

HR data: Ready to be part of next 'big' thing?

The idea of "big data" is to gather input from multiple data sources into a single analytical tool. And the term big really does mean big — we're talking petabytes, not gigabytes.

That definition automatically excludes all but the largest organizations — which represent the smallest percentage of the business community. Medium-sized and small organizations simply do not have enough data for the term to apply. But the concept of integrated data is just as important (or perhaps even more important) for smaller-scale organizations.

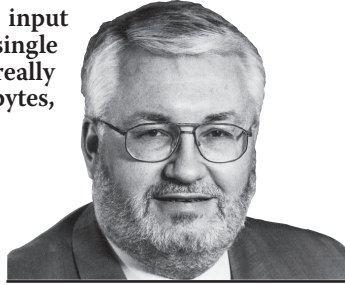
The ability to analyze operations and sales, predict trends, make comparisons with market intelligence, understand and manage labour distribution — all of these are essential to an organization's degree of success.

As valid as the idea of big data or integrated data may be, its foundation must include accurate, reliable data available to all who need it in a timely way. You will find very few human resources practitioners who believe their organization has that foundation.

In almost all organizations, the information tools for human resource data are a mishmash of official automated systems, quasi-unofficial automated systems, spreadsheets and various hard-copy documents, including notes in a supervisor's back pocket.

This is completely analogous to the idea of business process engineering (or re-engineering). Business processes are not normally "designed" — they emerge over time as someone somewhere has a need for something to occur. At no point does anyone provide a complete overview by revisiting the process from the defined need to the final output.

Why is data



Ian Turnbull
GUEST COMMENTARY

not managed properly?

When we look into a typical organization's human resource data, we usually discover that data is not managed. There are a number of reasons and explanations for this:

- There is no documented data management plan. This inevitably means there will be gaps and overlaps, plus issues of completeness, timeliness and accuracy. How many versions of one data element exist in a typical company?
- Whatever limited definition that does exist is contained in the design of the automated systems (almost always several unintegrated or poorly integrated systems). And even if every automated system has a complete and accurate data dictionary — and most don't — no one has ever done a thorough map of it all.
- There is no structured plan of data ownership. For example, every data element is assigned to someone (or some unit) who is the custodian of it. The concept of custodianship is an important one. Not only does it impact quality, accuracy and timeliness, it also extends to issues of data privacy. Who has the need and the right to access human resource data? In some organizations, the IT department takes that responsibility, but IT stands for information technology — they focus on

the technology, not managing the data.

- Without a clear custodian of a data element, who defines the parameters of the data? Which of the multiple system options represent the primary data source? What systems can update the data? With what frequency? With what edits or review?
- Back to data privacy for a moment — there are no clear guidelines regarding data retention and destruction. And beyond very high-level security issues, no one is looking at how data can be used effectively while respecting individual privacy.

The sad state of affairs outlined above is almost always true, even when an organization has recently implemented a single HR-wide information system.

Issues of data source and integration were likely considered during implementation but it is equally likely the work was less than complete (pressure on timely delivery and budget) and almost certainly not maintained.

Payroll already knows the score

Of course, this 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 (though the duplication is not exactly the same).

Think about the most basic data element — a worker's name. In most organizations, this data is collected over and over again, some with a middle initial, some with a middle name, some concatenating the hyphenated surname for lack of room, and so on.

If we can't even get one clear piece of data here, how can we ever consider big data?

We in human resources (and

also lawyers, arbitrators and legislators) tend to think and speak of the employee file. We all know that in most organizations there is no such thing — instead, there are paper and electronic files resident in every HR functional and operational area.

Co-ordinating 2 databases

But what about an organization that has a large operational department, with its own HR systems totally separate from the official HR department's HR system?

There is absolutely no co-ordination of the two databases. Much of the data exists in both places but is not exactly the same, and gaps certainly exist. OK, now hand over "the file."

Integrated data management and analysis is an extremely worthy goal. It 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.

While I am particularly concerned about human resources data, I want to emphasize that much of the foregoing is equally true for all functional areas. The management of data — from definition to collection, from retention to usage guidelines and destruction — is very often a process that falls through the cracks.

Ian Turnbull is managing director of Toronto-based Laird & Greer Management Group, specializing in HR, payroll and time system selection and management. He is the author of the HR Manager's Guide to Managing Information Systems, published in 2014 by Carswell. He can be reached at (416) 618-0052, iturnbull@lairdandgreer.com or on Skype: ianturn.