MOBILE

# BIG DATA
# 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

## HADOOP

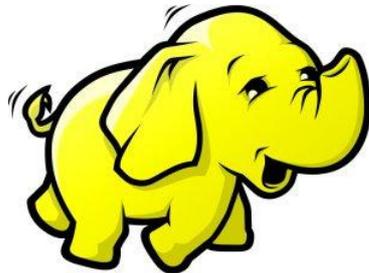# Hadoop: Executives & Managers Need to Read the Warning Labels

**Robert Plant**, Associate Professor, School of Business Administration, University of Miami
11/23/2012
Comment
11 comments

Login

50%   50%

If you were going to buy your child a soft toy, an $18 elephant for example, you would probably look carefully at the warning labels: Is it flammable, does it present a choke hazard, or contain poisonous substances? Assuming no present or latent threat, you may proceed to buy it and name it... Hadoop, for example.

Companies would be advised to follow the lead of parents and read the warning labels of technology products that they buy. These are not the ones that say "Hazard 110V," but rather the ones that are understood by the cognoscenti. Unfortunately, they are also the ones overlooked by the popular media who hype the latest and greatest technology and imply as if it was built out of granite.



**Look after your big data toys**
Hadoop the software package is in danger of becoming a big corporate toy. A toy which, if looked after (like the toy elephant), could bring bountiful happiness and reward to its owner. Unfortunately, it could also, under some circumstances, lead its owners to the ER.

CXOs looking at the stories of Hadoop happy endings in the press need to be extremely wary of failing to understand that it's only on version 1.0.3. And that open-source does not mean there is no potential for a downside.

This is not like buying a spreadsheet and plugging the data in. While the attraction is understandable, success stories abound, and names like Google, Facebook, Amazon, Twitter, Walt Disney, Yahoo, IBM, and Netflix all use this technology, if we take a step back and reflect, we see that these companies have some very striking similarities. They are all giant tech companies, have ultra-high tech-IQs, extensive tech support capabilities, deep pockets, a long history of leading tech change, and a management that understands Hadoop.

**Two key questions**
If your company is thinking of joining the big data revolution, and your technology team is talking to you about funding a Hadoop solution, here are two litmus test questions before starting the journey.

First, ask them what MapReduce is and where it came from. If they can't define it in non-technical English, this is not a good sign. If they reply, "it's a way of representing or placing products on a map," then you're definitely not ready. It is, in fact, the programming model that underlies the computational implementation.

Second question: What is the programming basis of the MapReduce? The answer should be a pure functional programming model, technically a language one without side effects, such as Haskell, work that originated in the 1980s by the likes of David Turner, and members of the Oxford University Computer Lab (Programing Research Group). This technology isn't so new that people haven't had a chance to see it, and technically aware IT staff should know this. For a more complete list of things Hadoop is not, the Hadoop "Warnings Page" makes fun and interesting reading.

Fundamentally, understanding how and why a technology works remains important, and we need to remind ourselves of this. It's seductive to think we're all technologists now, thinking due in large part to being intoxicated by the simplicity of a smartphone or tablet. We forget these are consumption devices, and not devices our organization is ever going to directly program or write an App for.

## 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

Hadoop Radio
Latest Archived Broadcast
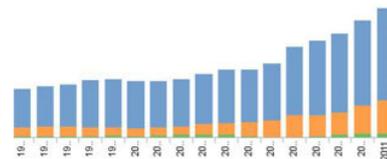Stop Kidding Yourself: Cloud Realities You Need to Embrace
Join this radio show to truly understand what a CIO needs to do to build a successful private cloud and what skills and values the IT team will need to embody.
**Listen**

Full schedule   Archived shows   Download to calendar

## DATA VISUALIZATION SHOWCASE

### Anatomy of a Visualization



This Tableau visualization of international debt demonstrates how simple visualizations can give great insight
Explore this data here.

More Data Visualization Showcase

## BDR IN YOUR INBOX

Enter email          SUBMIT

**Four outcomes if you don't understand your own tech**

This inability within an organization to understand technology at a deep level has four potentially worrying outcomes:

- First, firms are forced to rely on third parties to do the technical development work. This can lead to lags in cycle time; difficulty in developing meaningful SLA contacts (the technology is so new there is no prior history); and higher development costs as the demand for the product drives up the rates.

- Second, firms without technical depth are automatically laggards, handing over the technical advantage to rivals who do have depth. These rivals are able to spot innovative technical opportunities and act on them, further extending their advantage.

- Third, technically shallow firms won't know what they don't know. So when they see a label that says Version 1.0.3, they may fail to understand this is a warning label, not just a version number. For example, the following is a "Release Note" from Hadoop 1.0.3, referring to changes since 1.0.2:

  > Build fails with non-Sun JREs due to different pathing to the operating system architecture shared libraries

  > The src/native/configure script used to build the native libraries has an environment variable called JNI_LDFLAGS which is set as follows:

  > JNI_LDFLAGS="-L$JAVA_HOME/jre/lib/$OS_ARCH/server"

  > This pathing convention to the shared libraries for the operating system architecture is unique to Oracle/Sun Java and thus on other flavors of Java the path will not exist and will result in a build failure with the following exception:

  > [exec] gcc -shared ../src/org/apache/hadoop/io/compress/zlib...

- Finally, your system may work, but produce the wrong results. Worse, you might then act on them.

Yes, you can point the finger once again at the technology group, but if you just lost the revenue race in the end of year holiday retail sales period, the finger pointing may not be much comfort when you're asked to explain why the business went along with the big data suggestions. Saying "We just did not understand the technology" may not help!

Perhaps now is the time to not just develop a technology roadmap for your next big data acquisition, but really work to develop a human capital one as well. After all, its better to know what you don't know than find out the hard way.

- The Beauty of Big Data Models
- Big Data VC Shifts Towards the User Friendly
- Hadoop Won't Turn Your Dump Into a Data Warehouse
- Delivering on the Big-Data Promise

— Robert Plant, *Associate Professor, School of Business Administration, University of Miami*

Email This   Print   Comment

## COMMENTS

Newest First | Oldest First | Threaded View

PAGE 1 / 2   >   >>

MDMConsult, User Rank: Exabyte Executive
11/30/2012 | 7:02:47 PM

**Re: Bypassing the malpractice**
With the growth of Hadoop and the more widespread adoption becomes within the enterprise, the pressure will increase on both ends and offering education will just normally adapt. Having the right skills drives demand in skill, in salary, etc al

Login

50%   50%

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

Anna Young, User Rank: Exabyte Executive
11/29/2012 | 5:05:31 PM

**That's somebody else's job**
I hate to disappoint you but I don't check labels on products and I certainly don't examine them for flammability! I admit that's stupid but I expect the government regulators to do that and they've done fairly well so far. However, your point is well noted. Perhaps I should be paying more attention to labels, aside from the ones on food items, that is.

Let's be honest. We've all become so reliant on government bodies to safeguard us we've abandoned the things we should do to assure our well being.

Login

50%   50%

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

smkinoshita, User Rank: Exabyte Executive
11/29/2012 | 2:29:41 PM

**Re: more training**
*"I think some companies don't take ownership of the technology by training employees and building up knowledge within the organiation."*

@mharden -- You're not kidding. I've found that problem popping up in all sizes of organizations. It's even worse when the people who actually have the training leave -- without passing their training on to others. Then even when things were started up properly, they're not maintained properly or other applications are added without fully understanding how they all have to connect.

Imagine if an organization begins using Hadoop properly, but the only person who really understands it leaves, and those who are left have just enough knowledge to be dangerous?

Login

100%   0%

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

mharden, User Rank: Exabyte Executive
11/24/2012 | 9:16:34 AM


**FEATURED VIDEO**

**Big Data Explained: What Is YARN?**   21

Getting to grips with Hadoop 2.0? Here's a simple video explaining YARN.
Watch This Video

More Video Blogs

**FOLLOW US ON TWITTER**

**Big Data Republic**@BigDataRepublic   15 Apr
AllJoyn: A Common Language For Internet Of Things - Open-source lingua franca aims to let IoT devices and services... ow.ly/2FHbY7
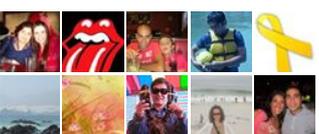Retweeted by Yakamoto
Expand

Hajime Ozaki@hajozaki   22h
BDR Top 20 #BigData David Smith David Smith List of top tweeters on Big Data from @BigDataRepublic wp.me/p4Aavi-2
Show Summary

Marcus Borba@marcusborba   28 Apr
#BigData #Analytics #BusinessIntelligence #FS @informationweek @AllAnalytics @BigDataRepublic @BigDataNetwork @BigDataClub

**LIKE US ON FACEBOOK**

**Big Data Republic**   Like

4,168 people like Big Data Republic.

Facebook social plugin

**ACCOLADES**

MEDIA PIONEER AWARDS WINNER 2013

Media Pioneer Award Winner

**more training**
I think some companies don't take ownership of the technology by training employees and building up knowledge within the organiation. I actually have a client that has been that way on a different application. Your "Four outcomes if you don't understand your own tech" section exactly describes this customer.

Login

50%  50%

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

---

SharCo, User Rank: Petabyte Pathfinder
11/24/2012 | 4:26:14 AM

**Re: Bypassing the malpractice**
I would think that the bigger or more relevant Hadoop becomes, the bigger will be the need to introduce courses on the program to students. The earlier, the more comprehensive, the better.

Login

50%  50%

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

---

SharCo, User Rank: Petabyte Pathfinder
11/24/2012 | 4:24:55 AM

**Re: Bypassing the malpractice**
Misuse or impoper usage (or ignorant usage, if ever there is such a thing) are common things that happen when the IT tech at hand don't really have a clue what they're using or supposed to be doing with it. I agree that user training is key here; the team has to know what they're dealing with and what they can do to take full advantage of it so that their efforts and the software itself don't go to waste or get a bad name.

Login

50%  50%

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

---

legalcio, User Rank: Exabyte Executive
11/23/2012 | 9:28:46 PM

**Re: Bypassing the malpractice**
The industry is woefully behind on training professionals to step into companies and, if you will, become Hadoop experts, or at least data experts. Good blog today on Saas-in-the-Enterprise on Cloudera's attempt to finance the training of thousands of students in Hadoop, but it's going to take more than one vendor. When will universities start to offer a big data/Hadoop curriculum?

Login

50%  50%

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

---

Saul Sherry, User Rank: Blogger
11/23/2012 | 10:25:47 AM

**Re: Bypassing the malpractice**
That's the crux, and something shortsighted IT planners skip - Hadoop is free, until it starts costing you money in broken promises.

Login

50%  50%

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

---

kiran, User Rank: Petabyte Pathfinder
11/23/2012 | 10:22:34 AM

**Re: Bypassing the malpractice**
Yes Saul. It almost always comes down to how your end-user interacts with the system, which leads to user training. User training is the key, and in case of Hadoop it matters to an added extent. It's not rare to see application malpractices bringing organizational activities to a halt and causing huge financial losses.

Login

50%  50%

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

---

Saul Sherry, User Rank: Blogger
11/23/2012 | 9:31:35 AM

**Re: Bypassing the malpractice**
@Kiran - I think another element here is the 'Big Data Hype'. I know we've spent a lot of time pointing the finger at Big Data Hype (or BDH, if you will) in this community, but I feel Hadoop especially has been hard done by. It's a great solution, but only when used correctly - Robert here manages to spell out exactly where is is (and isn't) understood.

Login

50%  50%

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

---

PAGE 1 / 2  >  >>

---

**BIG DATA REPUBLIC**

ABOUT US  |  CONTACT US  |  HELP  |  REGISTER
TWITTER  |  FACEBOOK  |  LINKED IN  |  GOOGLE+  |  RSS

in partnership with **InformationWeek**   UBM Tech   powered by **UBM DEUSM**

---

bigdatarepublic : /bigdatarepublic/section/2635 : /bigdatarepublic/section/2635/254548

UBM TECH

OUR MARKETS: **Business Technology** | **Electronics** | **Game & App Development**

Working With Us: Advertising Contacts | Event Calendar | Tech Marketing Solutions | Corporate Site | Contact Us / Feedback