



Overview of Thunderstone's TEXIS

Executive Summary

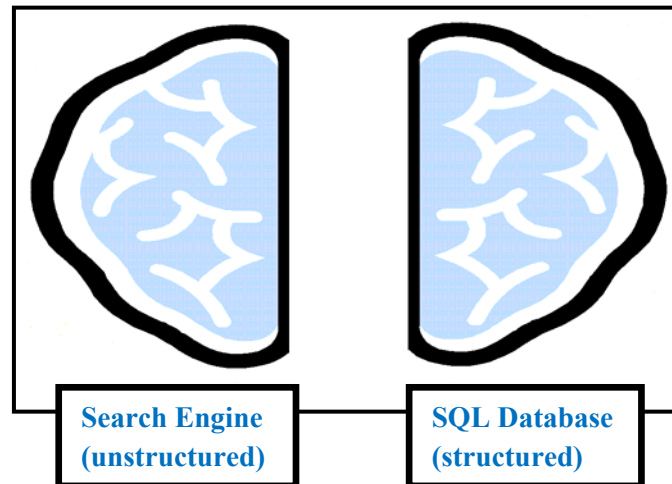
- TEXIS is the only fully integrated SQL RDBMS optimized for full-text search.
- TEXIS has high-performance ability to intelligently query and manage databases containing natural language text, numeric values, standard data types, geographic information, images, video, audio and other payload data.
- TEXIS powers real-time applications with zero-latency data insertion, providing immediate search availability of key data without waiting for scheduled index updates.
- TEXIS efficiently sorts and groups search results by any field(s) in the data. It can quickly sort tens of thousands of hits or more.
- TEXIS, the innovative development platform behind Thunderstone's entire line of enterprise search products, lets users and developers incorporate their own unique knowledge and expertise into customized search solutions that easily integrate with other applications.

What makes Taxis different from other search engines and databases?

Thunderstone's Taxis is the only search engine developed from the ground up as a fully integrated SQL RDBMS optimized for full-text search, and it's the only relational database that can store and search text documents of unlimited size within standard database tables.

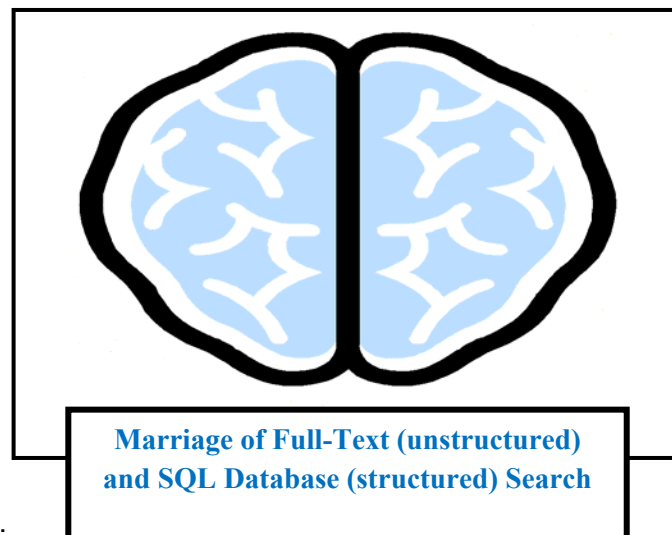
Used by hundreds of thousands of database application developers around the world, Structured Query Language provides many advantages for satisfying complicated search requirements. SQL also holds great promise as a reliable, well-defined path for implementing unanticipated new search functionality in the future. All other search engines offer a much narrower range of possibilities based on proprietary interfaces.

While typical search engines do a nice job of searching unstructured text and traditional databases have an impressive ability to handle queries on fielded or structured data, text searching and relational database management each rely upon radically different paradigms for organizing and retrieving information. They both developed and matured over decades as completely separate technologies, and they don't "marry" easily.



Thunderstone is the only company that has accomplished the true marriage of a full-text search engine and a powerful SQL relational database in a single platform. Addressing this challenge, the simultaneous searching of structured and unstructured data, remains one of Thunderstone's core competencies.

Deep in the heart of the Taxis RDBMS resides Thunderstone's Metamorph, a concept-based natural language search engine utilizing advanced lexical set logic.



Metamorph has often been classified as a form of Artificial Intelligence, since its functions fall into the categories of knowledge acquisition, natural language processing and intelligent text retrieval. The software attempts in its own way to understand your search queries, to represent

its understanding to the data in the files and to come up with relevant responses as retrieved portions of full-text information which best correspond to your submitted queries.

Metamorph's starting vocabulary has 250,000+ word connections, constructed in a dense web of associations and equivalences. Search parameters can be adjusted to dynamically dictate surface and deep inference. The program's responses can be controlled so that they are direct or abstract in relation to user queries. Proximity of concept can be fine tuned so as to qualify degree of relevance, providing matches which are sometimes concrete, sometimes abstract, as desired.

Metamorph allows for editing word sets. This means that you may select which associations you would like in connection to any search. You can create your own concept sets permanently for future use. You can fine tune the search to use associations of only a certain part of speech. You can enter all known spelling variations of any particular search word in the same way. **You can generally customize the program to include your own nomenclature and vocabulary, making it increasingly intelligent the longer it is in use.** When you want to control exactly what associations are made with any or all of the words or expressions in your searches, you can do so by editing the equivalence set associated with any word already known by the Equivalence File or by creating associations for a new or created word not yet known.

You can call up the ApproXimate Pattern Matcher (XPM) and tell it to look for a certain percentage of proximity to an entered string, finding misspelled names and typos. You can also look for numeric quantities entered as text, thanks to the Numeric Pattern Matcher (NPM) which recognizes that "four score and seven" is the same amount as "87."

Metamorph allows users to search for intersections of sets of lexical items, while also performing prefix and suffix morpheme processing. Users can specify, right in their queries, the delimiters of choice: i.e., they can look for lexical intersections within a sentence, a paragraph, a page, a designated amount of text or some other defined textual unit such as a memo.

Texis, with Metamorph inside it, provides a modular set of tools to attack the formidable problem of how to get at and deal with a large volume of information when you don't really know precisely what you need or where to find it. Thunderstone's Texis gives you the power, speed and flexibility to rapidly implement a customized search solution that will accomplish your data access/retrieval objectives in the most dynamic, efficient and pragmatic way possible.

When you want to control exactly what associations are made with any or all of the words or expressions in your searches, you can do so by editing the equivalence set associated with any word already known by the Equivalence File or by creating associations for a new or created word not yet known.

Thunderstone's Taxis has a number of characteristics and built-in advantages that differentiate it from other search solutions:

1. A fully integrated SQL database management system (DBMS) that follows the relational database model, Taxis is optimized for addressing the inclusion of unlimited quantities of narrative full text. It provides a method for managing and manipulating an organization's shared data, where intelligent text retrieval is harnessed as a qualifying action for selecting the desired information. Taxis simultaneously provides full-text, fielded and Boolean searching of both structured and unstructured content.
2. Taxis powers real-time applications with zero-latency data insertion, providing immediate search availability of new data without waiting for scheduled index updates. Unlike other search tools, Taxis ensures that all information which has been added to any table can be searched immediately—regardless of whether the table has been indexed and regardless of whether it has been suggested that an index be maintained on that table or not. Sequential table space scans and index-based scans are efficiently managed by Taxis so that the database can always be searched in the most optimized manner with the most current information available to the user.
3. Taxis enables searchers to sort and group query results by any field(s) in the data. And Taxis can quickly sort tens of thousands of hits or more. Other search tools either bog down sorting more than a few hundred items or else their sorting features are much more limited than the capabilities of Taxis.
4. Taxis allows you to treat the concatenation of any number of text fields as a single “virtual field.” As a single field you can create an index on the fields, search the fields and perform any other operation allowable on a field.
5. Taxis has high-performance ability to intelligently query and manage databases containing natural language text, numeric values, standard data types, geographic information, images, video, audio and other payload data. While Taxis excels at purposeful manipulation of textual information, it also performs useful mathematical operations on your data. You can construct queries that combine calculated values with text search.
6. Taxis lets you create an unlimited number of independent search collections—each with their own unique data types, fields, attributes or parameters. It also empowers users to submit queries to multiple search engines and/or multiple collections and have the results displayed together or combined.
7. Taxis gives developers prototype-friendly customization tools, extreme flexibility, rapid deployments and a feature-rich API. It supports multiple Search User Interfaces that offer specially-defined views of query input and results for different audience types or even for each unique individual. Thunderstone's Taxis imposes no user interface requirements. Taxis Web Script (Vortex) maintains “neutrality” with regard to whatever HTML markup (or JavaScript or other user interface technology) is employed for the user results presentation.

Which enterprise search applications require the robustness and flexibility of Taxis?

Taxis is the premier solution when large-scale, mission-critical and/or complex information retrieval challenges call for full-text searching tightly integrated with traditional structured database querying.

Businesses, governments, NGOs and educational institutions use Taxis in a wide range of applications such as online catalogs, auctions, classifieds, automated categorization, litigation support, intelligence collection/analysis, risk assessment, quality control, CRM, knowledge discovery, document and multimedia management, internet publishing, vertical portals, real-time message handling, web searching and many others.

Taxis powers the search on many high-traffic, heavy-content Ecommerce sites.

Sellers with extremely large numbers of customers and products have deployed Taxis-based online catalogs containing tens of millions of SKUs.

Taxis enables users to quickly search databases, XML repositories, HTML pages, PDF files, Flash/JavaScript content (including links) and other relevant sources.

With Taxis you can track user context, capture sessions for data mining and analysis, control browser cookies, set up log-in and security features, validate and modify user or data input according to custom-configured rules, create special interfaces, plus much more.

Thunderstone's Taxis provides the ideal development platform for rapidly deployable, custom-designed applications that require both unstructured and structured types of searching:

- **Online catalogs** contain unstructured text (product name, description, etc.) and structured content (style/size, price, in-stock availability, etc.) Users expect the ability to search by item description, to navigate by price range or to do both in combination.
- **Knowledge management systems** demand very efficient and secure enterprise-wide information retrieval across multiple repositories that serve different types of users, who all want dynamic, context-sensitive views of defined content (structured data) with the ability to refine results through full-text searching (unstructured data).
- **A Thunderstone solution provider customer** has deployed Taxis in a "brute force" full-text search scenario for its DoD Intelligence Community customer, using Thunderstone's Taxis to search the contents of a massive Oracle database in a counter-terrorism effort. Taxis is being used as an adjunct to Oracle full-text search because of its ability to scale while still providing superior performance in both rate of ingestion as well as search. Thunderstone's Taxis enables this customer to search across a 20 terabyte index, ingesting 70-80 million new records per hour and returning typical search results in < 10 seconds.
- **A Fortune 20 customer** is using Thunderstone's Taxis as the search platform for what they describe as the "single largest knowledge management system currently deployed at any corporation in the world." The application encompasses knowledge, people and processes, and it is used globally within the organization to access more than 30 terabytes of data. Users access the application 20+ million times per day, retrieving and sharing information from across the global enterprise. The application is the most-used corporate I.T. resource after e-mail.

Thunderstone's Taxis lets users and developers incorporate their own unique knowledge and expertise into customized search solutions that easily integrate with other applications. For additional information call +1 216 820 2200 or visit us online at <http://www.thunderstone.com>.