

Solution Brief

That's News to Me: Investigative journalists find vital, unexpected relationships

News organizations rarely hesitate to proclaim their ability to provide “all news all the time” or “all the news that’s fit to print.” Yet no tagline claims that reporters have regularly provided all the news about a particular story. That’s left to investigative journalists who spend months and, sometimes, years tracking down every lead and checking every fact.

The result is stories that reveal unexpected connections, formerly hidden or unknown facts, and revelations about corruption, inexplicable successes, human failures and triumphs, and secret negotiations. The published stories lead to increased circulation or ratings and, frequently, Peabody Awards and Pulitzer Prizes.

Of course, the facts were there all along. It just took research and phone calls and legwork to pull them all together and discover their inter-relations. Yet, in a 24-hour-news-cycle world, the months and years required by investigative journalism are far too long. YarcData’s Urika™ graph analytics appliance can shorten that cycle to days and, potentially, hours.

Its ability to work with structured and unstructured data – without conforming it to a single format – and find information that was hidden or unexpected (which a typical specific search wouldn’t find) let journalists discover, for instance, that a majority of college students contributed to Republican candidates in 2012. Given the traditional liberalism of college students, Urika helped journalists tie the students to parents who had contributed the maximum legal amounts and used their children as conduits for additional funds – a potential violation of election and money-laundering laws.

The ability to unveil relationships among seemingly unrelated facts can also help journalists track down financial irregularities and fraudulent activities on Wall Street, tie local cancer rates to industrial run-off, provide information about prior or simultaneous research that wasn’t mentioned in news coverage of a medical breakthrough, and uncover sources who can shed light on the players influencing international trade agreements.

Each story has the potential to move markets. Each can lead to the implementation of new procedures to prevent future occurrences. And each relies on YarcData’s Urika to reveal to journalists the things they didn’t know that they didn’t know. And that’s the most powerful knowledge there is.



The Urika™ Solution

Urika's speed and thoroughness are directly related to its huge, globally shared memory of up to 512TB which holds every historical and current database resource in-memory and, using the SeaStar™ chip which can handle up to 350TB of I/O per hour, can quickly update relationships dynamically as new data becomes available.

The massively multi-threaded architecture of the Threadstorm™ (128 independent threads) is specially designed for analyzing graphs and allows threads to continue executing even if some are waiting for data to be returned from memory. Plus, the extreme flexibility of the Urika system architecture lets I/O or compute nodes be independently added as the application demands. This design facilitates relationship exploration.

Query results are provided in human time – minutes rather than overnight or multi-day runs – so that refinements can be made multiple times if necessary within the time-sensitive window of daily or even 24-hours news cycles.

News analytics opportunities

Once investigators have, with help from their IT team, created the appropriate query logic for their hypotheses (using the industry-standard software bundled with the Urika appliance), they can develop a library of queries that can be applied to similar stories in the future. That, in turn, reduces investigative time even further by eliminating the query-building step.

Good guys get acknowledged faster, corrupt practices come to light sooner, and the public has the chance to understand the dynamics that affect their daily lives.

About Urika

YarcData's Urika is a big data appliance for graph analytics. Urika helps enterprises gain business insight by discovering relationships in big data. It's highly-scalable, real-time graph analytics warehouse supports ad hoc queries, pattern-based searches, inferencing and deduction. Urika complements an existing data warehouse or Hadoop cluster by offloading graph workloads and interoperating within the existing analytics workflow. Subscription pricing or on-premise deployment of the appliance eases Urika adoption into existing IT environments.

About YarcData

YarcData, a Cray company, delivers business-focused real-time graph analytics for enterprises to gain insight by discovering unknown relationships in big data. Adopters include the Institute of Systems Biology, the Mayo Clinic, Noblis, Sandia National Labs, as well as multiple deployments in the US government. YarcData is based in the San Francisco bay area and more information is at www.yarcdata.com.

