



Enterprise Search for Business Transaction Data

*Moving Towards the Definition of a
New, Disruptive Market Category for
Access to Structured Enterprise Data*

January 2006

“Everything has changed but our way of thinking.”

-- Albert Einstein

Many Times, The Business Problem You Think Is A ‘Database Problem’ Is Really A Search Problem.

Conventional wisdom holds that data management business problems are solved by databases. While this may sound obvious, it is no longer true...and adherence to outdated ideas is costing business billions in unnecessary expenses and lost revenue-generating opportunities. Many data management business problems actually have nothing to do with databases, and can be solved for 10% of the cost of a database, or less, while giving enterprises richer strategic firepower than they can currently imagine.

Imagine the competitive advantages of unlimited, immediate access to a complete record of business transactions reaching as far back as a company wants to go, without the usual cost and complexity associated with a database. To cite just one example, Marketing and Sales departments could capitalize on instant access to each customer’s purchase history the moment that customer walks into a store, or contacts a call center, website, or other touch point.

Imagine the capabilities of business intelligence tools fast enough to get inside the daily decision cycle armed with the right business data – buried inside months or years of stored history among billions of other transactions – that captures as much of the past as desired, along with transactions that happened just minutes ago.

All of this, and more, is possible simply by changing one’s way of thinking, shaking off obsolete data management assumptions to consider the benefits of enterprise search for business transaction data. This new way to index, search, and retrieve unchanging, structured data meets the definition of a truly disruptive technology – delivering an order-of-magnitude improvement in performance with an order-of-magnitude reduction in cost.

The first step is for enterprises to understand that many so-called data management business problems are actually much simpler data access problems that can be solved using cost-effective, non-invasive search technology. Once that insight is achieved, enterprises can, for the first time, store all their unchanging, historic data; keep it indefinitely; and access it immediately for a variety of strategic purposes...for a fraction of the cost of database solutions.

CopperEye believes recent technological innovations have primed the data management market for this new category of data access solution – augmenting, not replacing, relational databases in enterprise data management strategies – that is uniquely suited to the real-world demands of business and business intelligence. Contrary to what IT has been taught over the past 20 years, all data must not, of necessity, be managed by costly relational database management systems (RDBMS).

Many business problems can be solved simply by retrieving historical, unchanging transaction data when needed. Putting historical transaction data in a database against the future need to retrieve it is overkill – like using a

gold-plated Swiss Army knife when a toothpick would do. Databases are designed with sophisticated and complex features for managing hundreds or thousands of concurrent users and rapidly-changing dynamic data, and their costs reflect those capabilities. Considering the costs of hardware, database and application software licenses, and the administrators needed to manage terabyte-sized databases, storing structured, unchanging transaction data in a database can be 15 to 50 times more expensive than storing the same data in a flat file, according to Steve Duplessie, founder of the Enterprise Strategy Group analyst firm. An ideal data access solution for historic transaction data, therefore, would combine the swift retrieval capabilities of a database with the low cost of file storage.

What's Search Got To Do With It?

A primary purpose of any data management solution is to give business access to the exact data it needs as quickly as possible, and cost-effectively. To accomplish this in a business context, an ideal solution indexes the data without the need to move, load, or otherwise transform it, and then delivers precise results to highly selective queries – classic characteristics of a search solution.

Despite the similarity in name, all “searches” are not created equal. Google, Yahoo, and other popular consumer search engines focus on unstructured data, such as documents and Web pages, to return numerous, ambiguous results that users can peruse at their leisure. Business transaction search, on the other hand, leverages unique indexing technology to access enormous volumes of structured, unchanging, historic enterprise transaction data, providing business users and business intelligence applications precise, immediate results that would otherwise require the cost and complexity of a database.

Business transaction search has the ability to make monumental problems trivial by approaching data access as a search problem, not a database problem. To return for a moment to the most familiar form of search engine, it would be impossible for Google to put all four billion Web pages it searches into a database. The same holds true for business transaction data that in many enterprises is five to 10 times more extensive than Google's four billion Web pages. If a company is already storing its data on its network, and only needs to access the data when needed, there is no reason to put that data into a database when search can do the job for a small fraction of the cost.

Business Transaction Search – The 100% Solution To The 45% Problem

“Business transaction data” is the gold ore of a business. Whether generated by stock trades, phone calls, point of sale transactions, Web purchases, text messages, ATM and credit card activities, vendor interaction, or RFID transmissions, business transaction data contains everything there is to know about a company's customers, operations, vendors, partners, and market.¹

¹ (As used in this paper, “business transaction data” is a subset of “business event data,” a more inclusive term encompassing data that is collected but is unrelated to business transaction activity, such as application log files, user logon/logoff, etc.)

Concurrent trends are driving enterprises to make business transaction data accessible both more quickly and in greater historical depth at the same time the sheer volume of such data is exploding at an estimated rate of 100 percent annually. It is no wonder that 62 percent of CIOs report a serious application backlog, according to CIO Magazine. While compliance with government regulations in the U.S., E.U., and Asia mandates storing business transaction data for longer periods of time than ever before, compliance/regulatory and competitive business needs demand increased visibility into near-real-time transaction history. IT is being asked to store more data, keep it longer, and provide faster access to it. As Gartner has observed, “The problem of storing static data is likely to be the greatest enterprise storage challenge for the next 10 years.”²

Surprisingly, 90 percent of large companies have no data retention strategy beyond keeping data in a production database according to a leading global analyst firm. Not only is storing, indexing, and accessing all required business transaction data in a production application database prohibitively expensive in hardware, license, and headcount costs, but the deadweight of old historical data bogs down the system’s performance for new transactions. Some enterprises believe their database backups qualify as data retention strategies, but they do not. Database backups are really just for disaster recovery, and cannot support routine queries, much less deliver results as quickly as business demands.

Many of the 10 percent of companies with a data retention strategy have moved to data archiving as a strategy, using either custom built solutions or vendor packages. While the growth of archiving proves that business is looking for a lower-cost alternative to storing all its data in a production database, archiving is an imperfect approach. It will prove to be an expensive and ultimately a pointless exercise, as archiving fails to attack the root of the problem: that business transaction data was unnecessarily loaded into the production database to begin with. Archiving may lighten the production database’s load and speed applications, but the archive is just another expensive database, and archiving technology has no fast, simple, inexpensive way to access archived data when it is needed by the business.

Important though it is, business transaction data is unnecessarily clogging the world’s data management systems. According to Forrester, 80 percent of enterprise data is structured, and 56 percent of that 80 percent is historic, unchangeable, transaction data (histories of mobile phone calls, customer retail purchases, banking transactions, etc.). That means 45 percent of enterprise data now retained in expensive databases may not need to be there.

The new business transaction search category ends the loose-loose tradeoffs between data storage and data access caused by database cost. It frees IT from choosing among today’s unpalatable options:

- Spending too much to achieve the goal of keeping business transaction data in a database;
- Saving money on databases but, as a result, not saving enough business transaction data; or
- Saving the data more economically by simply storing it in a file system, but at the expense of such poor access the data is essentially inaccessible.

² Gartner quoted in *Information Age*, April 2005

Business transaction search frees companies from the usual constraint of deciding how much relevant data they can retain based on how much database and hardware they can afford. Organizations can now save unlimited data on low-cost storage anywhere on a network (where it probably resides already), then use a search approach to access it immediately – for business users or BI applications – exactly as if it were in a database.

Transaction Processing, Analytical Processing and Search: It's All About The Index

It is a little understood fact that one of the principle speed bottlenecks for large databases is indexing performance. Indexes are necessary for quickly finding and retrieving specific data when needed, but they impose a formidable performance overhead. Big databases are not slow because of the volume of data they contain, but rather because they tend to have big indexes which are slow.

Until now, business has dealt with the index/performance conflict by compromising, and by over-spending. Online transaction processing applications (OLTP) that require rapid query response – such as those used by service representatives in a call center – make do with fewer indexes and therefore support only limited reporting capabilities. Online analytical processing (OLAP) applications – data warehouses, business intelligence, and other analytics – are rich with indexes, but to speed up these tools enough to enter the daily decision cycle, companies must deploy massive hardware resources to combat the poor performance that rich indexing creates.

The need for compromise and for over-spending is now at an end. The core innovation of business transaction search is a software-only indexing technology that overcomes the performance bottleneck of current indexing approaches (B-tree, bitmap, etc.), delivering performance that is 100 times faster than previously achieved without resorting to the massive hardware and memory architectures usually required for comparable performance.

This new kind of index enables companies to search for and quickly retrieve exactly the records needed from months or years of history and billions of business transactions stored on low-cost file systems. The result is OLAP/BI application performance at the speed now associated with OLTP, made even more powerful by deeply enhanced access to historical data – while saving 80 percent to 90 percent on data management infrastructure costs.

With Business Transaction Search, Enterprise Storage Vendors Could Redefine The Game

The enterprise information architecture of the near future will differ significantly from today's database-centric alignment. The need to store, manage, access, and leverage today's torrent of enterprise data has rapidly increased the importance of storage vendors, who may over the next 10 years equal or even surpass the strategic value of database vendors as they learn to solve customer problems that previously required expensive databases.

In an example drawn from real life, imagine a mobile carrier with 12 million subscribers making 200 million calls a day, or 73 billion calls a year – 16 terabytes of data. Due to government regulations and the carrier's own fraud-detection strategy, the call detail records cannot be batch-loaded into a database at night; the carrier needs to query the records within five minutes of

a call being made. Fulfilling this requirement with a standard, database-oriented data management approach would cost millions of dollars. Using business transaction search, however, the carrier writes the data to log files, indexes it with technology requiring 99 percent less disk I/O than a standard B-tree, and accesses it as though it were in a database – but no database is required.

Imagine the sales opportunity for a storage vendor that could promise the mobile carrier a solution to this problem using no database at all.

Data Retention – You Have To Do It. Now Maybe You'll Be Glad You Do It

For our example, let us remain with the communications industry and emerging government requirements for data retention on a scale never before encountered. As a result, telecommunications providers, Internet Service Providers (ISPs), and other communications companies are ideal candidates for the benefits of business transaction search. Few industries compare in terms of the sheer volume of constantly-generated transaction data stored in enormous quantities, with government regulatory mandates and business-strategy requirements demanding a simultaneous telescope/microscope approach – more data than ever, reaching back into history farther than ever, stored longer than ever, made available for retrieval faster than ever, and in finer detail than ever. Failure to do all of this would not only be bad business, it would also be illegal.

Think about the other ways communications providers are under pressure: Pricing pressure is ruthless. Competition requires 24x7 innovation which leads to product cycles with very short life spans. New technologies, such as text messaging and 3G services create masses of new data. Customer fraud is an unfortunate reality. Meanwhile, there is government regulation trying to protect the public from being defrauded by the company. Last comes, and certainly not least, the issue of national security in which communications providers play a vital role. Telecoms have always dealt with large data volumes, but never at today's volumes. They have always retained data, but never for as long as they are being told to retain it, by law. They have always needed to access their data, but never as fast and as accurately as competition and law enforcement now demand.

Business is taking a look at this problem from a database perspective and starting to panic. The European Telecommunications Network Operators Association and five trade groups representing telcos, ISPs, and other electronic services providers in Europe have appealed to the European Union to do something about the “excessive” costs they are facing in meeting this proliferation of data retention/data access requirements.

The price of storage and hardware is decreasing, but not fast enough to stave off this avalanche of potential database cost...unless the industry discovers business transaction search. Literally within a few weeks and without any additional hardware investment, communications operators can deploy business transaction search to retain 100 percent of their data indefinitely on inexpensive flat files, then retrieve specific data within seconds from tens or hundreds of billions of stored transactions, at enormous savings.

Data retention is a necessity in the communications industry, but with business transaction search it need not be a painful necessity. Once providers discover the riches available in their stored transactions, they will go beyond

mere data-retention-at-the-lowest-price to data-retention-as-a-revenue-driver, unlocking the competitive advantage that, until now, has been buried in the data. Necessity may turn out to be the mother of invention.

Customer Loyalty: How To Turn Your Cryogenically Frozen Log File Data Into Gold For Your Business

The following example looks into business intelligence applied to customer loyalty programs in the supermarket industry. This example was chosen for its universality – everybody goes shopping. For all of us it is possible to see parallels between customer transaction data gathered by supermarkets and our own business. Following this section is an overview of other solution areas where business transaction search promises a price/performance revolution.

Supermarket shoppers who are not IT experts believe supermarket chains do something analytically sophisticated and marketing-related with the information gleaned from the swipe of loyalty cards at check-out. People who know supermarket IT understand that this is not the case; in fact, within some sectors of the supermarket industry, loyalty programs have come to be known as “margin-reduction programs” where loyalty-card users receive a discount in return for simply buying what they were going to buy anyway.

The volume of transaction data generated by swiped loyalty cards – (or, to broaden the focus, the volume of transaction data generated by your company’s applications, e-commerce systems, and security firewalls) – is so huge it would overwhelm even the biggest of today’s data warehouses indexed the old-fashioned way. Like many companies, supermarkets don’t want to throw this data away since it may be useful someday, when someone figures out how to use it cost-effectively. So they store it in inexpensive, quickly-written log files waiting for an innovative way of using it one day. (Does this sound familiar?)

This brings to mind people who have their bodies cryogenically frozen in anticipation of the day when medical science finds a cure for what killed them.

In the most common scenario today, customers who present supermarket loyalty cards at checkout receive printed coupons for future purchases of the same products they just bought. This is not a smart marketing strategy. It’s just a costly way to reinforce current behavior. The supermarket is handing out discounts while receiving no value in return. It certainly is having no material influence on its customers’ buying behavior.

All this can be changed if and when supermarkets (and enterprises like them; i.e., any enterprise that has customers) add business transaction search to their current data management strategy. Then their log files become gold mines of customer history, and the delivery of coupons and other promotions is elevated into one of marketing’s Holy Grails – one-to-one marketing to a Market of One.

With a few simple algorithms and without much in the way of hardware investment or time, a supermarket chain with timely access to an individual customer’s purchase history can effectively employ a business intelligence solution to deliver to customers highly-tailored promotions specific to them, and do so as they enter – not leave – the store.

As a customer enters a store, they swipe their card in a kiosk. The store knows this particular customer buys detergent every 60 days, and it has been six weeks since they last purchased the detergent. The kiosk offers a coupon for detergent, providing an incentive for the customer to buy the detergent today and not at a later time, potentially in someone else's store. Another example of the value of detailed customer purchasing habits is the ability to target specific campaigns at exactly the customers of greatest interest. The system knows the customer is a lifelong Pepsi buyer, and so can sell a competing beverage company, such as Coca-Cola, access to the customer at the very time the customer is most likely to pay attention – when they are in the store shopping. Following Coca-Cola's script, the kiosk starts issuing discount coupons for Coke in escalating amounts. If the customer keeps buying Pepsi despite the coupons, the kiosk increases the offer on each subsequent visit. If all offers are refused, on the next visit the system can alert the checkout clerk to put a free two-liter bottle of Coke directly into the customer's bag on the way out.

A drugstore chain in the U.K. that uses current data warehouse technology recently installed coupon-printing entrance kiosks at enormous expense, but as of this writing the project has not delivered the expected results because the kiosks don't issue relevant offers to individual shoppers. Every shopper gets the same offers. After a while, people walk right past the kiosks. But if the chain used business transaction search to make personal histories available at the kiosk – making historical data from log files available to BI applications as though coming from a database – marketing algorithms could slice and dice each customer's patterns and preferences, making the kiosk investment a smart one. In this scenario, business transaction search enables laser focus on one customer's history, applied to an entire population of customers, at extremely low cost. Intelligently customizing offers to individual shoppers increases a marketer's "share of customer," the classic prescription for delivering greater profitability.

With its capacity for storing enormous volumes of data for as long as a company needs that data stored, and then delivering precisely the data needed in near-real-time, business transaction search holds out rich promise to many industries beyond retail and telecommunications. Hosted software, for example, becomes a more viable undertaking because business transaction search drastically reduces infrastructure cost. Customer service evolves from a cost center to a profit center by going beyond today's limited OLTP applications to give customer service representatives in-depth customer histories, turning them into true salespeople. Homeland security benefits because vast amount of data can be stored indefinitely, then accessed immediately to find patterns of behavior that could foretell or forestall a terrorist strike – all without the cost and complexity of a database.

Conclusion: You Can Solve That Problem Using A Database, But Why Would You?

It is a unique solution that provides both superior performance and lower cost, and when faced with such a solution it is customary for buyers to be leery of the promises. Add to this two solid decades of advertising, public relations, and other forms of indoctrination from database companies to convince business and IT decision-makers that every data management problem requires a database solution, and people's natural skepticism becomes even more difficult to overcome. But the simple fact is that business

and technology have changed a great deal in the last twenty years, and it should come as no surprise that there are new solutions available to meet new changes with greater performance and improved cost-effectiveness.

Once again, it should be made clear that the business transaction search category is not being proposed as a wholesale replacement for databases -- databases are wonderful business tools, but because of their expense and complexity they should be used when the business situation absolutely demands a database. Using a database when a more powerful and less expensive alternative data access strategy is available, however, is simply wasting money.

Saving money is a good business practice, but the business transaction search category promises more. There are business problems that organizations are currently not solving because solving them with databases is simply unaffordable. Companies can now solve those problems cost-effectively, and then go many steps beyond – realizing the full richness of their business by mining the treasures embedded in their business transaction data. All that is required is for business to approach evaluation of this new category with an open mind, and without preconceptions that might turn out to have been leapfrogged by innovation.

UK Corporate Headquarters

CopperEye Ltd, Suite 47, Aztec Centre, Almondsbury, Bristol BS32 4TD
t +44 (0) 1454 203610 f +44 (0) 1454 203330

US Headquarters

CopperEye, 101 Federal Street, Suite 1900, Boston, MA 02110, USA
t +1-617-342-7173 f +1-617-342-7080

e contact@coppereye.com
w www.coppereye.com

©2004-2010 CopperEye, Ltd. All rights reserved. CopperEye and the CopperEye logo are trademarks of CopperEye, Ltd. COPPEREYE_WP_CATEGORY_30JAN2006 A4