

**\*\*\* If your Internet provider or business filters incoming e-mail, please add [CDI\\_Alert@tcdii.com](mailto:CDI_Alert@tcdii.com) to your list of approved senders to make sure you receive the e-mail alerts and newsletters to which you've subscribed.**

**\*\*CDI Alert\*\***

## **IBM Places Bets on CDI Components via SRD Acquisition**

Monday, January 31, 2005

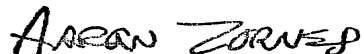
(Note: If you have trouble accessing the hyperlinked articles, please go directly to the alert at [www.the-CDI-Institute.com/Alert](http://www.the-CDI-Institute.com/Alert))

**Customer Data Integration (CDI)**. Do you know how many articles are published each week on CDI? Especially since Oracle re-announced the availability of its Customer Data Hub (while pre-announcing hubs for citizen data, financial data, RFID data, etc. etc.). Talk about information overload!! So many articles, so little time ... This newsletter focuses on newsworthy events that have not been rehashed *ad nauseum* in the press. We also provide the *story behind the story* that industry analysts such as Gartner/Meta and Forrester/Giga are unable to provide due to their internal editorial practices and retainer-based business relationships with the CDI vendors. And lastly, since our CDI Advisory Council has requested this, we will devote future space to *best practice* strategies they should/must consider. The intent is to provide ongoing advice for CIOs, CTOs and CDI project leads in helping develop their CDI strategies – **independent, authoritative, and relevant analysis**.

Something that you may not have caught in the press is the January 7, 2005 announcement that **IBM acquired CDI vendor SRD**. If you don't think this is news, think again. Yes, during 2004 other vendors acquired CDI technologies (SAP acquired A2i [July 2004 [press link](#)]; Initiate Systems bought Journé [July 2004 [press link](#)]; Siperian bought Delos [October 2003 [press link](#)]; and D&B bought Group1/Sagent [July 2004 [press link](#)]). And, of course, IBM also acquired Trigo [March 2004 [press link](#)] to bolster its master data capabilities. **But SRD is a different animal as it had the distinction of pioneering “anonymous entity resolution” matching algorithms**

**Bottom line: IBM is serious about master data integration and has taken a leap ahead of the CDI vendor crowd** by beefing up its atrophied/lackluster CDI investments-to-date (IBM Client Information Integration Solution and DB2 Information Integrator) while simultaneously cornering the market in **anonymous entity resolution** for increased matching accuracy (vital to homeland security as well as valuable to most other industries seeking customer data integration).

“Heads up” from the CDI front lines,



Aaron Zornes  
Chief Research Officer  
[The CDI Institute](http://The-CDI-Institute.com)  
<mailto:editor@the-CDI-Institute.com>

To automatically subscribe, enter your e-mail address on the form at [this link](#).

# CDI Alert: IBM Acquires SRD

Monday, January 31, 2005

## What's New?

With this acquisition, [IBM significantly extends its fundamental capabilities for customer data integration](#). IBM is serious about master data integration and has taken a leap ahead of the CDI vendor crowd by beefing up its atrophied/lackluster CDI investments-to-date (IBM Client Information Integration Solution and DB2 Information Integrator) while simultaneously cornering the market in an anonymous entity resolution for increased matching accuracy (vital to homeland security as well as extremely valuable to other industries). But this is not the first acquisition for IBM in recent year. For example, there was also AlphaBlox for business intelligence dashboards and Trigo Technologies for product data management. [It is good to see IBM behaving like a world class software business again.](#)

## Why Do We Care?

[Customer data integration \(CDI\) is going to be a huge software market](#). And not just because Larry Ellison and Chuck Phillips are pitching it as one to Wall Street and the Oracle user groups. [Database vendors cannot stand by and let start-up CDI vendors or application package vendors corner the market in master data management – especially given that customer data integration will rise to the Top 5 list of many Global 2000 enterprises' strategic investments during 2005-06.](#)

[Conceptually, the dominant database vendors \(IBM, Microsoft, Oracle\) have an inside track advantage in their ability to execute/integrate pattern matching and other key CDI capabilities within the database itself to provide better performance, reliability and scalability](#). However, we all know what happens to databases that attempt to accomplish too much! [*Does anybody remember Informix when they acquired Illustra for \$400 million in early 1996? Whatever happened to Oracle's video-on-demand server from five+ years ago? Remember data blades, data cartridges and abstract data types?*] [It is our belief and general industry consensus that there's more to customer data integration picked than pattern matching and query optimization](#). For most observers, customer data integration/master data management but pretty far up the software stack and infringe upon the data models and process/workflow inherent in application packages. While we do believe that the enterprise application package vendors (Oracle/PeopleSoft, SAP, Siebel) will have great success with their CDI product families, these are necessarily generic to a great extent, and not as industry-specific that is the more nimble (and somewhat-endangered) products from the standalone CDI solutions vendors (DWL, Initiate Systems, Siperian). As we previously predicted, the standalone solutions need to go as vertical as fast as possible and as well as allowing quickly with certain large scale systems integrators. So we see DWL in banking and insurance; Initiate Systems in hospitality, governments, and retail pharmacy; and Siperian in pharmaceutical and life sciences. But what about IBM? IBM recently more-or-less "sunsetted" its Client Information Integration Solution (CIIS) by moving the software development teams from their industry-specific focus (Ireland and financial services) and blended them into the Software division. [But IBM's Software division does not develop applications, and, yes, we believe that for the next two years that CDI is more of an application than an infrastructure decision for most enterprises.](#)

[Longer term, through 2007-08, it is clear that that the long-term viability of IBM's and Microsoft's database strategies will depend upon their use as the underlying database for enterprise applications – including master data management capabilities such as CDI.](#) [A quipster or wag might say "By 2006, a

*DBMS lacking native identity management and cross-DB entity resolution will be like an 'abacus missing some of its beads'.* Yes, IBM, Oracle and Teradata will continue to dominate on the high end while Microsoft will dominate in the small to medium business sector. It is important to note, however, that most enterprises have acquired (and will continue to acquire) application packages from more than one strategic vendor. And often, the same packages do not all get deployed using a single database brand. Yes, there may be a center of gravity around SAP or Oracle/PeopleSoft or Siebel, but this does not provide an incumbent advantage to the degree one might think. Often the driver for a CDI solution is to provide an insulation layer (service-oriented architecture of business services via web services) to link together these disparate application architectures (data models, process models, middleware stacks, etc.). [Through 2006-07, there will be at least three camps providing CDI solutions: mega app vendors \(Oracle, SAP, Siebel\), standalone CDI solutions \(DWL, Initiate, Siperian\), and data service providers \(Acxiom, Experian, Harte-Hanks\). And therefore IBM is growing its own standalone solution, albeit closely tied to DB2, through its acquisition of SRD.](#)

[We have been involved in 100+ CDI evaluations over the past year, and as a result we have been evolving a set of "top 10 CDI evaluation criteria" through our briefings and workshops with IT organizations around the world](#) (see below for this short list). Note that IBM is now backing off on the number one evaluation criteria customer data models as a result of its downplaying the Dublin Assets. Furthermore IBM does not have a strong story (yet) when it comes to process flow or workflow ("business services") for its CDI solutions. For a detailed discussion of these evaluation criteria, see our recent report titled "Customer Data Integration: Market Review & Forecast for 2005-2006" (download via [this link](#)).

Our client experience shows that in most cases there are ten key criteria categories that an IT organization needs to focus on during its assessments of CDI solutions, including:

1. Data model
2. Business services
3. Identity management
4. Data management
5. Architecture
6. Infrastructure
7. Connectivity
8. Analytics
9. Developer productivity
10. Vendor integrity

[Given that enterprise CDI represents "mission-critical infrastructure" it should come as no surprise that this endeavor is most often lead by \(and even championed by\) IT management.](#) In the larger enterprises, it is common to outsource the creation of a request for proposal (RFP) that is tailored to the enterprise's specific CDI requirements – often the systems integration vendors such as IBM BCS or Accenture are chosen to lead such requirements gathering and classical gap analysis.

[In our view, the acquisition by IBM of SRD is essentially a move to bolster DB2 relative to Oracle, and less of a bold move into the standalone CDI solution.](#) No doubt, IBM will sell a good deal of SRD to large businesses and the U.S. government intelligence agencies. However, the SRD product by itself is not a tremendous standalone CDI solution. Nor is it tied specifically to IBM DB2 (at the moment).

## What is SRD and Why Did IBM Buy into It?

IBM is in the software, hardware and services businesses. Anything that moves one or more of these as a large ticket sale is very attractive to IBM. CDI solutions are at present a software and services play for the vendors (but arguably represent insurance for IBM hardware to stay in place as well).

[SRD \(Systems Research and Development\) was a privately held company that specialized in providing analytics and high-end algorithms for the gaming industry](#) (it was rumored that all but one of the big casino conglomerates had major investments in SRD software and services). While the gaming industry is a growth industry in North America, and, indeed around the world, SRD's momentum had somewhat peaked in the past twelve months. Despite an infusion of former IBM executives for executive management and business development, the product was still not sufficiently packaged as a software solution and remained a high-end services play with a somewhat secretive, smallish software startup (secretive in the sense that SRD publicly touted the CIA connection through its In-Q-Tel investment arm as an investor). [The software tools NORA \(as in "non-obvious relationship awareness"\) and ERIK \(a product that "determines who is who"\) received sufficient kudos from technical evaluators at large firms as a good fit for "identity resolution software". More specifically, the SRD software helped enterprises gain an accurate view into individuals and relationships -- in real-time and did so without physically merging/collating databases.](#) There is no doubt that ferreting out relationships ("associations") between good guys and bad guys, or bad guys and other bad guys, enable an interesting and wide range of new business analytics. And without the expense of creating a massive/expensive physical data warehouse. But can't one do this with [federated query products and distributed metadata managers](#) (Avaki, InfoGlide, MetaMatrix, et al)? No – in fact, otherwise IBM's DB2 Information Integrator would continue to get lead billing in IBM's business analytics marketing. Until now – because [with SRD one gets "anonymous" entity resolution in real-time and with increasing accuracy.](#)

No doubt, IBM Global Services (IGS) and IBM Business Consulting Services (IBM BCS) will bring the SRD tools into engagements because they are just that ... tools, and not an application. Which means more expensive high-end programming talent needed to successfully deploy this solution (remember IBM Intelligent Miner?). However, these are mission-critical applications (homeland security, customer data integration, fraud detection/prevention, database campaign management). IBM had made statements that it plans to apply SRD technology with its Banking Data Warehouse and Crime Information Warehouse. This is good, as these "mature" product lines could use a mid-life kicker as well!

## What are the products?

[The new family name is "IBM's Entity Analytic Solutions" with family members: DB2 Identity Resolution, DB2 Relationship Resolution, and \(unannounced\) DB2 Anonymous Resolution.](#) These technologies help enterprises (business and government) increase their analytical insight by delivering an accurate view into individuals and relationships in real-time – "associations that previously were nearly impossible to discover, enabling a wide range of solutions" per IBM.

**IBM DB2 Identity Resolution** ([IBM product page link](#)): This is the new name for the basic name/address matching capability ("built-in heuristics techniques such as name stemming and address variations" for identity recognition) – formerly known as "ERIK" and formerly marketed as [real-time identity recognition software](#) (pre-IBM acquisition). Using proprietary, patent-pending technology, Identity Resolution can correct prior decisions in real time when new data exposes the truth. Without the use of database table scans or joins, Identity Resolution can digest real-time data feeds while preserving the full attribution and integrity of every record, regardless of the quality of the source data. And all of this is maintained independent of underlying/source databases and data warehouses which allows IBM to claim that "DB2 Identity

Resolution grows identities and self-corrects over time, which means that users don't have to periodically reload, and the accuracy and precision of identities compounds". Nice! In its pre-IBM days, we were aware of several sites that literally had hundreds of real-time data feeds into their SRD algorithms. Hmm... sounds like the U.S. federal government might to consider this in lieu of its proposed 10 year "death march" to force healthcare entities to implement the as-of-yet-undefined National Healthcare Identification Number ([see an industry response](#)). Now Initiate Systems has some good company via IBM in its campaign to persuade the Feds that the U.S. can live without an intrusive national ID for healthcare consumers ([see Initiate's white paper](#)).

**IBM DB2 Relationship Resolution** ([IBM product page link](#)). Formerly, this was (quite aptly) marketed by SRD as NORA as in [non-obvious relationship awareness](#) and [beyond 360 degrees](#). Not to mention, the "Kevin Bacon six degrees factor" ("everyone is related to actor Kevin Bacon by six degrees or less" – [see link](#)) as in "relationship detection across 30 degrees of separation" per IBM. Additionally, false negatives (when two identities are the same but not recognized as such), and false positives (when two identities are determined to be the same when they are fact not) are greatly reduced. Most importantly, this is [real-time access against hundreds or thousands of discrete data sources to ferret out both obvious and non-obvious relationships](#).

**IBM DB2 Anonymous Resolution:** This unannounced product is further designed to keep anonymity sacrosanct from "those who do not need to know" as when ERIK and NORA make a case for nabbing their villain (remember "Nick and Nora" of "The Thin Man" film series? SORRY! Film buffs, please cross check with [www.IMDB.com](http://www.IMDB.com)) In other words, this tool enables organizations to [compare identity data with each other without making that data visible to the other organization -- to check identities while protecting the privacy of customers or citizens](#).

**Bottom Line:** [IBM is serious about master data integration and has taken a leap ahead of the CDI vendor crowd](#) by beefing up its atrophied/lackluster CDI investments-to-date (IBM Client Information Integration Solution and DB2 Information Integrator) while simultaneously cornering the market in [anonymous entity resolution](#) for increased matching accuracy (vital to homeland security as well as extremely valuable to most other industries seeking customer data integration).

## \*\*\* About The CDI Institute

To provide feedback on our client's CDI initiatives we have two levels of sponsorship for IT organizations: (1) free membership (by invitation) in our CDI Advisory Council providing unlimited CDI consultation by phone, and (2) free membership in our CDI Business Council (survey base) which provides bi-weekly updates on key CDI trends and issues via an email newsletter.

- **CDI Advisory Council™** of fifty organizations who receive unlimited CDI advice to key individuals, e.g. CTOs, CIOs, and CDI project leads
- **CDI Business Council™** of 450+ Global 2000 IT organizations who receive a limited distribution, bi-weekly newsletter with CDI industry updates
- **CDI Alert™** bi-weekly newsletter provides IT organizations, CDI vendors, and investors hard-hitting insights into best practices as well as market observations derived from interactions with the **CDI Advisory Council™** and the **CDI Business Council™**. Initially free to qualified individuals, the **CDI Alert™** is expected to become a fee-based product 2H2005. The intended audience includes: CDI project managers, CIOs, CTOs, chief customer officers, chief privacy officers, data quality managers, data stewards, market analysts, metadata managers, and project teams responsible for CDI solutions and infrastructure, data quality, data warehousing, customer relationship management (CRM), enterprise resource planning (ERP), product data management (PDM), supply chain management (SCM), partner relationship management (PRM), and business intelligence.
- **CDI MarketPulse™** monthly survey results, e.g. budgets, success/failure rates, mindshare based on ongoing surveys of the CDI Advisory Council and the CDI Business Council
- **CDI Fast Track™** quarterly 1-day workshop -- fee-based and rotating through the major North American metropolitan areas

**To automatically subscribe**, enter your e-mail address on the form at [this link](#).