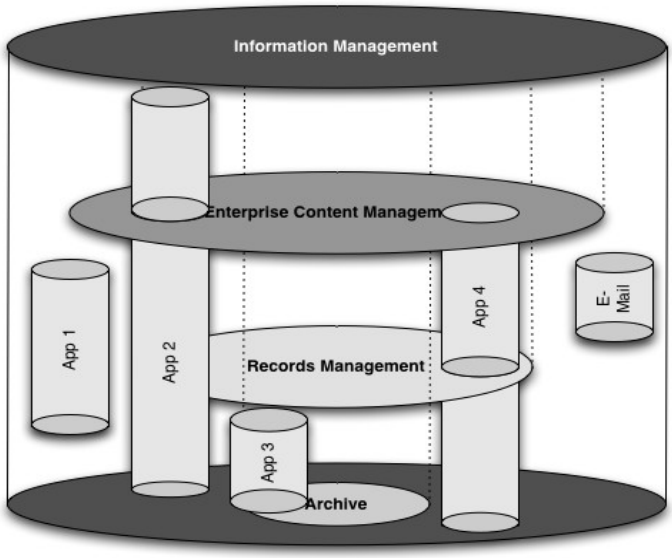
RECORDS MANAGEMENT – HOW TO COPE WITH DATA DELUGE

The traditional separation between archived data and productive data seems outdated – but how can an integrated approach work?

Dr. Bruno Wildhaber LLD, CISA, CISM, CGEIT

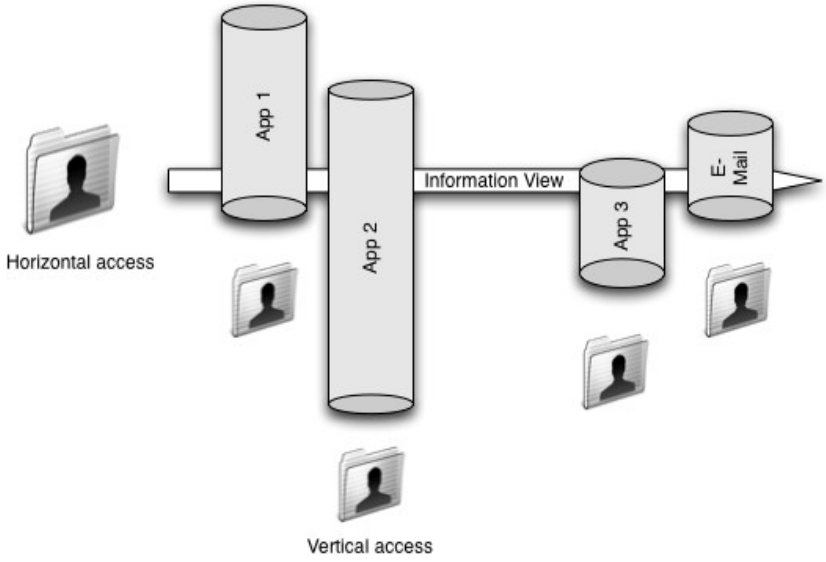
The increased importance of information and data processing and the virtualization of many vital business processes has produced an immense amount of data. Masses of data which need to be managed in every organisation – but how to approach this challenge? In the last few years, several new disciplines have popped up. Buzzwords like Enterprise Content Management (ECM) , Content Management (CM), Document Management (DM) Knowledge Management and Records Management (RM) even BPI (Business Process Integration) pop up when talking about managing data. They have different names and have different meanings, but aim at the same spot: A better management of documents, data and information in your organisation.

The following picture shows the different layers of data management in an organisation:



The entry level for most organisations is the bottom or inner layer - the compliance oriented archive level. Archiving has been done for many centuries, it is a historical discipline which does not require much explanation. Most organisation have an internal organisation and processes to archive physical = paper data. However, most organisations are helpless when it comes to cope with "data deluge".

The fact that storage requirements increase permanently leads to the procurement of more and more storage hardware. But non-organised data storage leads to something we call the "digital landfill". Data which ends on thousands of file servers or other storage systems – valuable data which is normally lost for future use because cannot be accessed – because it cannot be found. Data in organisations is produced in so-called "stovepipes", applications which have no interconnection and which produce enormous amounts of data. But useful information can only be produced if the right data combination can be found:



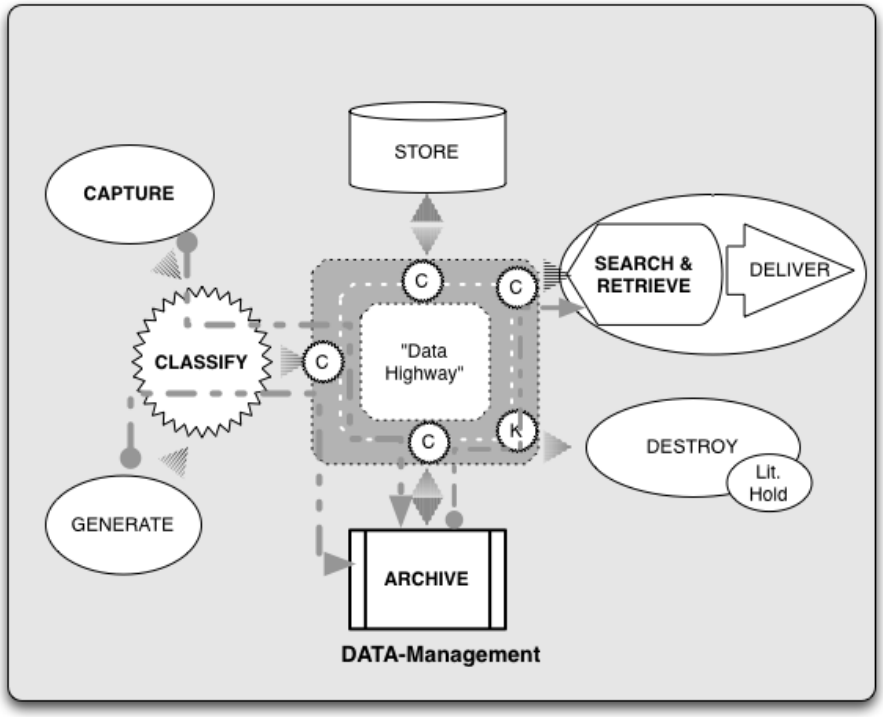
What is required to allow this horizontal access and to produce useful data collections? In most cases, this business requirement is successfully dealt with in traditional archiving systems. Data is collected from different applications, logically assembled and stored in a separate storage facility. But can be done logically, what has been for ages in the physical world? Traditional archives basically satisfy the lawyers, auditors and historians. Records Management (RM) basically addresses these target groups and tries to find answers how to retain information in different forms in a way, that it can be retrieved years later – complete and secured and in an efficient and effective way.

Data and information are only of value to an organisation and its stakeholders, when they have been managed carefully. This can only be done if the life-cycle of the data is known. All data which is created in an organisation follows a specific life-cycle. Unstructured data such as a word document or an Excel calculation is created by an individual, stored on a storage device, changed and at a certain point in time it is not used any longer and can be deleted. The intensive usage phase is very short. Most documents will be actively used for one week to one month, after this phase, they are more or less inactive. In most cases, they will remain stored until somebody dares to delete them or they have to be moved to a legal archive. The latter requires that already at creation of the document, a classification process has taken place. For every document type, this business classification has to be done.

In the traditional sense this was done by creating retention lists or schedules, today experts are using more and more the expression "taxonomy" which is based on the idea that also organisations can be structured based on their processes and business. However, these schemes need to be long-term oriented. Imagine that in 15 years from now somebody has to search for a client folder in a specific context. All organisational schemes will fail because organisation will change x-times, only business oriented models will allow to find the information.

One of the most important aspects about Records Management or ECM is the fact, that organisations will not be able to store the amount of data they actually produce. Data growth outnumbers storage space. Even more dramatic, the cost per stored GB will increase, not decrease. Already today, a hardware storage cost of 10 Eurocent/GB must be compared to the management and handling cost of this GB which will be around € 3000. Or in other words, only organisations which will be able to delete and destroy data will be able to reduce operating cost.

A combined and integrated ECM model could look as follows:



Such an integrated model allows for the access of all data needed, independent of where it is stored. The model also makes clear, that the true value of an ECM system lies in the search and retrieval process (S&R), not in the storage process, this is the easy part!

It is important to stress out that it is impossible to classify data after creation or to rely on search engines. In most cases, it is vital that the integrity of the information you are looking for has been protected. Using a search engine might deliver 80 pages of a hundred page document, but it might also deliver 100 pages of hundred, but in the wrong version. In the future, the delivery of complete sets of data will be a crucial factor in legal cases. The US already today applies E-Discovery laws which require the delivery of a complete set of data to the counter-party in a pre trial procedure. In addition, it will be impossible to destroy the data which has to be destroyed by legal reasons or for your own protection.

The downside of an efficient ECM system will be the exposure to claims from the data protection advocates. Concerns which are justified because as soon as personal data is store, a combination of this data from different sources could, if misused, become a serious threat to the privacy of an individual. Another reason why it is absolutely vital to exactly know which data is stored where, in what form and when it will be destroyed.

Of course the big benefit for the organisation lies in the fact that the production factor information can be used more efficiently and this should enable organisation to make use of their "hidden gold" – know-how. Especially knowledge driven industries will profit strongly from a Records Management and its future developments.

However, it is recommended to start with Records Management as the bases for all other activities, because RM is a must for every organisation. Organisations which struggle with a a lot of regulations will definitely want to go this way. However, it makes sense to achieve higher goals than just compliance. But be aware that most organisations are on a very low maturity level when talking about RM. Some have not even reached the status where we could say that they have identified the problem and the need for action!

There is no silver bullet what the implementation or technical solutions is concerned. A heterogeneous environment demands other solutions than a one solution shop, organisations which process high amounts of structured data will have to find other solutions than organisations which produce unstructured data only.

Wildhaber Consulting
Postfach 115
CH 8603 Schwerzenbach
Switzerland

Tel. +41 44 826 21 21
Fax. +41 44 825 31 60

info@wildhaber.com / www.wildhaber.com