# THE BUSINESS SHOULD CARE ABOUT DIGITAL E-WASTE

Nick Bakker
Senior Business Consultant
Rendabel Groen
Nick.bakker@rendabelgroel.nl

## Table of Contents

Introduction ........ 3
What is digital e-waste? ........ 3
What's the cause of digital e-waste? ........ 5
The volume of digital data is growing fast these years. ........ 5
The average life-time of data is becoming shorter ........ 6
Digital data is easily created and duplicated ........ 7
The average size of document is becoming larger ........ 7
Why should we care? ........ 8
Increasing Total cost of storage ownership ........ 8
Degrade system performance ........ 9
Business continuity, due to back up and restore problems ........ 9
Security issues ........ 9
Ecologic footprint of the datacenter ........ 9
Who should care? ........ 11
What can the business do? ........ 12
Reduce ........ 13
Reuse ........ 14
Recycle ........ 14
Recover ........ 14
Dispose ........ 14
How can the IT department support the business? ........ 15
Capacity management - reporting ........ 15
Financial management ........ 15
Demand management ........ 15
Conclusion ........ 16
Appendix ........ 17

## Introduction

With the current exponential growth of information, digital e-waste—the bits that are no longer wanted—is getting more attention. The rapid increase of digital e-waste is paired with several issues, for example:

- Increased TCO of storage
- Decreased performance
- Business continuity, due to backup problems
- Security issues
- Ecologic footprint of the data center

The problem of digital e-waste didn't exist previously. IT managers just bought bigger boxes to solve the issue, archived all the data on tape, or simply deleted files when there was no legal or business reason to store the data.

New technology has emerged to reduce the storage volume, for example compression and deduplication methods. However, as the digital universe continues its rapid growth, these methods will not be sufficient to stop the growing need for additional storage and mitigate the above mentioned issues.

Because data is an important corporate asset, the business should be made aware and held responsible to address this problem. Though not yet comprehensive, data governance and information management literature have begun to cover the nascent issue of digital e-waste management.

In this article, I will answer the following questions:

- What is digital e-waste?
- Why should we care?
- Who should care?
- What can the business do?
- How can the IT department support the business?

## What is digital e-waste?

There are many definitions of waste. However, all definitions have some things in common. Waste is something that is discarded and not wanted anymore. It's a dynamic concept; the same object can be waste or non-waste, varying from the user perspective.

Waste, like beauty, is in the eye of the beholder[10]. Data that became useless for a business process might still have value when used for data analytics purposes. Even machine data can have business value and therefore should not be thrown away without consideration.

Following the Waste Theory of Pongrasz a.o[8] digital e-waste is defined as: data that is in the given time and place, in its actual structure and state, not useful to its owner or has an output that has no owner, and no purpose.

This definition of waste is derived from the object oriented modeling language PSSP™, which uses the concept that every physical object can be characterized by four attributes: Purpose, State, Structure, and Performance.

Phoyola and Transkanen[2] used these attributes to classify waste into four waste categories. This classification is useful to define different waste management approaches, as will be explained later. This classification can be applied to define four classes of digital e-waste as well. The four waste categories and their equivalents for digital e-waste are presented in the table below[1].

| Class | General description | Digital E-waste classes |
| --- | --- | --- |
| Class 1 | Non wanted things, creation not intended, or not avoided, with no purpose | Unintentional data |
| Class 2 | Things that were given a finite purpose, thus destined to become useless after fulfilling it | Used data |
| Class 3 | Things with well-defined purpose, but their performance ceased being acceptable due to a flaw in their structure or state | Degraded data |
| Class 4 | Things with well-defined purpose, and acceptable performance, but their users failed to use them for their intended purpose | Unwanted data |

**Table 1: Classification of waste**

In this article we will focus on class 2, 3, and 4, as the story is told from a business perspective. Unintentional data is created by the IT department only.

## What's the cause of digital e-waste?

Digital e-waste is caused by the tremendous growth of the digital universe due to the shortening of the average lifetime of data and the increasing ease of creation and duplication of digital data.

**Volume of digital data is growing fast**

Digital e-waste is digital data at the end of its lifecycle. So as the digital universe is exploding, digital e-waste grows with the same factor. A recent study by IDC[6] forcasts the digital universe will reach 40 zettabytes (ZB) by 2020. A predicted volume, that is about 50 times as large as the situation in 2010.

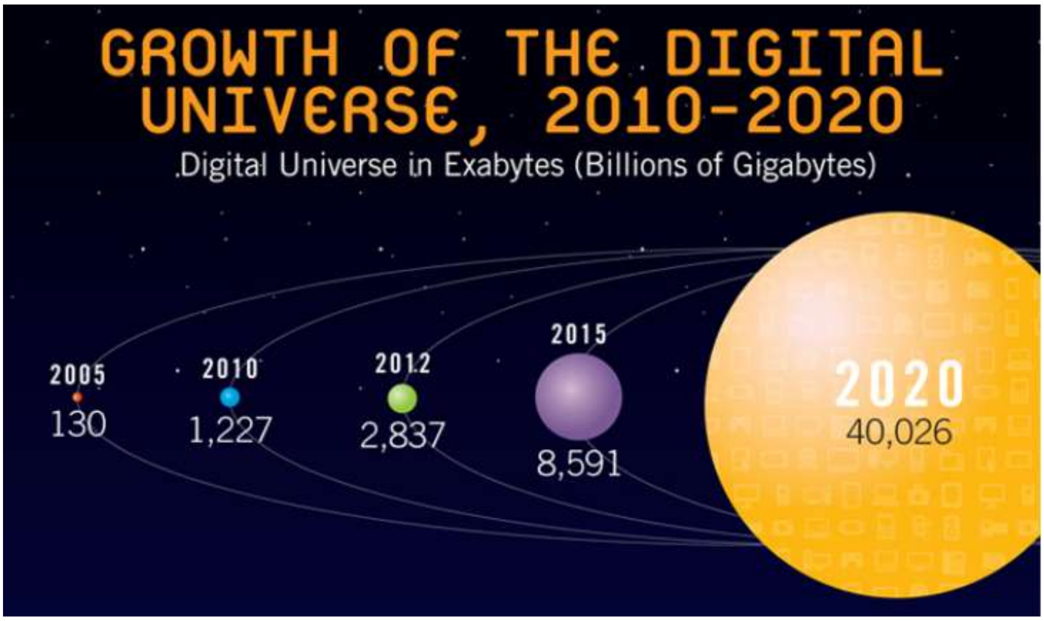


**Figure 1: Growth of the digital universe (IDC Annual Digital study 2012, sponsored by EMC)**

The IDC Digital Universe study of 2012 also revealed that only 1% of digital data is analyzed and 20% is protected. In 2007, IDC reported that only 10% of the stored data was perceived as useful by organizations.

Based on these results, we could formulate the hypothesis that over 50% of all digital data can be considered waste.

New technologies will increase the volume of digital data that is analyzed. It is common that not all data which should be protected is properly protected. It's widely known amongst storage capacity managers that most of the data is created only once and never accessed again. The increasing use of spreadsheets is a good example.

**Average lifetime of data is becoming shorter**

Purpose—or the lack of it—is one of the key attributes to define waste. Information Lifecycle Management (ILM) is based on the concept that data should be stored on media according to its business value. Several studies have been conducted to calculate the potential benefits of the ILM approach.

An example is a 2004 study by Berkeley University. It revealed that the possibility of reuse of stored data declined to 50% after just three days. After 60 days, this possibility is decreased to a mere few percent.

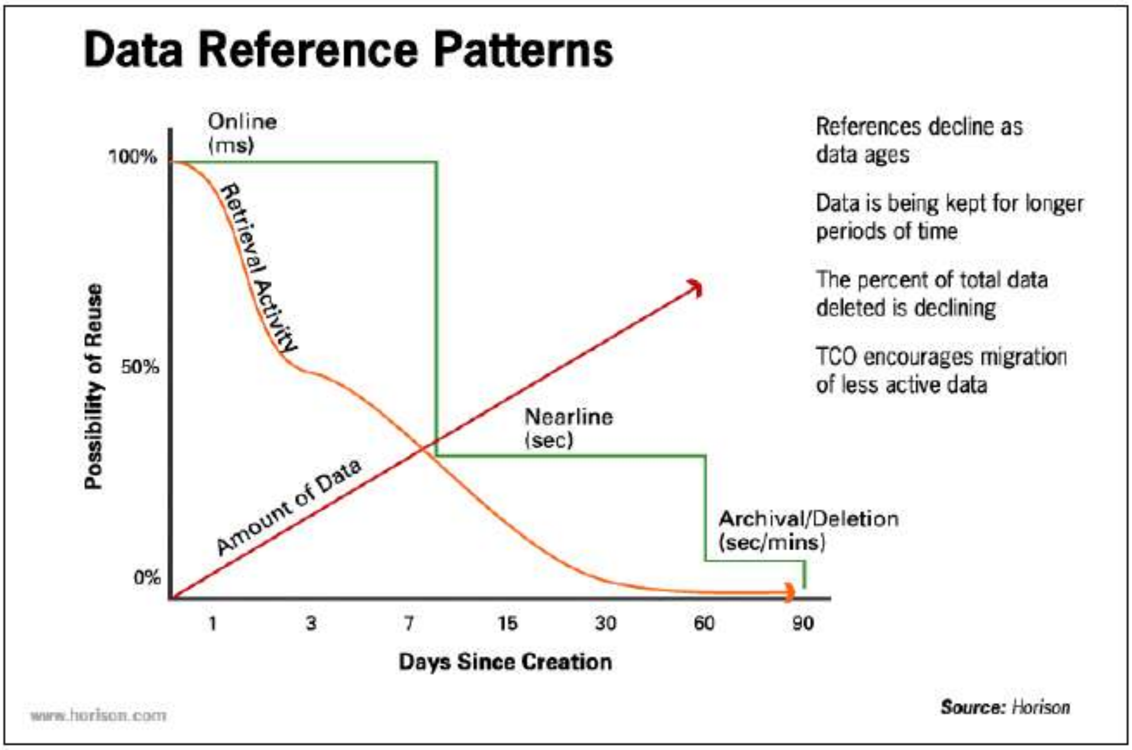


**Figure 2: Graphical presentation, Horison study 2004**[8]

These days, the average lifetime of a bit continues to shorten. Most data growth comes from unstructured data, such as tweets, pictures, movies, etc. as shown in Figure 3. Every data created in turn creates multifold data about that data by logfiles.

For most users, the average lifetime of a tweet is a fraction of a second. The same goes for most Facebook messages. Another example is the bulk of sensor data being created by an increasing number of devices. Only a small percentage delivers value to its users.

Big Data analytics tools can transform used data into information, but in a volatile world, the historic period is limited. Big Data algorithms are applied to parsed data, a subset of the original data volume.

### Digital data is easily created and duplicated

As shown in Figure 3, data volume is driven by new applications. Social media and the Internet, make it easy to create and distribute digital data. Multiple copies are made and duplicated in a split second.

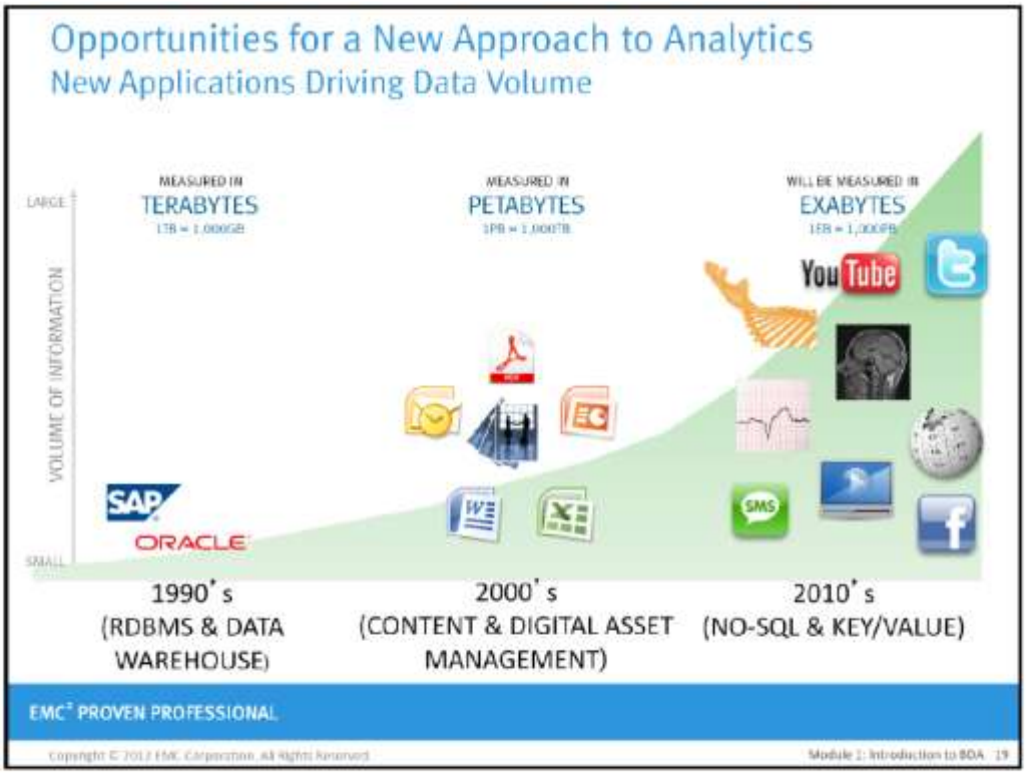


**Figure 3: New applications driving data volume**

For instance, how many followers do you have on the different social media platforms? You may be surprised at how many copies of the original version are made and stored on a huge number of devices. As technology development will never end, it's not unreasonable to predict that digital e-waste will continue to grow in the coming years.

### Average document size is becoming larger

Most stored bytes are representations of unstructured data files. In a click, pictures and movies are added to a document. Today, reports are made by putting colour presentation slides together. Pictures and movies are replacing words in documents, resulting in rapid growth of the average document size.

## Why should we care?

For citizens, garbage collection is a discreet service. Household waste simply disappears by truck once a week. But, landfills have finite capacity, as the people living in Naples, Italy noticed at the end of the 90's[5]. Local government had decided to use the slope of Mount Vesuvius as a new area to enlarge the landfills. Local protest by both citizens and trash delivery firms made it clear that creating new landfills is not a structural solution and even worse will create serious health issues for people living in the neighborhood.

Similarly, this applies to digital e-waste; cleanup performed by the IT department is invisible to the end user. Users don't seem to care until they are presented with a bill, or experience serious business issues, e.g. decreasing system performance facing continuity, security incidents, or environmental impact. These issues are addressed in more detail below.

### Increasing total cost of storage ownership

The myth of decreasing storage costs is still alive. Indeed, the price of hardware per GB lowers every day, but unfortunately hardware is only 1/3 of the total cost of ownership. As shown in Figure 4, the purchase (purple) part is a relatively small part of the total stack.

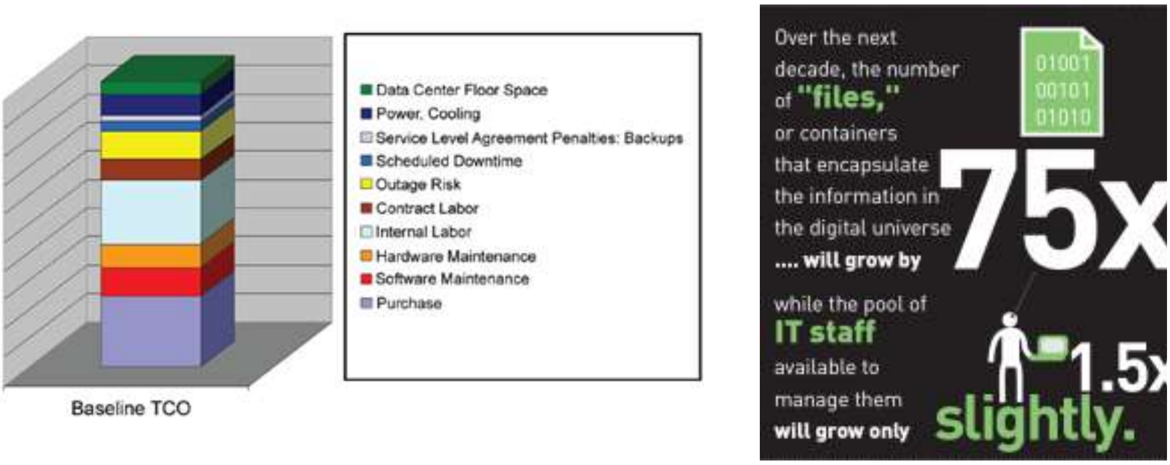


**Figure 4: TCO of storage baseline**

More important is the fact that other component costs do rise these days; especially the cost of power and the risk of outage as data is the new lifeblood of companies. Also, internal and contract labor cost will probably rise due to the scarcity of IT staff with the appropriate skills[6].

**Degraded system performance**

The presence of waste data may degrade system performance. If the files which are accessed frequently can be separated from those rarely used, system performance is improved. Obviously, the average speed of access in the database depends on the number of records which have to be processed. Waste will also result in slower performance.

**Business continuity due to back up and restore problems**

Waste data is often included in the standard backup scheme of an organization. As the volume grows, serious issues may arise to meet service levels concerning the backup window and Restore Time Objective (RTO). Unavailability of business data has a negative impact on the quality and output of business processes.

**Security issues**

Data which is considered waste by an intended user, might have value for another user. For example, private sensitive data may be useless for most users, but may have some value for commercial organizations. As a consequence, lack of control about waste data can lead to serious security incidents.

**Ecologic footprint of the data center**

The ecologic footprint of the data center consists not only of consumed energy and emitted carbon, but also of the raw materials and embedded energy to produce the hardware and tapes. These products will unavoidably end up at the landfill. Due to its relatively long lifetime, the ecologic footprint of storage is mainly consumed energy and carbon emissions.

Storage systems are among the most conspicuous consumers of power and cooling in the data center[9]. To be usable, data storage drives must be fully powered and running.

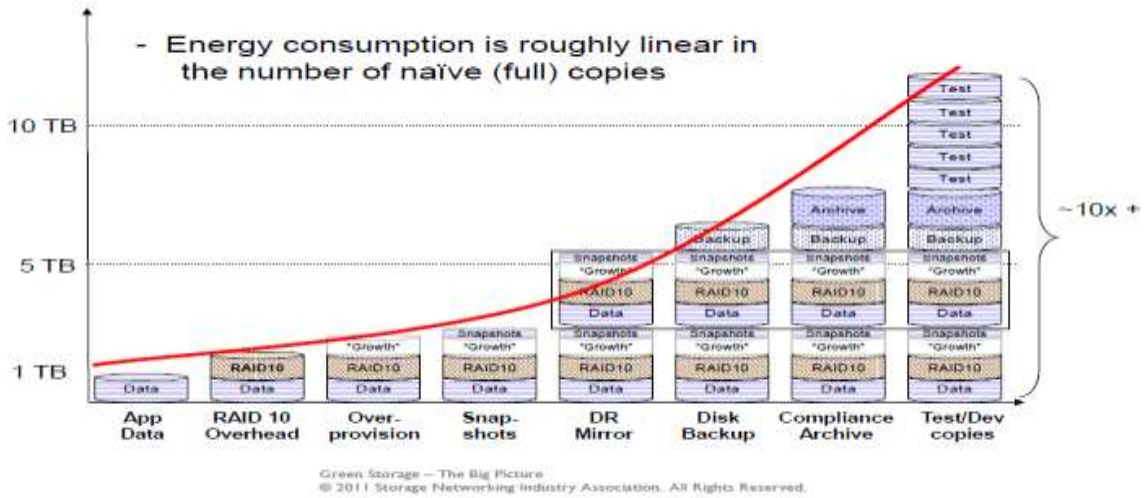


**Figure 5: Energy consumption is roughly linear in the number of naive (full) copies**

According to a recent study by the Storage Industry Network Association (SNIA)[4], energy consumption is almost linear with the number of naïve (full) copies, as shown in Figure 5. No doubt these copies include digital e-waste as well. The consumed energy is not limited to the storage hardware, but also includes the supply infrastructure to keep the core IT infrastructure running, e.g. UPS and cooling equipment. Most data centers still do have a PUE of 2, meaning 1kW hardware requires 1kW for supply equipment as well.

The recent focus of energy reduction projects in data centers has been on servers and cooling. For example, consolidation and virtualization projects brought quick wins for the energy bill. As a consequence, the percentage of storage of the total energy consumption is increasing even without growth.

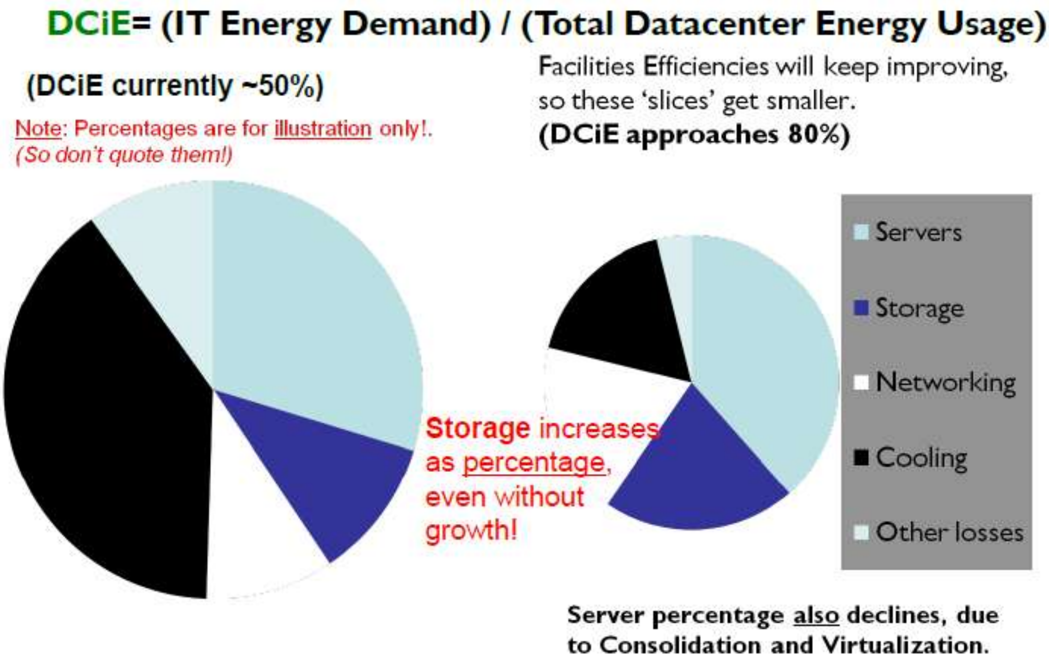


**Figure 6: Storage gets more visibility[4]**

It's expected that the share of storage in the total energy pie will increase as volume grows. To reduce energy consumption by the data center, storage has to be actively included in the scope of demand and capacity management. Consequently, storage will get more visibility by data center management.

## Who should care?

Data is a valuable asset for organizations and it should be managed accordingly. Traditionally, the IT department is held responsible for storage of the bytes and the business manages the information. Two different worlds that speak their own language and have their own mindset, organization, and dynamics.

Data governance is the intersection of Business and process and Technology. According to Orr[3]: "Data governance is the discipline of administering data and information assets across an organization through formal oversight of the people, processes, and lines of business that influence data and informational outcomes to drive business performance".

Digital e-waste, except unintentional wanted data, is created to become information for the user and eventually becomes just physical zeros and ones at the end of the lifecycle. Data ownership lies in the domain of the business.

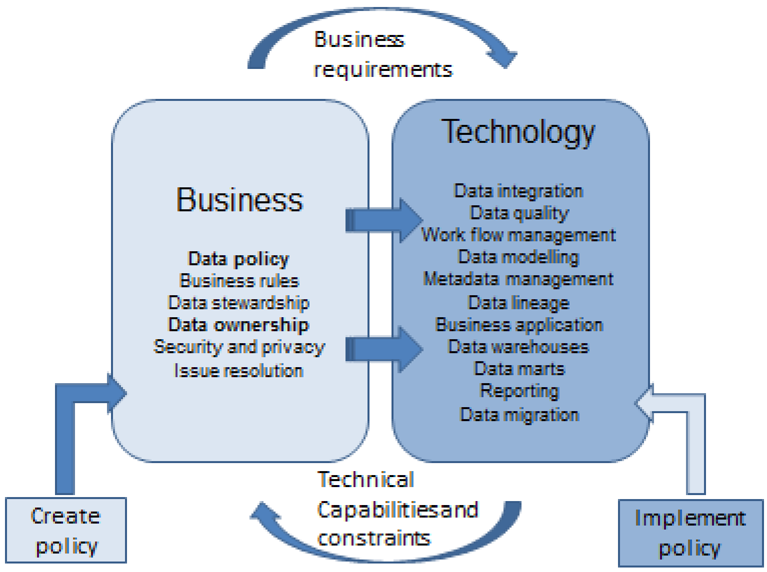


**Figure 7: Data governance; the intersection of Business process and technology.**

In most organizations however, the IT department is made accountable for the digital e-waste. All data is stored—at least for a period—to be compliant with law and regulations, and as long as the physical capacity is sufficient, there is no reason for change.

The value of data for an organization is determined by[11]:

- Probative value – organization has to be compliant to law and regulations
- Information value – data is required to execute and govern business processes
- Historic value – data used to keep track of the history of the organizations

We'll use the term business value to express all three types of values. Only the business is able to judge the business value of the data.

The business should make clear that they own the data during its entire life, from creation until disposal. The business should also be responsible for digital e-waste as well.

In the end, it is the business who is confronted with the growing issues related to digital e-waste, such as raising TCO of storage, performance, business continuity risks, security, and the ecological footprint of the data center.

## What can the business do?

As mentioned above, the business is the owner of the stored data within the organization. Digital e-waste is a subset of the entire data collection. Referring to the defined classification of e-waste, the business can be held responsible for three classes of digital e-waste:

- Used data
- Degraded data
- Unwanted data

We can apply the waste management hierarchy shown in Figure 8, which has been developed to reduce trash in the, let's say, physical world.

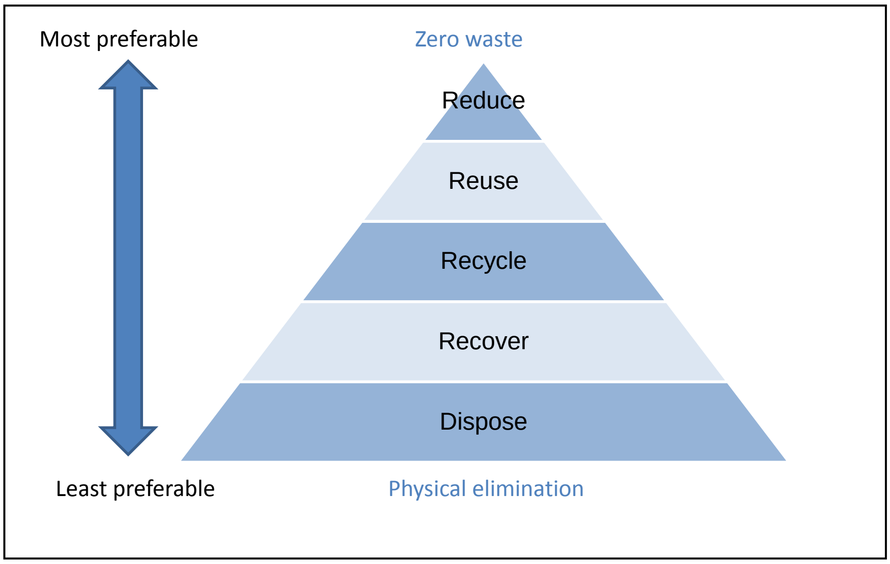


**Figure 8: Waste management hierarchy**

The most preferable choice is to reduce waste by design. At the drawing board, the product is designed with the end-of-life in mind and to minimize the volume of consumables during its entire life. Second best is to reuse the product when its functions still fit the requirements of others. Second-hand cars are a good example of products which lost their purpose for first-owners, but can be used by new owners with different needs.

The next option is to recycle the complete product to create another product with its own Purpose, Structure, State, and Performance (PSSP) characteristics. When full-recycling is not feasible, components might be recovered but ultimately, the product has to be disposed at the landfill or incinerate.

**Reduce**

The business can take several measures to reduce all class 2, 3 and 4 waste data. In reverse order, the interventions are presented, starting with unwanted data as the most effective approach.

**Class 4 - Unwanted data**

Unwanted data is waste data which never had any value for the user. Mailboxes contain a lot of these data, e.g. attachments which are never read. Common measures to reduce unwanted data are:

1. E-mail policy: Limit personal mailbox capacity as well as attachment volume.
2. Use of collaboration tools to centralize document storage and facilitate information pull strategy.
3. Active policy for storage of non-business data.
4. Rationalizing the application portfolio by deleting applications which are no longer or hardly used.
5. Data entry procedures to avoid data entry errors or worse, duplicate records.

**Class 3 – Degraded data**

Degraded data is data with a well-defined purpose, but whose performance is no longer acceptable due to defects in its structure or state. Examples of this category are data stored in formats not readable by new software versions.

Degraded data can be reduced by:

1. Defining data retention policy that complies with law and regulations and delete data which definitely lost business value.
2. Funding a project to convert the old data to a new format that is readable by the latest application.
3. Making budget available to parse the data for data-analytics purposes.

**Class 2 – Used data**

Used data is data that lost its purpose. Examples include monitoring data, most mail messages, and input parameters which have been processed. While used data lost its purpose for the primary user, it may have value to others.

Therefore, this data should be treated carefully and can only be deleted when assured that it has definitely lost their business value. Used data waste can be reduced by:

1. Defining data retention policy that complies with law and regulations and delete data which definitely lost business value.
2. Centralized storage.
3. Deduplication to avoid multiple copies of the same data.
4. Monitoring the access rate and access date of data.

### Reuse

Reuse is using the same object (structure and state) for different purposes and possible change in performance. Enterprise data warehouses and recently, analytic sandboxes for Big Data analytics, are good examples of reuse of data to create value. Transaction data which has lost its operational value are beneficial to perform analytic functions.

### Recycle

Recycling is changing the structure of an object to create a new purpose. A bottle of glass is recycled to make a new object of glass. Recycling of waste data is typical an IT activity and also outside the visibility of the business. An example of recycling is the transformation of old application components to a new application.

### Recover

Recover is to modify the state of an object. Though the structure is not changed, the purpose and performance may change.

Recovering waste data is to parse obsolete data to enable data mining. Data which has lost its initial purpose can be recovered for analytic purposes.

### Dispose

Dispose means a transformation of Purpose, State, Structure, and Performance. This activity is the domain of the IT department. However, the business, as data owner, should provide the security policy to ensure the data is safely disposed.

Information security processes should be in place to guarantee that all data is erased before the equipment leaves the data center.

## How can the IT department support the business?

The IT department, as storage service provider, can support the business to reduce the volume of digital e-waste by:

- Demand management
- Financial management
- Capacity management

These key service management processes add value to the business.

**Capacity management - reporting**

The first step in any improvement is to measure the actual state. Without measuring, you don't have control and are not able to take appropriate actions. Examples of reports that can be created to make the customer aware of the presence of digital e-waste are:

- Volume trend report, split by file type
- Reports on latest access date
- Dormant users report

Just as energy companies provide tools to measure energy consumption to their customers, storage providers should present information about the usage of storage.

**Financial management**

Money is an important driver to change consumer behavior. However, in most organizations, IT services are not charged directly or not split up per service; hence, the low awareness of the impact on cost. Cloud storage providers submit invoices on a pay per use basis, so IT departments may charge the costs of storage to their internal customers.

As soon as customers realize the costs of digital e-waste, they are triggered to reduce its amount.

**Demand management**

Demand management is about understanding the demand for services. Last but not least, the IT department can demonstrate partnership with the business by advising on costs, risks, and ecological impact of storage, as explained earlier.

## Conclusion

As the digital universe continues to expand, digital e-waste—the bytes no longer wanted—will grow in parallel at the very least. Waste management theory is applied to handle this problem. The object classification Purpose, Structure, State, and Performance can be used to define several classes of digital e-waste.

The business should be made aware that as data owner, they are accountable for and must confront the recurring issues of digital e-waste: rising storage costs, degraded performance, security, and the ecological footprint of the business.

Though treatment of digital e-waste is the domain of the IT department in most organizations, a change of mindset is required to effectively manage digital e-waste. It is the business that needs to take the initiatives to mitigate the risks.

This article recommends applying general waste management principles to tackle the problem of digital e-waste, the most preferable actions being prevention, reduction, and recovery.

As storage provider and 'waste company', the IT department is able to support the business by charging, reporting, and consulting on the state of digital e-waste.

## Appendix


1. Life and Death of Unwanted Bits: Towards Proactive Waste Data Management in Digital Ecosystems, Hasan R, Burns R, Department of Computer Science, John Hopkins University, Baltimore, United States, 2011
2. Phenomenon driven process design methodology: Formal representation Phohjola VJ & Tanskanen J (1998). Proc. CHISA'98 Conference, Aug 23-28. 1998, Prague, Czech Republic. CD ROM of full texts. Czech Society of Chemical Engineering. Software & design Magicware Ltd.
3. Data Governance for the Executive James C. Orr, , 2011
4. Green storage, the big picture, Storage Networking Industry Association, 2011 by Worth SW,
5. In Naples, A clash over Trash, 2010, Timeworld by Faris, Stephan,
6. Annual study digital universe IDC, 2012, sponsored by EMC
7. Report growth of information, 2007, IDC,
8. Information Life Cycle Management, Eine Methode zur Wertzuweisung von Dateien, 2009, Darmstadt, Germany by Turczyk, L.A..
9. Evolving the theory of waste management defining key concepts, University of Oulo Finland, 2004 by Pongrasz E, Phillips P.S. & Keiski , 4.
10. Green IT for Dummies, 2009, Wiley Publishing, Inc, Indianopolis, Indiana. By Baroudi C, Hill J, Reinhold A, Senxian
11. Selectiecriteria voor bewaartermijnen, 4e druk, Publicatie 2008-1,® 2008 NVBA