

# Data Collections and Ethics



For some time, people have been aware of their social responsibility, not only personally but also professionally. We have become aware that we have to deal in a proper way with the privacy-sensitive data of our fellow men. The computer applications with which we capture data from a wide variety of transactions and

processes are provided with advanced security. In addition, all actions that users perform are recorded in log files. Apart from capturing, access to these systems is also subject to authorization. Our data collections are well protected against unauthorized use. Well organized, right?

And then the need arose to combine the data from the various data collections – or source systems – in order to squeeze even more knowledge and more information from that data. Data Warehouse and Business Intelligence reporting and analysis systems were introduced. There is only a select group of experts who work in these environments. The reporting and analysis applications provided are usually well protected and access is only available to authorized users. Again, at first sight, well organized too.

But one day DWH/BI specialists started thinking about what was going on.

They realized that there is no limit to the transfer of data from various source systems into the data warehouse, and that they have unlimited access to all the data. DWH/BI specialists bear the great responsibility of dealing with the data in a 'proper' way. DWH/BI specialists may feel particularly uncomfortable because they have unrestricted access to highly sensitive information. Of course they are aware that this data may not be available to everyone. Due to the large amount of data and the particular technical aspects of the processes, DWH/BI specialists will not really read all the data itself, but it can be done! There is a big difference between, on the one hand, developing applications for capturing data and, on the other hand, making provisions for consolidated storage, reporting and analysis capabilities.

Unlike the application developer, the DWH/BI developer works with 'real' data. Thus, the data itself makes the difference and it is there that the great responsibility lies.

As long as data remains within the company, a confidentiality agreement is the company's legal safeguard. But what does the DWH/BI specialist do when a command is given to release data to a third party? What are the guidelines for this and what is allowed to be released and what not? According to current legislation, a company fulfils its legal privacy obligations by appointing a privacy officer. Within the company, there might be a privacy committee present to ensure that the applications that capture the company's operations and processes are sufficiently and properly authorized. But in many cases it can be overlooked, or people are not adequately aware that a lot of data in the data warehouse can be combined. In this way, the question of 'what is the right thing to do' often remains the responsibility of the DWH / BI specialist.

The question of 'what is the right thing to do' is a question that ethics does not want to answer; instead it only provides us with considerations and trade-offs. Ethics is a branch of philosophy that for time immemorial has asked people the question 'what is the right thing to do?' There's been a lot written about this subject, so we do not need to think all by ourselves. One problem is that there are different schools of philosophers who approach this question in different ways. These schools of thought can be broadly divided into two streams. In one school of thought, people look at the consequences that an act entails: consequentialism. The consequences of an act can be assessed on a scale of the greatest good for the greatest number of people. Translated into a BI example: most people like it when a supermarket sends us a shopping list based on our recent buying habits. Evidently, capturing the buying behavior of customers is ethically permissible. The other major philosophical stream is duty ethics. This school assesses if the act is in itself good by applying the popular saying: don't do to someone else anything you do not want to be done to you. We do not want everything about us to be registered as a generally accepted law. Translated into our BI example: we do not like the idea that the supermarket remembers what we bought. This makes us feel less free, we cannot reflect properly on our own choices. Not because the supermarket forces us to do something, but because we feel observed.

Whether it is ethical or not, Google remembers the search strings we use for a long time. Yet once again it looks as if "Google is your friend".

If you search for it, it turns out that there are a lot of articles on BI and ethics on the internet. But every outline of the answer also delivers new questions. Should we find the answer to the question 'the right thing to do' in Computer Science or is it the responsibility of the Business? What do we think of 'big data' when it comes to everything that is recorded through the internet, and we're not even talking about the "Internet of Things" .....

This article is certainly not meant to be an extensive philosophical essay. Rather, it's meant to be an impetus to thinking about a directive that can be used by BI/DWH specialists.

Undoubtedly, there are BI specialists who say that they only run commands and who do not worry about whether this is okay or not, because after all this is part of the job. But there are also numerous BI thought leaders who make statements about the responsible actions of the BI profession. In order to begin a search for what has been written about BI and ethics, the following is a possible series of addresses that are worth visiting.

#### Ethics General

Ethics timeline <http://www.ethics.org/resource/business-ethics-timeline>  
Wikipedia Ethics [http://en.wikipedia.org/wiki/Ethics#Modern\\_normative\\_ethics](http://en.wikipedia.org/wiki/Ethics#Modern_normative_ethics)  
A Framework for Universal Principles of Ethics <http://www.ethics.ubc.ca/papers/invited/colero.html>

#### Computer Science and Ethics

International Review of Information Ethics <http://www.i-r-i-e.net/archive.htm>  
ACM Code of Ethics and Professional Conduct <http://www.acm.org/about/code-of-ethics>  
Computers & Society: Computer Ethics Links <http://www.sigcas.org/ethics>  
Ethics in Computing <http://ethics.csc.ncsu.edu/abuse/>  
International Center for Information Ethics <http://icie.zkm.de/>

#### Business Intelligence and Ethics

Exploring Ethics in Business Intelligence <http://www.b-eye-network.com/view/8506>  
Ethics in Business Intelligence <http://a-erickson.com/ethics-in-business-intelligence/>  
Data ethics <http://www.computerweekly.com/news/2240162744/Data-ethics-Author-warns-of-ethical-pitfalls-of-data-collection>  
BI Ethics (Richard Hackathorn) <http://www.bolder.com/ethics/>

#### Business and Ethics

An Ethics Toolkit for Managers <http://managementhelp.org/businessethics/ethics-guide.htm#anchor26548>  
Business Ethics <http://www.businessethics.ca/>

#### Data

Big data as a civil rights issue of our generation <http://radar.oreilly.com/2012/08/big-data-is-our-generations-civil-rights-issue-and-we-dont-know-it.html>  
Big data ethics expert <http://www.informationweek.com/global-cio/interviews/why-youll-need-a-big-data-ethics-expert/240145500>  
The Internet of Things <http://www.huntonprivacyblog.com/2013/04/articles/ftc-seeks-input-on-the-internet-of-things/>  
Personal data <http://www.details.com/culture-trends/critical-eye/201305/sharing-biodata-on-apps-and-devices#ixzz2S3Gt5zqt>  
or simply search for 'data ethics principles'

In addition to the above (far from complete) internet resources, Frank Buytendijk has written a series of articles which appeared on his BeyeNETWORK blog 'In Search of Wisdom' <http://www.b-eye-network.com/channels/5567/>

He has also included an extensive reading list. This series of articles was also published in book form: *Socrates Reloaded: The Case for Ethics in Business & Technology*

How many DWH/BI specialists are considering the ethical aspects of their work?

It is certain that answers cannot be found overnight, but perhaps a significant step is taken when, together, we start thinking and encouraging our companies and organizations to continue to think.