QUEEN'S UNIVERSITY IRC

Business Intelligence, Big Data, and HR

Ian Turnbull, BA, MBA
Laird & Greer Management Group

Our people are our most important asset, or so we hear, so data about those people – workers, or employees, if you prefer - should be central to our organization's total data set! To understand where HR data fits, you first have to understand your organization's overall data management strategy. How is data collected, organized, and managed? And how do you analyze that data to obtain information?

"Business Intelligence", the idea of transforming raw data into useful and actionable information, has become an oft-discussed concept. It allows management to gain historical insight and to produce predictive analytics for competitive advantage. And Business Intelligence arises directly from "Big Data", the process of bringing together raw data from multiple data sources into a single analytical tool. That tool can be used by management to produce Business Intelligence.

"Your recent Amazon purchases, Tweet score and location history makes you 23.5% welcome here." [1]

The next time that you use Google, or some other search engine, do a search for some unusual item; something that you haven't searched for before. Then spend some time on sites that you visit often.

You will notice that ads related to the unusual item will pop up beside your search results for the more common items. That is Big Data at work in a marketing context. Google has picked up your first search and is now displaying pages that its algorithm predicts will be of interest to you based on that search. And Google (and other providers) charge advertisers for this. It is the core of their economic model.

Multiple Data Sources

Why are there multiple data sources that should be amalgamated into Big Data? Historically, information technology (IT) systems have always been stand-alone; conceived, designed and built on the concept of functional data subsets representing core business processes: finance, marketing, sales, manufacturing, inventory control, etc.

This organization of data management has mirrored the organization structure as illustrated by the typical organization chart (figure 1).

---

1 Big Data graphic. (n.d.). Retrieved August 13, 2015, from
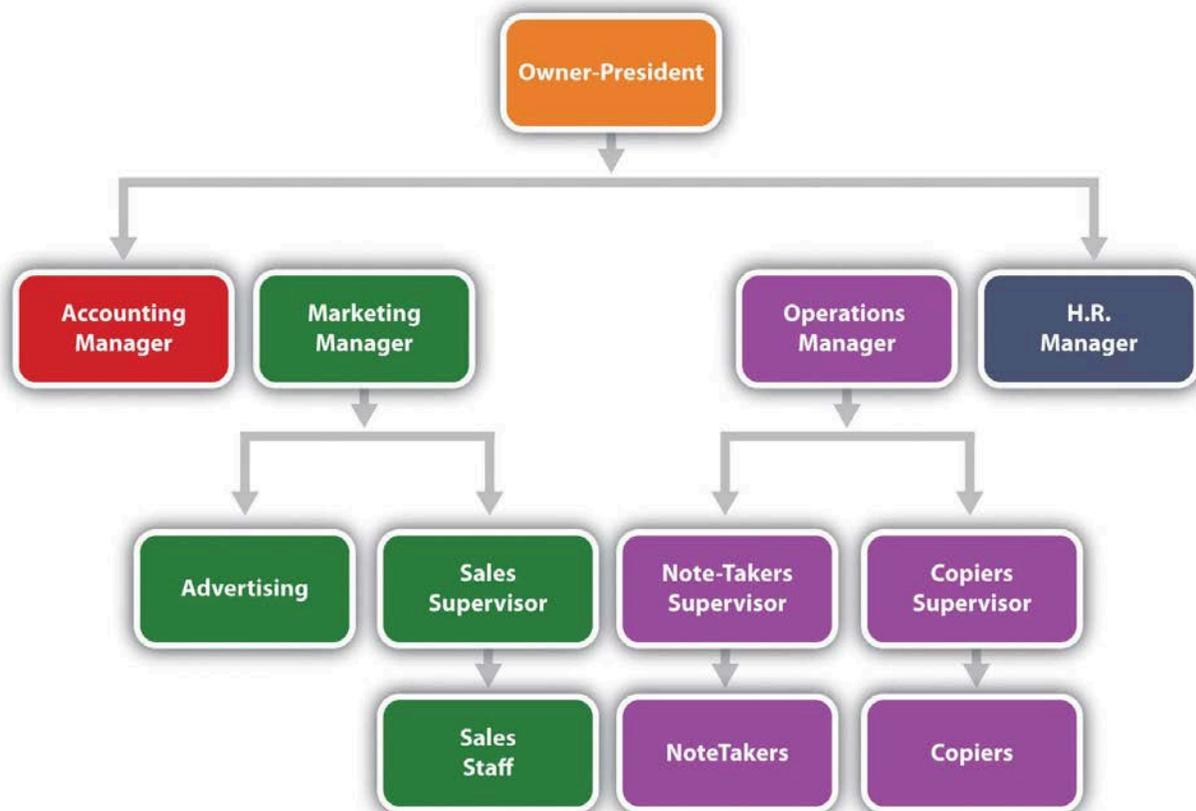https://upload.wikimedia.org/wikipedia/commons/b/b3/Big_data_cartoon_t_gregorius.jpg

Figure 1: Typical Organization Chart[2]

We have all heard the story of the CEO who asks both the CFO and Chief HR how many employees the organization has, only to be given two different answers. This frustrating result is because most finance and HR departments have different automated systems, AND because they have never agreed on how to define employees. Are part-timers and casuals counted as one each, or as a percentage of full-time equivalents (FTEs)? Do you count only active employees, and not those on leave without pay? Do contractors/ consultants count?

Interfaced Systems

These multiple data sources are typically because of the best-of-breed software philosophy – each function has a right to select the software that best meets their own functional needs – i.e., multiple systems from multiple vendors; marketing systems are used by marketing staff, finance systems by finance, and so on.

---

[2] Typical Organization Chart. (n.d.). Retrieved August 20, 2015, from http://images.flatworldknowledge.com/collins_2.0/collins_2.0-fig06_008.jpg.

In this model every system operates independently with its own tool sets, separate, and often different: data models, data element definitions, data capture, entry, storage, security, retention rules, standard and ad hoc reporting tools, user interfaces, security, and analytics tools.

As you might suspect, this approach has a strong tendency to discourage data sharing and co-operation. It also lends itself to capturing duplicate or supposedly similar data.

Integrated Systems

Integrated systems, often known as enterprise resource planning (ERP) software, offer a wide-range of functional systems ("modules" in this scenario) that use common report writers and other tools with their fully integrated set of applications.

The ERP Overview of the modules offered by SAP illustrates this:

Figure 2: ERP overview of the modules offered by SAP [3]

The theory is that the software developer, SAP in the example above, has designed the software to maximize effectiveness and efficiency, and that the conflicts seen between interfaced systems don't exist.

[3] SAP Modules Training. (n.d.). Retrieved August 11, 2015, from http://sapmodules.org/wp-content/uploads/2013/01/SAP-Modules.gif

Big Data

The term "big" means "BIG" – petabytes, not gigabytes. That definition automatically excludes all but the largest organizations; the smallest percentage of the business community. Medium and small organizations simply do not have that much data and so the term doesn't apply. But the concept of combining data from different parts of the organization and various external sources IS valid for every organization.

As valid as the idea of Big Data, or integrated data, may be, its foundation must be the existence of accurate, reliable data available to all who need it in a timely way. And you will find very few HR practitioners who believe that their organization has that foundation.

In almost all organizations the information tools for human resource data are a mish-mash of official automated systems, quasi-unofficial automated systems, spreadsheets, and various hard-copy documents, including notes in a supervisor's back pocket.
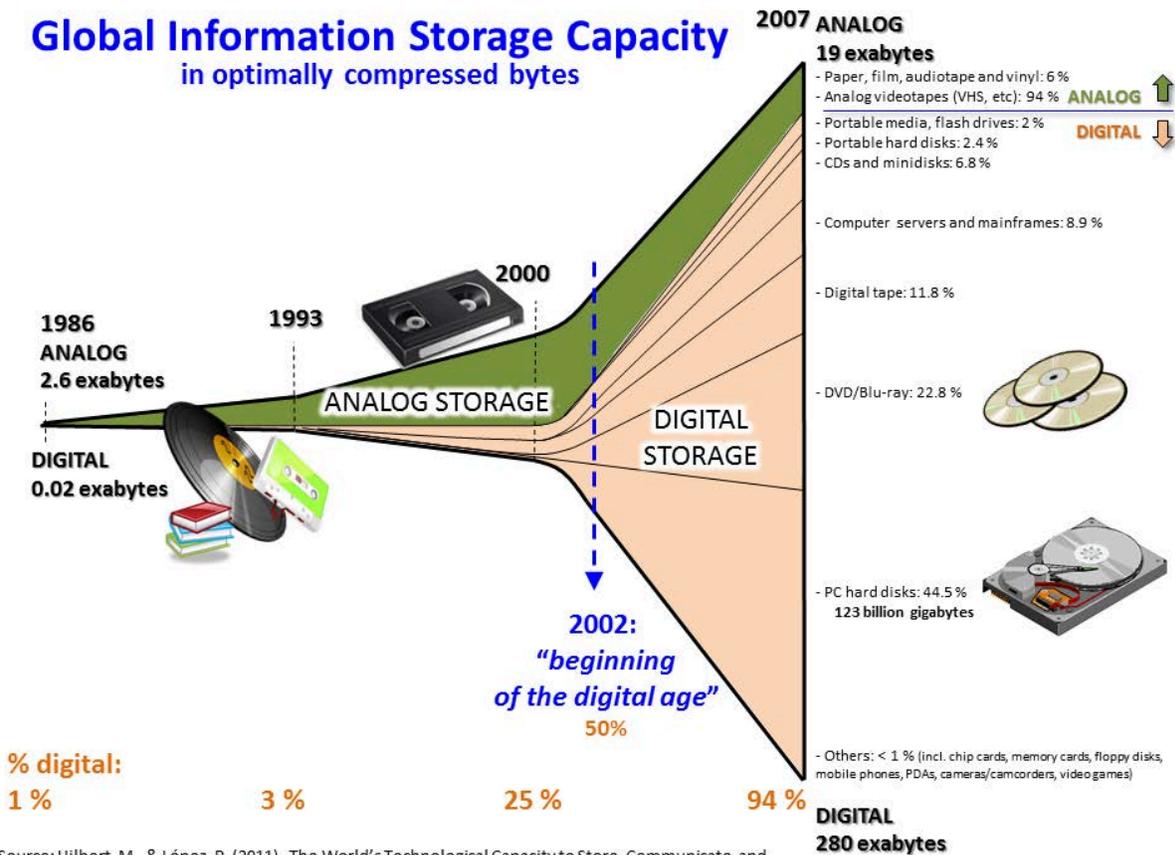
Figure 3: Global Information Storage Capacity[4]

---

[4] Global Information Storage Capacity. (n.d.). Retrieved August 11, 2015, from
https://upload.wikimedia.org/wikipedia/commons/7/7c/Hilbert_InfoGrowth.png

But the concept of integrated data is just as important (or more important?) for smaller scale organizations. The ability to analyze operations and sales, predict trends, compare to market intelligence, understand and manage labour distribution; all of these are essential to an organization's degree of success. This is completely analogous to the idea of business process engineering (or re-engineering). Business processes are normally NOT "designed". They happen through time as someone somewhere has a need for something to occur. At no time does anyone provide a complete overview by revisiting the process from the defined need to the final output.

The largest challenge of Big Data is that most organizations do not manage their data in a coordinated and centralized manner. The "library" is normally a department focused on external data, not internal. And the department known as information technology, or information systems has come to focus on the technological side of the equation, abandoning the data side to the functional departments.

But most functional departments have no documented data management plan, and even if they do, what department is responsible for coordinating the plan? The obvious answer is that inevitably there will be gaps, and overlaps, and issues of completeness, timeliness, and accuracy. You will discover, likely to your horror, how many versions of just one data element exist in your organization.

Every organization needs a structured plan of data ownership; e.g., every data element assigned to someone (or some unit) who is the custodian of it. The concept of custodianship is an important one, and IF the HR department is responsible for organization design, then the failure to document these roles is HR's to shoulder.

Without a clear custodian of a data element, who defines the parameters of the data? Which of the multiple system options represent THE data source; e.g., which system is the primary system for that data element? What system/s can update the data? With what frequency? With what edits/review? The importance of these data management tasks grows with every step towards Big Data.

Not only do these issues go to the challenges of quality, accuracy, and timeliness, it extends to the issues of data privacy. Proper data privacy requires clear guidelines regarding data gathering, retention and destruction, and most importantly, which users have the need and the right to access data about people.

What About HR?

The state of HR systems and HR data almost certainly mirrors the situation in the rest of your organization. There may be an HR/payroll/time management system (Human Resource Management System – or, HRMS), but most often there is a mix of single function systems of varying age and effectiveness, and no coordinated data map. Many organizations have a different system for each of applicant tracking, recruitment, training, basic HR admin, benefits, compensation, pension, labour relations, and so on.

Figure 4: Example of an Integrated Database[5]

[5] Example of an Integrated Database, Courtesy of DLGL Technologies Corporation, Blainville, Quebec.

A quick glance at the detailed functional map from DLGL Technologies demonstrates just how many functional elements there can be in a complete HR/ payroll/time management system.

Imagine each of these from a different vendor and consider the incredible task of trying to keep all of the technical AND functional elements in sync.

Use of a fully integrated HRMS – like DLGL (or others) – would be a huge step in moving towards big data capability because ALL worker-related data will be standardized and available. In fact, when compared to a mix of interfaced systems a well-integrated HRMS is big data – about people.

That is the challenge of system implementation – to work through all of the data and process complexities. Organizations that fail to do this, or worse, fail to maintain it, will discover that their HR data is not accurate, complete, or timely. Of course this frequent condition of human resource data is no secret to payroll. That is why so many payroll managers want their own data system separate from any HR database (and why one organization's payroll and HR systems duplicate most data; duplicate, but not identical).

Think about the most basic data; worker name. In most organizations this data is collected over and over again, some with middle initial, some with middle name, some shortening the hyphenated surname for lack of room, etc. If we can't even get ONE clear piece of data here, how can we even consider Big Data?

The Ideal Solution

Every organization system needs data about workers; each needs to track users and assign appropriate security. That alone suggests that the core HR system should be the foundation system for all system security. After all, it is designed to track worker movement from position to position, and therefore, from one user security profile to another.

And WHEN (not if) your organization creates a data custodian responsibility map, including data privacy, what better organization system to manage it than the HRMS? With front-office systems like HRMS and finance providing a well-organized foundation, the rest of the organization's information systems can be re-oriented to make Big Data and Business Intelligence a reality.

The Impact of the Cloud

Do you use Dropbox on your home computer to store pictures? Or maybe you are an Apple user and rely on iCloud? In either case you are a cloud user.[6]

What the heck is cloud computing, and does it impact on the foregoing discussion?

---

[6] The Impact of Cloud. (n.d.). A curated report by the Economist Intelligence Unit. Retrieved August 8, 2015, from http://www.economistinsights.com/sites/default/files/The%20impact%20of%20cloud.pdf

In the simplest terms, cloud computing means storing and accessing data and programs over the Internet instead of your computer's hard drive. The cloud is just a metaphor for the Internet. It goes back to the days of flowcharts and presentations that would represent the gigantic server-farm infrastructure of the Internet as nothing but a puffy, white cumulonimbus cloud, accepting connections and doling out information as it floats. [7]

Okay, but does it matter when discussing Big Data, Business Intelligence and HR?

As illustrated by the graphic, the storage of data will heavily shift to the Cloud. Instead of installing software on your organization's own servers the applications can also reside in the Cloud.

So yes, the Cloud has the potential to impact your organization, especially IT and its internal infrastructure, and make new organization structures possible – all topics for other articles.
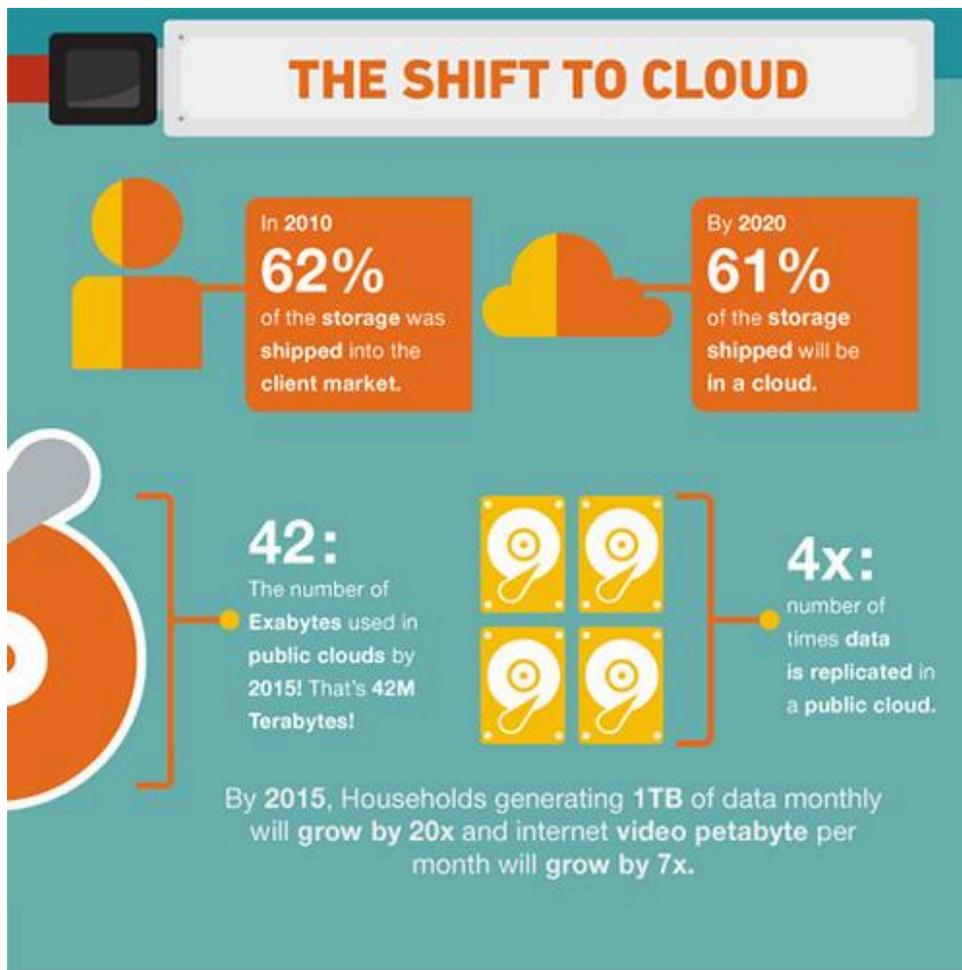

Figure 5: The Shift to the Cloud[8]

---

[7] Griffith, E. (2015, April 17). What Is Cloud Computing? Retrieved August 11, 2015, from http://www.pcmag.com/article2/0,2817,2372163,00.asp

[8] The Shift to the Cloud. (n.d.). Retrieved August 11, 2015, from http://cdn.phys.org/newman/gfx/news/2013/thedemandfor.jpg

But in considering Big Data and Business Intelligence, the Cloud is of secondary importance. Why? Because the challenges of Big Data and Business Intelligence are primarily organizational. It doesn't matter if two systems and the data that they contain are on your servers, or in the cloud. What matters is, have you done the analysis and data cleanup necessary to bring the data into alignment?

The Bottom Line

Regardless of an organization's size, integrated data management (Big Data), and analysis (Business Intelligence) are extremely worthy goals. They could be operationally critical for an organization's success or failure.

But if one of the core foundational systems – dealing with human resources – isn't managed properly, one of the chair legs could be significantly shorter than the others.

Most organizations manage data poorly. It is as if you put four different tire types and sizes on your car. Poor performance, wear, and safety are almost guaranteed. By treating your organization's data as a matched set, meaningful, actionable Business Intelligence is a certain outcome, and that will be a significant business advantage.

About the Author

Ian Turnbull is Managing Director of Laird & Greer Management Group Corp, specializing in HR, Payroll, and Time system selection and management, and a Director of Earning Through Learning. A former Chair of the Canadian Council of Human Resource Associations (CCHRA) and of the International Association of Human Resource Information Management (IHRIM) his latest book (Carswell 2014) is: *HR Manager's Guide to Managing Information Systems*. Ian has a BA and MBA from Western University (The University of Western Ontario) and obtained his professional human resource designation, the CHRP, in 1992.

References

Big Data graphic. (n.d.). Retrieved August 13, 2015, 150813 from:
https://upload.wikimedia.org/wikipedia/commons/b/b3/Big_data_cartoon_t_gregorius.jpg

Global Information Storage Capacity. (n.d.). Retrieved August 11, 2015, from
https://upload.wikimedia.org/wikipedia/commons/7/7c/Hilbert_InfoGrowth.png

Griffith, E. (2015, April 17). What Is Cloud Computing? Retrieved August 11, 2015, from
http://www.pcmag.com/article2/0,2817,2372163,00.asp

Hilbert, M. & Lopez, P. (2011). The World's Technological Capacity to Store, Communicate and
Compute Information. Science, 332(6025), 60-65. Retrieved August 11, 2015, from
http://www.martinhilbert.net/worldinfocapacity.html

Integrated Database. (n.d.) DLGL Technologies Corporation, Blainville, Quebec.

SAP Modules Training. (n.d.). Retrieved August 11, 2015, from15-08-11 from:
http://sapmodules.org/wp-content/uploads/2013/01/SAP-Modules.gif

The Impact of Cloud. (n.d.). A curated report by the Economist Intelligence Unit. Retrieved August
8, 2015, from
http://www.economistinsights.com/sites/default/files/The%20impact%20of%20cloud.pdf

The Shift to the Cloud. (n.d.). Retrieved August 11, 2015, from
http://cdn.phys.org/newman/gfx/news/2013/thedemandfor.jpg

Typical Organization Chart. (n.d.). Retrieved August 20, 2015, from
http://images.flatworldknowledge.com/collins_2.0/collins_2.0-fig06_008.jpg

Industrial Relations Centre (IRC)
Queen's University
Kingston, ON K7L 3N6
irc.queensu.ca