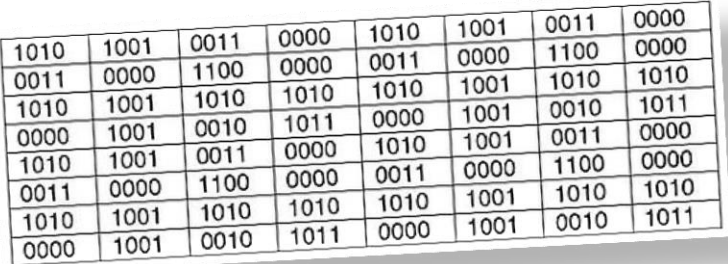# Big data small?

John de Croon   *21 March 2014*

Once when it was still cold, we wrote about how you can take decisions with little information. Then we particularly focused on how to maximally utilize the available data. To deliver more performance with the assets with the same asset management budget generally more and better data is required. Is it then useful to immediately go for a 'big data' initiative, or would you better start small?

figure: Wikipedia commons

For every decision, it comes down it is a sensible decision when the benefits exceed the costs. This also applies to information: when the cost of obtaining the extra information are exceeded by the value of the additional performance obtained (which can be the value of the decreased risk position as well), it makes sense to invest in it. And of course you should calculate with the net present value. We make an attempt and we first look at the cost.

The purchasing cost of data carriers decrease rapidly. In 1980 you had to pay about 90 euros for a Megabyte, now we get 2 Terabyte for it[1]. So this will make you happy as asset manager you would say. However, this does have a downside. When the asset manager for instance wants to compare maintenance data of assets (e.g. preventive maintenance costs, timing of failures and associated costs) with actual production data then usually a link between two systems is required (the ERP or EAM system with the production system[2]). The production data contains a magnitude of a few Terabytes per year, so approximately 200 euros on storage costs. Quite often the link between the information systems in cases does not yet exist. This means that including design, build and test it could cost an additional € 100k for just a link. In addition to this roughly 20% of maintenance costs per year needs to be paid. If you're unlucky, for the import of the production of data in the ERP / EAM system customization is needed and again some € 100k is required. Then there is an upgrade several years afterward, and the customization must be adjusted because otherwise the system does not work anymore. So on the cost side, for a period of five years we are talking about a few hundred euros in storage costs plus some € 400k to IT integration and maintenance of the system. And we have not even talked about a possible business warehouse implementation of at least € 500k in which all kinds of sophisticated cross-section analysis can be made. If your ambition is to apply for an ISO 55001 certification, paragraph 7.6 of the standard sets requirements for the maintenance of all kinds of data. That maintenance costs money as well and one can become disappointed. But do not worry, because we try to quantify the benefits of information and if the costs are exceed by the benefits that provides opportunities. However, there are some pitfalls.

Suppose you have pumps which no longer meet the required yield. You found it out during an inspection and it was just in time. For these kind of situations defining the benefits is not that hard. There is no need any more for physical inspections (efficiency gains) and when the aging process is not entirely known, with a periodic inspection regime you face the risk you are too late and the pump crashes if the output is not measured periodically. Through an automated link where the output is

---

[1] http://nl.wikipedia.org/wiki/Harde_schijf

[2] ERP: http://en.wikipedia.org/wiki/Enterprise_resource_planning EAM: Enterprise Asset Management Systems. Maintenance Management systems which also often support related processes like (maintenance) purchasing and inventory management.

measured, an early detection is possible. The replacement can be done on time and failure costs (repair costs and loss of income) are prevented. When failure data is known, then with a probability calculation the benefits can be calculated and it can be determined whether the investment is worthwhile.

It becomes more difficult when we have several options to avoid a jamming pump. As a cheap option the impellers can be substituted by imitation parts, or more can be paid to replace the old ones by ones that are provided by the original manufacturer of the pump. The question then is whether it is likely for the cheap option you do not meet the required output, and if so how likely it is. If you have information available to calculate the trade-off between risk and benefits than making the choice is not so difficult. Is that information not directly available but that it can be made available, then the value of the information can be calculated[3]. However calculating the value of information is quite difficult and an additional problem is that you often only know what that value really was afterwards. This also means that you first create a business case for obtaining the required information, which in turn is part of the decision on how to deal with the pumps themselves. It all does not make it easier...... Also with more data we have the probability that time is wasted on non-relevant data. More data is not always better. If you already are capable to determine the value of the information for the impellers, then a firmly debate will pop up whether the savings are caused by the improved information, or by the measures itself to reduce the risk. After all the information itself does not reduce the risk.

Let us go back to the cost. We begin with two Terabytes of data for 90 euros and if we are not careful, we end up with a difficult business case that causes a lot of discussion.

What should be started is to ask the right questions: what is the problem and how big it is. If we want to solve the problem, we will have to consider what data is required. This process looks suspiciously like asset management decision-making, and that it is indeed!

Keep taking it in mind how much time is invested in the business case. Our advice: if possible for information provisioning do not go too much into depth in determining financial benefits on detailed asset level since this leads to nowhere as shown above. Though everything is quantifiable, basically you only need to do it in the right context. With information systems this does not succeed at individual asset level, but it usually does on system level. So you determine the value of your information policy and from that you derive what data is needed. Note that the implementation and maintenance of additional information can be 500 times higher than the cost of storage. So start small if possible and do not right start immediately with a big data initiative!

John de Croon is partner at AssetResolutions BV, a company he co-founded with Ype Wijnia. In turn, they give their vision on an aspect of asset management in a weekly column. The columns are published on the website of AssetResolutions, www.assetresolutions.nl/en/column

---

[3] For an example visit http://en.wikipedia.org/wiki/Expected_value_of_perfect_information