MOBILE

BIGDATA REPUBLIC | Transform Your Business With Data

ABOUT US   REGISTER   LOGIN

HOME   BLOGGERS   MESSAGES   POLLS   WEBINARS   RESOURCES   VIDEOS | STRATEGY   ANALYTICS   TECHNOLOGY   FB   TW   LI   G+   RSS

FINANCIAL SERVICES   HEALTHCARE   GOVERNMENT   EDUCATION   RETAIL

BIG DATA ANALYTICS

The Disappearing Database

Robert Plant, Associate Professor, School of Business Administration, University of Miami
11/21/2012
Comment
21 comments

Login
👍 👎
50%   50%

Technology never stands still. Just as big data starts to come into focus, there is an emerging alternative to the big data model: data-as-a-service (DaaS).

This model comes from the confluence of two areas: the cloud and search. Pioneering vendors have begun to create data marketplaces -- cloud-based platforms where companies can place datasets for sale. Microsoft has utilized its Widows Azure platform to provide its clients with tools and services for managing data and processing in the cloud (services that used to be performed inside corporate datacenters). This creates a stable, salable platform for data storage, management, and processing. In conjunction with this, Microsoft has created the tools needed to access the data from the cloud via the Internet, employing solutions from the Bing Search environment.

Here are some examples of datasets that are available.

- Infoconnect is a real-time Web service that provides business firmagraphic data.
- MetricMash US PPI- Industries provides access to PPI data by industry and covers more than 600 industry price indexes in combination with more than 5,000 product lines and categories from 1926 to the present.
- MacroTrenz gives access to more than 100,000 Federal Reserve Bank reports.
- Data.gov.uk offers access to UK government metrological data.
- The Stock Sonar provides sentiment scores of US public companies.

One advantage of this approach is that it is a pay-as-you-go model; you pay for what you consume. Additionally, companies do not have to store the data locally. They can keep it on the original cloud, reducing transfer costs and time. This has the added benefit of low security risks and built-in backup.

The DaaS model is a form of data mashup in an executable, standards-based environment in which companies can bring divergent data together on an as-needed basis. This mashup can be used in conjunction with corporate data extracted from internal systems or alone on a one-time basis. The advantage of using a platform such as Microsoft's is that the data uses common standards and thus is highly malleable, forwardly consistent, and maintainable. Additionally, the ability to provide processing power for these datasets is also scalable through the vendor using the SaaP-Saal facility. Again, this reduces the operating and capital expenses generally associated with an internally developed solution.

Early days for DaaS
Why has this not yet caught on? First, we are still in the early days, and competitors are watching the pioneers carefully. The major issues for executives and managers fall into two categories. The first category is associated with everything cloud: security, costs, public vs. private, moving large internal hardwired systems to the cloud, etc. The second category revolves around cloud-based data: data ownership, transportability, reliability, etc. All these concerns are being addressed by vendors, corporate legal teams, and legislators. However, due to the speed at which technology moves forward, there are sometimes lags in the legal and legislative frameworks within which products operate.

For managers, now is the time to consider how this may be implemented at their firm. The following four steps can be helpful in developing a DaaS strategy framework.

- Assess your data requirements in the immediate, short, and longer terms. For example, determine what data would provide quick wins if the data were available elsewhere, and whether those wins would otherwise be impossible due to the delay in organically building that dataset.
- Determine what aspects of the data required are already available on the cloud, as well as when new cloud datasets will become available. Why rebuild a dataset modeling an economic environment if a government agency already does that on a scheduled basis and has placed that data on a cloud marketplace for acquisition?
- Assess the cost benefit of moving your data to the cloud or accessing cloud data and linking it to your internal data in a public-private cloud mashup.
- Assess what data you may wish to provide to a cloud-based data market and thus create revenue from an otherwise dormant asset.

This approach to big data contrasts with the proprietary internally focused, organically developed big data model and provides a low-cost, flexible data management solution that allows opportunistic data mashups and avoids the creation of a data white elephant. As data explodes further, it will become increasingly uneconomic to hold big data internally. The security risks associated with internal database

---

MORE BLOGS FROM ROBERT PLANT

A Review: Big Data in 2013
39 comments
Data scientists, NSA, C-suite buy in... Big data changed a lot in the last 12 months.

What Would Seymour Cray Think of Big Data?
13 comments
How would the tech legend respond to the 4 Vs?

The 360 Degree View of a Customer
56 comments
Big data is helping companies like Netflix see what you really want (and even when you go to sleep).

Lessons for Big Data From President Obama's Healthcare Implementation
24 comments
Three big data lessons for CEOs from HealthCare.gov's bumpy start.

Is Big Data Ruining Sports?
29 comments
From sailing to the NBA -- big data could increase the gap between sporting haves and have-nots (while ruining the magic in the process).

More from Robert Plant

FLASH POLL

All Polls

DIGITAL AUDIO

Big Data Analytics Radio
Latest Archived Broadcast
Your Big Data Initiative Starts With a Great Team
The Big Data Conference and Nemertes Research have teamed up to provide expert advice on building a big data team and finding the right leader. Join us on Tuesday, Aug. 20.
Listen

Full schedule   Archived shows   Download to calendar

DATA VISUALIZATION SHOWCASE

Visualizing Sporting Value: Football Player Transfers
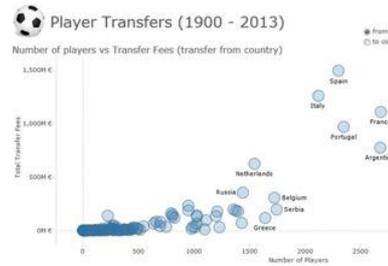
A great visualization of transfer fees leads us to think of deeper sporting analytics.
Explore this data here.

More Data Visualization Showcase

systems will become increasingly expensive to manage, and as collaborative data sharing becomes essential to business fluidity, the cloud is the only solution.

Related posts:

- Hadoop Won't Turn Your Dump Into a Data Warehouse
- Failure to Manage the 3 V's Is Not an Option

— Robert Plant, *associate professor,* *School of Business Administration, University of Miami*

Email This   Print   Comment

COMMENTS

Newest First | Oldest First | Threaded View

PAGE 1 / 3   >   >>

technetronic, User Rank: Blogger
12/5/2012 | 12:50:23 PM

Re: DaaS in Research Community
Anna, great point re: dataset efficiencies. There's definitely a lot of duplication, though to some extent it's because we need to have multiple sources anyway.

The thing with the fragmentation and the existing business research market is that the spread of misinformation is slower than is increasingly possible with digital systems (I hope). The additional risk of bad informatio nbeing spread means that we have to become smarter data users, and that's a topic that I think deserves attention.

Login

50%   50%

Reply | Post Message | Edit/Delete | Messages List | Start a Board

mharden, User Rank: Exabyte Executive
11/23/2012 | 5:33:17 PM

Re: Creativity from outside
Fortunately, Data-as-a-Service via the cloud has emerged as a vehicle to implement a new set of best practices for not only verifying data at the point of entry, but also verifying existing records: thus ensuring higher data quality. If companies get a comfort level with security this tech is definitely take off.

Login

50%   50%

Reply | Post Message | Edit/Delete | Messages List | Start a Board

mharden, User Rank: Exabyte Executive
11/23/2012 | 5:32:44 PM

Re: Creativity from outside
Fortunately, Data-as-a-Service via the cloud has emerged as a vehicle to implement a new set of best practices for not only verifying data at the point of entry, but also verifying existing records: thus ensuring higher data quality. If companies get a comfort level with security this tech is definitely take off.

Login

50%   50%

Reply | Post Message | Edit/Delete | Messages List | Start a Board

Saul Sherry, User Rank: Blogger
11/23/2012 | 6:13:54 AM

Re: Creativity from outside
@SharCo Legal issues as well as security issues - one would hope a full outsource issue would also provide a series of storage options (locations and security) as well as providing advice (or taking liability) for the legal portion of the discussion

Login

50%   50%

Reply | Post Message | Edit/Delete | Messages List | Start a Board

SharCo, User Rank: Petabyte Pathfinder
11/22/2012 | 1:56:59 PM

Re: Creativity from outside
Kiran, Agreed. As they say, it's never good to keep all your eggs in one basket. It's not only outsourcing, but also hosting them on a server located somewhere else but still within the country. A lot of security issues with the Cloud being brought up so I'm 50-50 on the outsourcing bit.

Login

50%   50%

Reply | Post Message | Edit/Delete | Messages List | Start a Board

Anna Young, User Rank: Exabyte Executive
11/22/2012 | 10:36:27 AM

DaaS in Research Community
DaaS may seem to be a new phenomenon but I have a feeling we already have a form of this in some crude shape already. Research institutions have always generated, collated, analyzed and packaged data for sale to institutional investors and business. This is not new. Perhaps what's new here is that the universe of organizations and institutions that is able to utilize this service is expanding as is the number of organizations offering the service.

What's fascinating about DaaS is that we may now have the opportunity to reduce waste in the process. Sometimes companies duplicate data generation, not realizing that what they are looking to review has already been done by another organization. If DaaS offers a private data generation opportunity, a firm may simply have the chance to search for organizations that may have done research into an area it's interested in and simply have them analyze the data based on its own parameter. The data may need to be reviewed or updated but you aren't starting anew.

Login

50%   50%

Reply | Post Message | Edit/Delete | Messages List | Start a Board

Saul Sherry, User Rank: Blogger
11/22/2012 | 8:46:14 AM

Re: Not So Early Days for DaaS
@Will... paranoia aside, that will be the hole that data inferencing will fill. Our behaviour reveals more than our words, so if marketing departments can be clever enough to read into actual actions, I think they will be fine!

Login

50%   50%

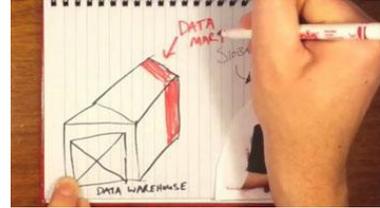Reply | Post Message | Edit/Delete | Messages List | Start a Board

Saul Sherry, User Rank: Blogger
11/22/2012 | 4:40:40 AM

Re: Not So Early Days for DaaS
@technetronic - an interesting premise. It's something you can see starting with geeky pranksters, seeing what how they can get away with - much like all the spurious updates you find on wikipedia. But oncethe business sabotaging potential is realized, could this become the next corporate warfare weapon?

Login

50%   50%

Reply | Post Message | Edit/Delete | Messages List | Start a Board

---

Will, User Rank: Bit Player
11/21/2012 | 8:55:55 PM

Re: Not So Early Days for DaaS
@technetronic I 110% agree with your paranoia, I'm not sure how effective some uses of big data in marketing will be longer term considering the sources it uses.

Login

(n.b. I've just deleted 2 paragraphs of paranoia fueled propaganda :))

50%   50%

Reply | Post Message | Edit/Delete | Messages List | Start a Board

---

technetronic, User Rank: Blogger
11/21/2012 | 8:38:20 PM

Re: Not So Early Days for DaaS
A somewhat sensationalist headline because databases are not disappearing, they're becoming *standardized* -- it's a great article.

Login

I'm excited by this idea because it's a big step toward unified "human" knowledge. If we can all pay for access to the "best" source of data, then at least all users are working from a common, comparable dataset.

50%   50%

Of course, the cynic in me worries about the insider information market. Not only are there concerns about hiding information for personal gain, but what about disinformation campaigns hitting a central database? This is more paranoid than I like to be but it feels important in these early days.

Reply | Post Message | Edit/Delete | Messages List | Start a Board

PAGE 1 / 3   >   >>