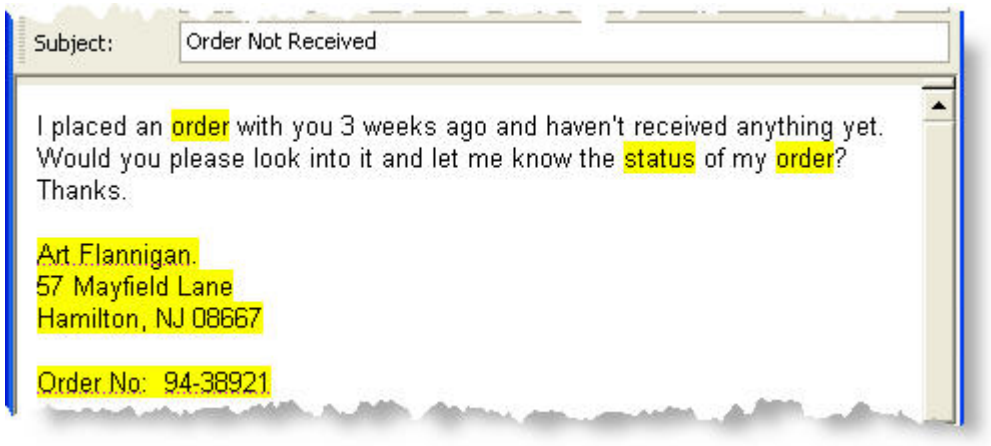
Incoming documents are essentially "existing" documents that the user opens as a recipient. The Veraterm System analyses documents when they are opened, or when they arrive in an email in-box. Clues are derived from the context of the email. In the following examples, the yellow highlights indicate important information to the Veraterm System.



The Veraterm System recognizes the above document as a request for information by a particular person or company. This knowledge can be used to automatically perform actions such as:

- Access purchase history if this is an existing customer
- Add them as a new prospect or customer
- Compose a reply email
- Create a purchase order for free samples in an order-processing system

Similarly, the Veraterm System identifies the customer and the specific order from the communication below, eliminating the need to manually find the customer account or the original order.



2. Smart Tags

Veraterm brings workflow integration to Microsoft Office productivity tools (Word, Excel and Outlook). Veraterm's "InfoTerms" are expressed to the user through the use of Smart Tags. Smart Tags provide users with the ability to associate text and data with actions.

Smart Tags

If you perform a Google search for the phrase "What are Smart Tags?" or "What is a Smart Tag?" you will read:

- Smart Tags are dotted lines under words in MS Office products, or
- Smart Tags are small pop-up icons that quickly enable user access to content-specific features with Office Apps, or
- A Smart Tag is data that is recognized and labeled so that it can be used to perform actions that you would normally open other applications to perform, or
- Smart Tags are user interface icons that dynamically recognize text and present the users with options based on that text

All of these are accurate and they concentrate on different aspects of Smart Tag technology - the underline, the icon, the data, the text and the actions. For those not familiar with Microsoft Smart Tags, here is what they look like. This example shows a Microsoft Outlook Contact Smart Tag.

In the Veraterm System, Smart Tags are windows exposing InfoTerms. Microsoft's Smart Tag technology creates external access to the words in a document, allowing the creation of customized links to other resources. Microsoft uses dotted underlines and drop-down menus to give the user a familiar way of doing this. Before the dotted underline (the visible Smart Tag) was created, however, the host application (ie: Word or Excel) sent the edited text to a "Smart Tag Recognizer" which had a chance to evaluate the text. When a contact was found, the recognizer was able to tell the host application to create the Smart Tag.

The Smart Tag technology built into Office applications provides:

- Notification that text has been edited
- The visible icon and drop-down menu...

...But not much more! Microsoft has created a valuable "action-in-context" mechanism which developers can use to create custom Smart Tag recognizers and actions. Microsoft has **not** provided all the tools and technology necessary to create an enterprise-level Smart Tag recognizer solution. That is where Veraterm comes in!

To a Smart Tag developer, the Veraterm **Technology Engine is a full-featured Smart Tag technology infrastructure that handles the underlying subsystems.** The result is that Smart Tag solutions can be created at a reasonable cost without the need for complicated system-level programming (parsing, caching, application monitoring, content analysis, data access and the like).

The native Smart Tag recognizer interface is best suited for short lists and simple in-memory list lookups, and is not well-suited for enterprise database lookups. The reasons include:

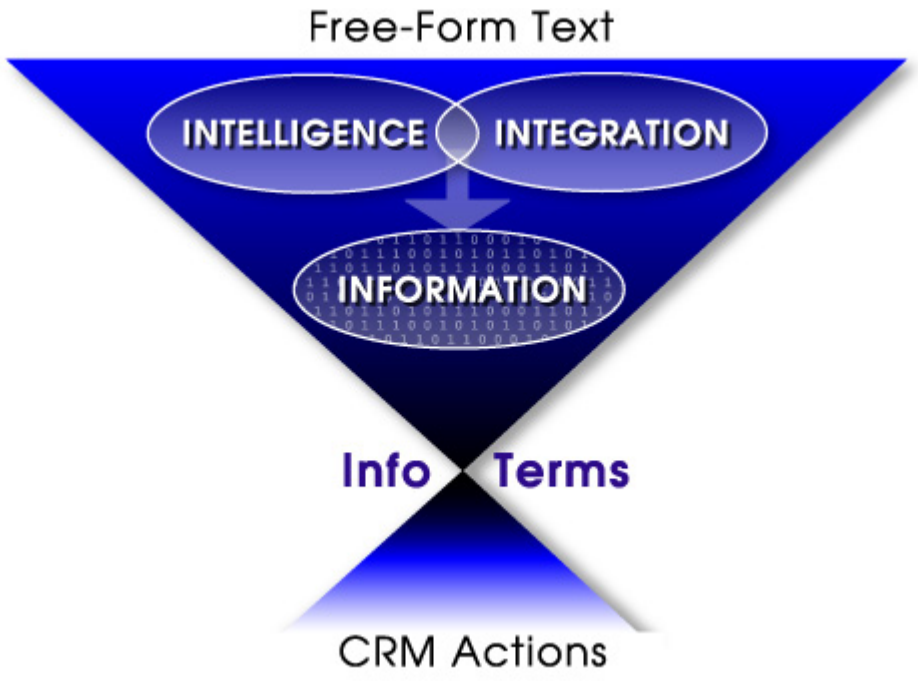
- No preparsers to identify potential person/company names, etc.
- Too many calls to the recognizer with the same text. So many, that it would freeze a workstation to query a database each time.
- No mechanism for local buffering of search results and data to minimize network traffic and to maximize speed.

Why Use Smart Tags?

Microsoft Office Smart Tags are a significant user-interface advancement that can be utilized to advance "workflow integration." Microsoft Office productivity tools are almost ubiquitous, and Smart Tags offer a powerful means for individuals and organizations to be better informed and more responsive, based on the context of their documents. Veraterm extends the functionality of the native Smart Tag interface and provides the tools necessary to create enterprise-level Smart Tag solutions.

3. System Overview

Veraterm Technology, at the highest level of abstraction, is the combination of **Intelligence** and **Integration** resulting in **Relevant Information** expressed as InfoTerms. In document workspaces Veraterm Technology is applied to **free-form text** to create InfoTerms for workflow automation.



Free-Form Text means:

- Text that is not labeled, tagged or necessarily organized in cells or fields
- Words without context or declared meaning
- For example: words in an email body, word processing document, spreadsheet cell (unlabeled)
- The operation of the Veraterm System does **not** depend on document templates or any specific formatting

Intelligence is applied, in the form of:

- Identification of relevant terms through the use of parsers and text mining techniques
- Knowledge of the presence or absence of a term in a database application
- Ability to share/relate information between terms
- Knowledge of how to store and retrieve information related to the terms
- Knowledge of actions already taken on a term

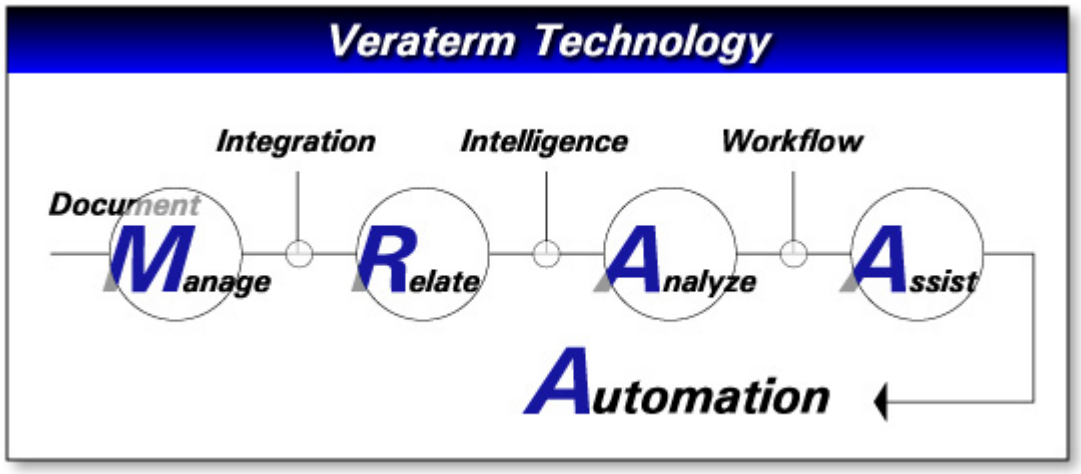
Integration principles are applied, in the form of:

- Connectivity to one or more database applications per term
- Automatic information retrieval ("push" vs. "pull")
- Real-time processing "as you type"
- Flexible security module
- "Workflow Integration" - Integration for the purpose of workflow automation

This process is brought to life by the Veraterm Technology Engine, which manages the process and provides support to database applications, users and developers.

4. Veraterm Processes

The Veraterm Technology Engine manages a set of processes behind documents in a host application. The diagram, below, depicts a series of process categories that represent functions inside the Veraterm Technology Engine. The process flow is called "MRAA" for the four main process categories.



Manage. The Veraterm System manages the process by:

- Monitoring the host application and its collection of documents
- Detecting content changes. For Microsoft Word, which does not have the equivalent of a "paragraph edited" notification event, this includes an object model that was built around MS Word. Word's native Smart Tag interface is not sufficient for this purpose.
- Parsing - Applying text mining processes to identify "interesting" text (ie: names that should be looked up)
- Running the "Term Recognizer" (TR) system. TRs are the components which are configured to make specific relations between text and data.
- Controlling the security mechanism and the pooled database connections

Relate. Term Recognizers orchestrate the integration process. TRs use:

- Data Access objects managed by the Veraterm System.
- Term Lists for text comparisons and for system metadata
- The Search Term buffer for recent query results
- Entity Containers for configurable XML data storage
- Data Atlases, Data Maps and Data Links (all configurable) which hold metadata about the structure of entities, their database systems and the relationships between them. This sub-system is responsible for the **updatable** composite views involving two-or-more applications and the **integration** between them.

Analyze. Term Recognizers also make use of Veraterm Intelligence tools:

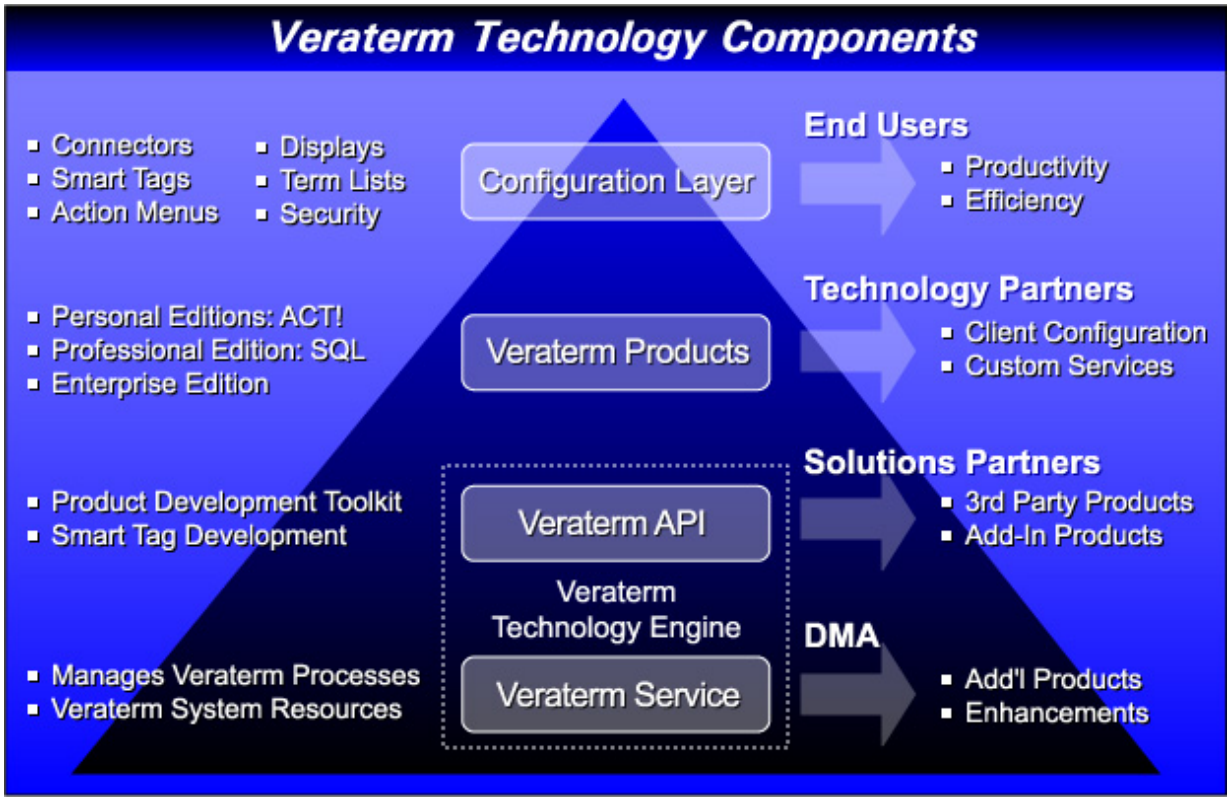
- The Veraterm System manages a set of "clues" or "facts" related to a document. For example, clues are used to determine that an email is actually a "request for information."
- Veraterm Events. An event subsystem is managed for the benefit of the Term Recognizers (TRs). TRs thereby have knowledge of the content and context of the whole document!
- A practical example of applying this technology is the resolution of duplicate names. The Veraterm System continues to look for clues which could be used to tell them apart - even if the clue is typed later. If an ID number, email, address, etc is found, the TRs are informed of this and the duplicate is resolved.

Assist. The goal of the Veraterm System is to assist the user in a significant way to improve workflow efficiency through automation. The system uses additional tools and processes to do this even after an InfoTerm has been created.

- Entity Containers load data in the background and can store the data locally for instantaneous access. This is beneficial when a full view of a customer is comprised of information from multiple systems including internet-based data sources.
- Entities (such as customers, accounts, orders, etc) have knowledge of how to display themselves. This makes the creation of Smart Tag actions much easier. The Veraterm System provides both the data and an internal display mechanism
- Entities know how to update themselves. If permissible, the Entity System can send inserts, updates and deletes back to the host application, making Veraterm Integration a 2-way link. The author of a Smart Tag action class does not need to be concerned with this level of infrastructure.
- Veraterm Actions. The Veraterm System does not only create Smart Tags as its only output mechanism. The system can be configured to initiate any number of processes based on the business logic contained in the Term Recognizers.

5. System Architecture

The Veraterm Technology components form four layers as shown in the diagram, below. The core of the Veraterm System, the "Veraterm Technology Engine" (VTE) is the combination of the service with the development toolkit and API (Application programmers' Interface). The VTE is included with all InfoTerm products and solutions.



Beginning with the foundation components and moving up, each layer is further described:

Veraterm Service

- Implemented as a standalone process
- Services multiple host applications and Veraterm Toolkit/API objects
- Manages most of the functionality described in the previous section (see MRAA diagram)
- Manages system resources such as Logins, Connections, Parsers, Term Recognizers, Metadata Objects and Entity Containers
- Accessible only to the Veraterm Toolkit/API. No external access allowed.
- The service component will be used by DMA for product enhancement and for additional products that utilize Veraterm Technology

Veraterm Toolkit and API

- Implemented as an apartment-threaded DLL. All public classes are documented in the Developer Reference Manual. This is the "application programmers' interface" (API) to the Veraterm System. At last count it contained almost 100 public classes.
- Services all functional aspects of Veraterm and Smart Tag solutions with important tools and processes.
- To Smart Tag developers, this layer is a Smart Tag Development Toolkit.
- Solutions Partners use this layer to develop 3rd party products and Veraterm Add-Ins.

Veraterm Products

- Using the VTE as a common foundation, DMA has packaged the product with different combinations of Connectors and functionality to service the needs of different markets.
- **Personal Edition**: For single users integrating with a single database application.
- **Professional Edition**: Full-featured product with the ability to connect to multiple database applications. Supports multi-user configurations.
- **Enterprise Edition**: For organizations needing enterprise-level features such as application-server integration, custom term-recognizer creation and group licensing.
- Technology Partners provide support, configuration, training and custom programming services on the Veraterm product platform.

Configuration Layer

- Most tools and processes in the system are **configurable** and **extensible**.
- The "System Manager" is a standalone application with configuration windows for all aspects of the Veraterm System
- Users have direct access to such tools as: Term Lists, Connection Parameters, Action Item Menus and Display Properties
- Users can install additional 3rd party Add-Ins, Connectors, Displays, Term Recognizers and Actions

Extensibility

The architecture of the Veraterm System is open, extensible and flexible. Veraterm integration has an unlimited number of potential uses across all vertical markets. Partners and third party developers require the ability to access its core components. Extensibility is supported in different ways:

- **Public Interfaces.** The interfaces and the API for all core objects are made public.
- **XML.** Internally, core objects are serialized and cached as XML strings. In the object model, all core objects have custom properties which can be used to add user-defined XML nodes. By this mechanism the Veraterm System preserves 3rd-party information attached to all core objects.
- **By Design.** The System Manager allows for the addition of user-defined objects in just about every category.
- **By Policy.** If there is a category of extensibility that DMA has overlooked and it becomes important to the developer community, then DMA is likely to accommodate the developers. Our aim is to promote the use of Veraterm Integration and Information Management.