

3-31-2009

Auditing Personal Information Management

Mark Gregory

Finance and Operations, ESC Rennes School of Business, Mark.gregory@esc-rennes.fr

Mario Norbis

Department of Management, Quinnipiac University, Mario.norbis@quinnipiac.edu

Follow this and additional works at: <http://aisel.aisnet.org/ukais2009>

Recommended Citation

Gregory, Mark and Norbis, Mario, "Auditing Personal Information Management" (2009). *UK Academy for Information Systems Conference Proceedings 2009*. 29.

<http://aisel.aisnet.org/ukais2009/29>

This material is brought to you by the UK Academy for Information Systems at AIS Electronic Library (AISeL). It has been accepted for inclusion in UK Academy for Information Systems Conference Proceedings 2009 by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

AUDITING PERSONAL INFORMATION MANAGEMENT

Mark GREGORY

*Finance and Operations. ESC Rennes School of Business
35065 Rennes Cedex. France
Mark.gregory@esc-rennes.fr*

Dr. Mario NORBIS

*Department of Management. Quinnipiac University
Hamden, Connecticut 06518, USA
Mario.norbis@quinnipiac.edu*

Abstract

This paper firstly introduces personal information management using computers. It then describes a suggested approach to auditing the information needs which are personal to a knowledge worker. As part of an overall research programme in personal information management, the paper describes a personal auditing approach which is positioned as part of the experimental research aspect of the overall research. The paper aims to give practical guidance including on the effective classification of an individual's data.

Keywords: PIM, Semantic Web, GTD – Getting Things Done, KFTF – Keeping Found Things Found

The problem domain: Personal Information Management

Knowledge and information workers (Drucker, Peter (1999)) work as individuals within virtual team structures.

As individuals and as team members, they acquire information, which they store in a large number of arbitrarily complex ways: some being paper-based, but increasingly computer-based. Computer-based tools can assist in the storage and management of such information. However, little is understood about how people use these tools, how they learn new ones, the ways in which the tools constrain the ways in which people work and think, and how best to educate people to make the right choice of the right tools.

The assumption underlying this research is that individuals working in groups should be encouraged and educated to make better use of the available tools, and that the tools themselves should evolve (or even be replaced) by better ways of representing information and knowledge (for example, expertise).

Structure of this paper

We identify as background, what was already known about *personal information management* (PIM). We discuss *methodological issues*: how can we research PIM? This paper takes as an example of a multiple-methodology approach, an experimental approach based on self-audit. It starts with a question we ask ourselves, and one we hope that you the reader might also find interesting:

How might you improve your own PIM?

We draw interim conclusions and identify the scope of future work.

Working better

In general, we as example knowledge workers are motivated by or paid for what we do, what we achieve: for our *work*.

Doing things involves *processes, resources, information* and *knowledge*. Work is viewed here as usually being done in a competitive context – where we as individuals or as part of an organisation have to do and be better than others.

The role of information management in work

Knowledge and information workers work as individuals within *organisational departments* or the virtual team structures sometimes called *projects*. *Work processes* require *information* which the worker stores in a large number of arbitrarily complex ways. Some are paper-based. But they're increasingly computer-based, and they're moving to the Web.

Research goals and questions

This paper corresponds roughly to the first half of a three-year research exercise, having as goals:

- Initial goal
 - Identify what information storage and retrieval approaches exist to assist individuals and small groups to store, classify, search and retrieve personal information. (*We regard this aspect as being substantially complete. See Gregory, Mark & Mario Norbis (2008a), (2008b).*)
- Second goal
 - Observe, experiment with and evaluate information organisations that allow more convenient storage and retrieval than current mainstream approaches so as to be able to make well-founded recommendations. This is our *current* goal. One major aspect is treated in this paper.
- Longer term goal
 - We do have as a longer-term aim the need to identify, or even assist in establishing, a new approach to small-group knowledge management.

Some “obvious” research questions

These have set the agenda for the early research.

- What is Personal Information Management, PIM?
- Why and to what extent is it important?
- Who currently uses PIM?
- Why do so many people not use PIM?
- In what ways can people improve their PIM? How useful is this in terms of productivity or work quality?

Literature review: What is already known about PIM?

We have reviewed the literature at some length elsewhere. See Gregory, Mark & Mario Norbis (2008a), (2008b).

In summary, we have found that:

Over the last two decades a significant number of computer-based tools (sometimes referred to as Personal Information Managers or PIMs) Teevan (2006) have been created in order to assist in the storage and management of such information. So far it is arguable that no tool has achieved ubiquity, whether measured in terms of the extent of its use or the generality of its application. There is little consensus about the most appropriate methodology for evaluating information technology or information systems Beynon-Davis (2004) and so far Information Systems evaluation remains underdeveloped and undermanaged - Love, P. & Ghoneim, A. (2004)

The underlying hypothesis of the research-in-progress presented in this paper is that individuals working in groups should be encouraged and educated to make better use of the available tools, and that the tools themselves should evolve into or be replaced

by better ways of representing information and knowledge. This study is further motivated by perceived deficiencies in current data management paradigms and other paradoxes, some of which are highlighted here.

Deficiencies of the current data management paradigms. Current database concepts materialized in software packages, fail to handle properly the complex, non linear relationships among data items or families of them. Examples of this situation arise in diverse areas as new data organization to be utilized at museums information centres (De Vorse (2006)).

The productivity paradox. Paul Strassman (Strassman (1999)) has suggested a productivity paradox (which he identified at the corporate and the country level): that there is little or no correlation between the amount which organizations spend on IT and their profitability.

Uncorrelated investment. The authors' experiences suggest that there may be relatively little correlation between the investment made in personal productivity tools and those oriented to small-group productivity.

The usage question and the expressiveness of information. Our overall research aims to examine the reasons why what should be widely-used techniques, are not. It is our belief that computer users voluntarily sacrifice freedom in favour of structure in order to facilitate storage, retrieval, and especially querying and communication; but they still do not achieve the level of precision and communication that they strive for.

Towards a better Group Information Management paradigm. In this context, a new paradigm should provide the user with the flexibility to accommodate a variety of interactions among data items maintained by individuals or groups: by defining structures and mechanisms to better accommodate particular needs while keeping a more rigid framework to guarantee consistency while allowing for expressiveness of the data to be present.

PIM Scope

Knowledge and information workers working as individuals within virtual team structures acquire information, which they store in a large number of arbitrarily complex ways: some being paper-based, but increasingly computer-based.

There are a number of computer-based tools, sometimes referred to as Personal Information Managers or PIMs (Kelly (2006) and Teevan et. al. (2006)) which can assist in the storage and management of such information.

PIMs are additional and complementary to the functionality of so-called “office suites” (sometimes “office productivity suites”). An office suite is a software suite (collection of component programs) intended for use by typical clerical workers and knowledge workers. The components are generally distributed together, have a consistent user interface and usually can interact with each other. The best-known current examples of office suites are Microsoft Office and OpenOffice.org.

The facilities offered by office productivity suites may include

- Word processor
- Spreadsheet
- Presentation program
- Database
- Graphics suite
- Communication
- Email client

They may also include

- Personal information manager
- Groupware

Office productivity suites focus on the production of documents of various kinds and offer various tools for managing personal information. But these tools can become a part of the problem, leading to “information fragmentation”. Different devices and applications often come with their separate ways of storing and organizing information.

Over the last two decades a significant number of computer-based tools (sometimes referred to as Personal Information Managers or PIMs) have been created in order to

assist in the storage and management of such information, see Teevan, Jaime & William Jones & Benjamin B. Bederson (2006) for a much fuller discussion. Examples include email, contact and event management software (e.g. Microsoft Outlook); hierarchical outliners; mind-mapping software (e.g. MindManager) and the use of general office applications such as spreadsheets and relational databases (e.g. Microsoft Access) applied specifically to personal and small-group information management.

So far it is arguable that no tool has achieved ubiquity, whether measured in terms of the extent of its use or the generality of its application.

The underlying hypothesis of the research-in-progress presented in this paper is that individuals working in groups should be encouraged and educated to make better use of the available tools, and that the tools themselves should evolve into (or be replaced by) better ways of representing information and knowledge.

Writing about Time Management, Project Management and Personal Information Management

A very great deal has been written about how people plan and manage their own time (so-called “time management”) and that of others (people management and project management).

On (personal) time management, see for example the work of Allen, David (2001).

Rather less has been written on personal information management. See for example the work of the University of Washington / Microsoft research group (for example, Jones, William (2007)). See also Boardman, Richard (2004).

Some related areas that are close to but outside the scope of this paper

We note that so-called “situational applications” exist (see for example Cherbakov, L. & A. Bravery & B. D. Goodman & A. Pandya & J. Baggett (2007)). Indeed, we can identify what we call an “applications space” dimensioned by the scope of application:

Scope	Systems approaches examples
Personal	PIM
	Office productivity tools
Small-group	GIM, situational applications
	Office productivity tools
Departmental, functional or SME	Office productivity tools
Corporate: enterprise	ERP
Intra-enterprise	Extended ERP

The evaluation of situational applications (to be discussed further in Gregory, Mark & Mario Norbis (2009b)) identifies issues which apply also to the evaluation of PIM approaches. However, we choose to ignore them in this particular paper.

Similarly, it is almost traditional to consider HCI issues when discussing information systems. Again, our chosen concentration is on *usefulness* rather than usability.

Findings to date: The Challenge

Following Strassman, Paul (1999), we can identify a productivity paradox: increasing technological possibilities raise the hurdles. As a result, although you can do some things quicker, overall you don't get much more work done!

Why? You have to do things better in order to compete with others who are reacting to and benefiting from the same new possibilities:

- You can no longer persuade, educate or sell using hand-written acetates...
- So you buy a PC and an expensive smartphone
- Which you then have to learn to use and exploit
- You and others do things *better*, but not necessarily *faster*!

Therefore: we should audit the way we work and the way we manage our information in order continuously to improve them.

Personal information management and processing

Personal information management: a brief introduction

Many of us keep a wide range of personal data, which we classify or sub-divide into areas such as:

- Agenda: list of appointments
- Address book: our contacts
- To Do list

Different groups of knowledge workers (Drucker (1999)) keep different kinds of personal information. Thus students and researchers keep more specialised (but still widely-used) data such as

- Bibliography: reference list
- Reading notes
- Project logbook

Some of us do this primarily on paper, in spiral notebooks or perhaps in more-specialised diaries and the like.

Computer-based Personal Information Management

Many of us also or alternatively use personal computers (desktop or notebook), digital PDA (Personal Digital Assistant) devices or smartphones. Some of us work in contexts where this kind of information is no longer exclusively ours, and we choose (or are obliged!) to share and merge this kind of information.

All of us are of course very careful to copy this personal data from one device to another, in order to safeguard it from corruption or loss. Some of us take additional care to synchronise this data; that is, when we store a new contact detail on our smartphone, we subsequently synchronise it into our desktop environment. An obligation to share this kind of data occurs if we have a secretary or administrative assistant who also collects this kind of data on our behalves.

Individuals, teams and organisations need to carry out business processes; they have to Get Things Done: GTD

See Allen, David (2001).

To do this, they also need to Keep Found Things Found: KFTF

See Jones, William (2007).

What is the difference between Getting Things Done (GTD) and Keeping Found Things Found (KFTF)?

- *GTD* is about *planning your work and doing it*, as an individual and in the various teams of which you are part. For example, for students:
 - Teams for coursework assignments
 - Work or project teams when doing internships

- Student clubs and micro-enterprises
- Football teams
- Whatever you do in small groups
- It includes things like diaries (agendas) and project plans
 - Those agendas may be personal to you, or include group agendas
- **KFTF** is about *keeping all the information you need* to learn, to work, to live. Examples are very numerous. For example, for students, they include:
 - Lecture notes
 - Reading lists
 - Contact lists (address book)
 - Shopping lists
 - Recipes for meals
- There is some overlap between these kinds of information, for example, a list of sporting fixtures (e.g. when your team is playing football) can be in either the GTD category or the KFTF category or both.

Some dimensions of personal information management

- Capture: of text, pictures, audio, video, web clippings...
- Access
 - Handheld, on PC, Web (“cloud”)
- Finding things again
 - Classifying them when found
 - Searching for them later
- Presentation
 - Visual aspects
 - Communication and Sharing (read-only, or shared-update)
- Secure storage
 - Keeping secrets
 - Preserving investment
 - Potentially across decades

Software for personal and small-group information management

There are a number of computer-based tools, sometimes referred to as *Personal Information Managers* or *PIMs* which can assist in the storage and management of such information. PIMs are *additional and complementary* to the functionality of the so-called “*office suites*” (sometimes “office productivity suites”).

Office productivity suites

An office suite is a software suite (collection of component programs) intended for use by typical clerical workers and knowledge workers. The components are generally distributed together, have a consistent user interface and usually can interact with each other. Best-known current examples of office suites are Microsoft Office and OpenOffice.org.

What do they do? Office suite functionality:

- Focus on the *production of documents* of various kinds
- Also offer various *tools for managing personal information*
- The facilities offered by office productivity suites may include
 - Word processor
 - Spreadsheet
 - Presentation program
 - Database
 - Graphics suite
 - Messaging
 - Email client
- Many people use general office applications such as spreadsheets (Microsoft Excel) and relational databases (e.g. Microsoft Access) specifically for personal and small-group information management

Using office software to improve your personal information management

AIM: To learn about, train in and exploit specific programs, often within a specific suite.

- Advantages
 - Availability
 - Consistency
 - Integration between programs
- Disadvantages
 - You begin to think about information only in the way provided by your chosen tool
 - You are potentially putting your mind into a straitjacket!

Some issues in personal information management

Using multiple tools can become a part of the problem, leading to “information fragmentation”

This is exacerbated by the multiplication of platforms – PCs, smartphones, cloud computing. Different devices and applications often come with their separate ways of storing and organizing information.

Problems associated with approaches based on particular programs in a suite

We can use an analogy:

- A carpenter who only uses hammers and nails tackles badly, or not at all, problems which need screws and screwdrivers!

Similarly:

- Someone who uses spreadsheets to do what should be done with a database
- Having a good toolbox doesn't make you a good carpenter...

Ready-made or Do-It-Yourself?

If you need a new kitchen, you may Get A Man In (GAMI)! The resulting kitchen will be totally customised, expensive, depend on a partnership between client and supplier and on accurate transmission of requirements. After which you the client change your mind. Alternatively, you might choose a Do It Yourself (DIY) approach using the products of Brico Dépôt, Mr. Bricolage, or B&Q. Another option is to buy a kit from Ikea. In a similar way, the organisation that needs a new information system usually has to choose: Bespoke (custom-built)? Packaged? Integrated-component system building?

Ways of managing your personal information

Applied to personal information, the alternatives appear to be to

- Write your own PIM program?

This is not usually a sensible option! However, the arrival of “Situational Applications” (to be discussed in Gregory, Mark & Mario Norbis (2009b)) may make this more common, especially in small groups or small enterprises.

- Build your own customised approach

By integrating parts of an office suite, e.g. using Excel or Access.

- Select and procure one or more ready-made PIM program(s)

In practice: some *combination* is common.

PIM programs

Computer-based tools (sometimes referred to as Personal Information Managers or PIMs) have been created in order to assist in the storage and management of personal information.

Example approaches include

- Email, contact and event management software (e.g. Microsoft Outlook)
- Hierarchical outliners (e.g. Microsoft Word outlines)
- Mind-mapping software (e.g. MindManager)
- Programs that call themselves PIMs (e.g. EssentialPIM)

Bought-in solution: “PIM” (Personal Information Manager) tools or packages

Various so-called “PIM” (Personal Information Manager) tools have been developed and marketed with varying degrees of success. We have a database of over 150 such programs and services. These tools are frequently based on an underlying relational database, whose existence may be visible to the user or hidden from her.

Bought-in solution: “GIM” (Group Information Manager)

Various “GIM” (Group Information Manager) tools have been developed and marketed with varying degrees of success. The most-established such tool is IBM’s Lotus Domino (2008) family of applications. This incorporates Lotus Notes, which has been widely used to provide email client and document storage and retrieval facilities which arguably constitute the basis for group information management. More recently, Microsoft has introduced a raft of related tools which address the same basic market need. Based on Microsoft Exchange Server (Microsoft Exchange Server (2007)) and Microsoft SharePoint, these tools (just as those proposed by IBM) share as characteristics an emphasis on structuring data and information so as to encourage its sharing and reuse; but with dependence on computing professionals to set up and maintain the shared document store and/or database.

PIM Functionality: What PIMs do

This section firstly discusses in a summary manner the meaning of data, before proceeding to list PIMs identified by the authors and beginning to identify and classify their associated *functionality*, that is, what users can do with them.

The meaning of data: semantics

Making lists and storing them is not rocket science. In fact, it isn’t even science. A list is only as useful as the meaning it conveys. Consider this list of (what most of us will read as) girls’ names:

Andrea	2007
Chantal	2007
Gabrielle	2007

What is this? Three members of a hockey team?

The addition of a column heading changes the story a little:

Hurricane name	Year used
Andrea	2007
Chantal	2007
Gabrielle	2007

What we have done is to classify the data, by naming the sets. The process of labelling or naming data gives so-called semantic significance to the data. To be meaningful, data needs *syntax* (rules for content and formatting) and *semantics* (rules for meaning). An alternative and equivalent formulation is that data needs metadata to give it significance. Classification is fundamental to science and to knowledge.

Structure and meaning

To make use of any computer based personal information management tools, we have to “structure” our data

Computer users voluntarily sacrifice freedom in favour of structure in order to facilitate storage, retrieval, and especially more precise querying (answering ad hoc questions about the data) and communication; but they still do not achieve the level of communication that they strive for. In order to use computers we have traditionally needed to sacrifice, to limit the expressiveness, of the information stored, where expressiveness is defined as the ability to communicate meaning. Well-structured data can be queried with greater precision; that is, more accurate and complete answers can be obtained to questions about the data.

To illustrate this point. If we extend the example above:

<i>Hurricane name</i>	<i>Year used</i>	<i>Meteorologist</i>
Andrea	2007	John Smith
Chantal	2007	Methuselah Gabrielle
Gabrielle	2007	Chantal Legros

With data structured in this way, we can achieve precise answers to different queries:

- Which hurricanes have been named “Chantal”?
Answer: one - Chantal
- Which hurricanes have been named by a meteorologist called “Chantal”?
Answer: one - Gabrielle
- Which hurricanes have been named after the meteorologist?
Answer: none

Note that free-text searching of the content alone, without taking into account the structure of the data, would give imprecise (inaccurate) answers. The words Chantal and Gabrielle occur more than once in different contexts and thus with different meaning.

But how do we express meaning? Individual knowledge workers are provided only with basic tools in which integration remains unintuitive. Indeed, each tool tends to highlight one or two information storage and presentation techniques to the exclusion of others. The worker then resorts to approaches such as managing tasks by leaving emails in the inbox, and keeps lists in linked spreadsheets. This creates isolated islands of under-managed and difficult-to-integrate data.

Structure imposed centrally is essential in some contexts and inimical in others

We sacrifice freedom in favour of structure in order to facilitate storage, retrieval, and especially querying (answering ad hoc questions about the data) and some aspects of communication; but we still do not necessarily achieve the level or effectiveness of communication that we strive for.

Some data is very clearly the property of a worker's employing enterprise, and some needs to a greater or lesser degree to be held and managed centrally. Standards vary widely according to the objectives and style of the organisation. A worker in a client call centre may not be permitted to store any data locally on a company owned computer. Conversely, universities may actively encourage information sharing. More common perhaps is controlled shared information – as in medical practice or business consulting. Many organisations seek to impose a standard way of capturing and storing data, which meets some purposes but defeats others.

Analysis and further research questions

We are confident that there exist both problem space(s) and solution space(s).

Methodology

Our overall approach is in the spirit of Denzin (1970)'s "multiple triangulation". Specifically we intend asking lots of "little" questions so as to feel our way towards a better understanding – all this with an emphasis on practical usefulness.

Methodological approach

Our methodological approach involves:

Reviewing the state of the art: largely secondary

Learning from small user populations

We intend identifying specific populations such as:

Obsessive serial fans – the people who love PIMs: Easily done via forums.

Influencers (reviewers, journalists, teachers).

People who resist PIMs (neo-luddites): Finding out who is resistant and why is more difficult, but a large proportion of student users – even after they have tried a PIM – still declare themselves uninterested in the approach. This research is an ethnographic approach, in which we need to take care in distinguishing self-justifying, "false" ethnography from "true": Weir, D. & Hutchings, K. (2005).

Learning from large samples

This detailed research has to be complemented by experimentation and a quantitative approach. This latter is based on exposing comparatively large populations to questions and suggestions about the nature of personal information management and the use of a personal audit approach which can be adopted by different kinds of knowledge workers. An initial experiment in this area has been conducted based on a group of bachelors-level students and this paper reports the initial findings of that survey. Further work detailed research is necessary, is introduced by this paper, and has started. 455 Students Bac+4 (Bologna M1) students were required to audit their personal information management as part of a group assignment. We are still carrying out Quantitative and Qualitative (content analysis) of the data obtained.

Auto-ethnography

This is another interesting angle in an overall triangulation. The authors themselves and close colleagues are carrying out a longitudinal study of how their personal information management evolves over time.

Experimental / action-research basis

Almost inevitably, we have become drawn to a more interventionist approach based on the premise that it is necessary and desirable to educate people towards effective PIM exploitation; to incite them to improve. The approach presented in this paper involves encouraging people to audit themselves as the basis for experimental use. Who is the target: well, why not you and me?

Getting smarter: Auditing your current approach to personal information management

Why not set yourself the task of improving your personal information management? You can ask for help from a friendly neighbourhood PIM specialist (if one exists); or you can help yourself by carrying out a Self-Audit, perhaps following some suggested first steps...

Questions to ask yourself about personal information management

- What computers and mobile phones do you currently use?
- List the computer programs you use for personal and work-related purposes
- List the web services you use for personal and work-related purposes
 - Have you ever considered alternatives? Which?
- How do you keep a list of favourite web sites? Do you use bookmarks (signets) in Internet Explorer, Firefox, Safari etc.? Do you keep an online list of favourites (e.g. GoogleToolbar bookmarks)?
- List the ways in which you store and manage personal information at the moment
 - How do you keep your agenda?
 - Which email clients do you use? Do you synchronise them?
 - Where do you keep your electronic diary (if any)?
- What personal information matters to you? Make a list of the various types of information you store, and how you currently do it
- List the processes you carry out to maintain and use this personal information

Kinds of data: What data do you keep?

- Some suggestions:
 - General information
 - Used by “all” knowledge workers
 - Semi-structured organization of small pieces of information (phone numbers, errands to run, books to read...)

- Contact management, address books, etc.
- Diary: Calendar and meeting scheduling
- To Dos: task management for self and others
- Journal: a record of the use of your time
- Document management
- Message management
- Resource management

Rôle-specific information

- Just as there is an overlap between personal information and shared small-group
- So also there's an overlap between
 - Generic information – the information that almost anyone might keep
 - Agenda, contacts, etc.
 - Rôle-specific information
 - For a student: references / bibliographic details, shared agenda, etc.
- Why? Different groups of knowledge workers (Drucker (1999)) keep different kinds of personal information

Processes associated with personal information: Basic processes

Among the processes associated with personal information are these; which matter to you?

- Capture
- Store
- Secure
- Communicate between devices and platforms
- Classify
- Find
- Show
- Present
- Share
- Reuse
- Publish

Processes associated with personal information: Less common processes

At a more detailed level, do you need:

- Diarising: record “everything” (See Gemmell, Jim & Gordon Bell & Roger Lueder (2006)).

This includes also:

- Personal notes/journal, annotations and note-taking in multiple media: a kind of electronic jotter
- Transcription between media, e.g. handwriting recognition, voice recognition
- Search across email, e-docs and other information forms; across multiple media types
- Hypertext authoring: writing documents that make links between each other
- Synchronisation between computers: Mobile/PDA devices and inter-device synchronisation
- Coordination between people in hierarchies and in projects
- Visualisation of information resources
- Graphing, charting, mind maps etc.

Understand your style of work: How do you work best?

- How organised do you like to be?
 - Do you thrive on organisation, or find that it stifles your creativity? People oriented towards structure may favour databases, or PIMs which offer powerful data structuring. People oriented towards spontaneity and personal creativity may prefer more visual approaches – or to stick with paper!
- What kinds of computer software do you like, feel at ease with, or want to master?
 - Are you at ease with classification and with rigour? Try database.
 - Do you enjoy a numeric, quantitative, algebraic approach? Try spreadsheet.
 - Do you think visually? Try mind mapping.
 - Are you brave enough to try novel approaches? Try a specialist PIM program.

Decide about personal information management that suits you

Think about, describe, analyse and even write notes on:

- How you Get Things Done (GTD)
 - That is, how you keep details of what you need to do and when, and how you plan and organise your time
- How you Keep Found Things Found (KFTF)
 - That is, what you do to store and manage all the personal information you need in order to learn well and to live well
 - How should you classify things in order to be able to find them again? On paper? On your computer and other devices (phone, music)?

Think about group tasks

- Describe and analyse how you (plural) Get Things Done, that is, how you keep details of what your various workgroups need to do and when, and how you plan and organise your time as a group
 - How might you use computer software (such as the programs you used individually, or others) in order to improve this management of time? Can you get group “buy-in” (commitment) to your suggested approach?
- Describe and analyse how you Keep Found Things Found, that is, what you do to store and manage all the information you need to work together effectively in groups
 - How might you use computer software (such as the programs you used individually, or others) in order to improve this management of information?

Choose software to improve your personal information management

- One possibility is to use a ready made PIM
 - We have already identified over 150 PIM/GIM (group information manager) programs
- What if that is too restrictive, or you can't afford it?
 - The other possibility is to "roll your own" personal information management using office tools
 - Word processing
 - Spreadsheet
 - Database
 - Etc.

Some important computer programs – examples only! find your own personal preference!

- Microsoft Outlook
- Microsoft Office 2007
 - See in particular, Word, Excel, Access, Outlook, OneNote, Visio, SharePoint, InfoPath, Groove
- Web Services
 - Examples include:
 - Backpack
 - Remember The Milk
 - Google Calendar
 - Plaxo
 - Digg, Reddit, del.ici.ous
 - Twine – semantic web
- Mind mapping
 - Examples include:
 - Visimap
 - MindManager
- Group Information Managers
 - Examples include:
 - Lotus Notes
- Specialist PIM applications
 - Examples include:
 - Chandler
 - Info Select
 - Tinderbox (Mac only)
 - Ecco and EccoExt
 - SQLNotes (InfoQube)
- Semantic desktop – research prototypes
 - Such as:
 - Gnowsis
 - Haystack

Acting to improve your personal information management

- Choose at least ONE software program which you will use over the next few weeks in order to help you IMPROVE the ways in which you Get Things Done and Keep Found Things Found
 - Start to use it NOW
 - Set a time limit on your experiment – say five or six weeks
 - Take a LOG of what you do with the program, how you plan its use, how you learn more about it, what your experiences are – good and bad
 - A log is a list of things you have actually done – what they were and when you did them.
 - At the end, complete your log with an EVALUATION of how effective your experiment was. What will you do in the future to improve?

Making it easier to find things - background

- We all classify things – that is, we group them by name or keyword
- A simple organization of kinds of things is to list them alphabetically. If we give a list a title which attempts to name or describe the items in the list, we begin to establish a

vocabulary. If we make a list of football teams, each member of the list “*is-a*” football team.

- Making items into lists and deciding which list each member is a part of is a process called *taxonomic classification* or just *classification*, and it is fundamental to science and to the communication of meaning: we are ascribing and defining a vocabulary, and grouping things by their classification or type
 - See Boardman, Richard (2004) for a fuller discussion, including the limitation of hierarchical (tree) structures

Classifying the information stored on computers

All computer users classify their files in accordance with various criteria, such as the type of activity they are used for (for example, school or university files, home files, car files); but sometimes also the kind of program which is used to manage them. Each file has a name, and it is stored in a particular folder or directory. The computer file explorer program shows the names of folders and within each folder the names of the files which it contains. In addition, various properties or attributes of the files are displayed by the explorer. These attributes including the type of file, the date on which the file was first created, the date of its most recent modification, and other more specialised properties.

Grouping together computer files in folders is a simple example of classification and categorisation. Some computer users take the contents of folders and group them together in sub-folders. Establishing a classification system which corresponds well with the way in which the computer user works is not always obvious. Furthermore, the classification system which is appropriate at one point may subsequently become inappropriate or obsolete. As a consequence, it is necessary to use the file explorer to rename folders, and to move files from one folder to another.

The screenshot shows an example of a computer user’s folder hierarchy.

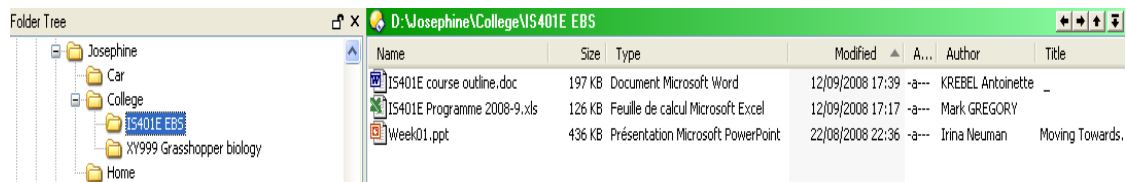


Figure 1: Example folder structure as maintained by a student

Making it easier to find things - application

- How do you group files into folders on your PC? Can you, should you, improve this? How have your classifications changed in the past, and what changes do you anticipate?

- Do you find the inherent limitations of the standard folder structure (which is strictly hierarchic) troublesome?
 - Do you need ways to store things in more than one place at a time?
 - Potentially you can take this further as you create an ontology.

Ontology

- An ontology defines a set of representational primitives which model a domain of knowledge
- **Ontologies extend taxonomy** by applying a larger variety of relation types than just “is-a”
 - The representational primitives are typically classes (or sets), attributes (or properties), and relationships (or relations among class members)
- An ontology is a data model that represents a set of concepts within a domain and the relationships between those concepts; it is used to reason about the objects within that domain
- Ontologies are used in artificial intelligence, the semantic web, software engineering and information architecture as a form of knowledge representation about the world or some part of it

Ontologies generally describe

- **Individuals:** the basic or "ground level" objects
- **Classes:** sets, collections, or types of objects
- **Attributes:** properties, features, characteristics, or parameters that objects can have and share
- **Relations:** ways that objects can be related to one another
- **Events:** the changing of attributes or relations
 - Source: OWL (2009)

Relevance to your self-audit: Why is it important to classify things?

As we decide what to do (GTD) and get the information we need to do our work and to live our lives (finding and keeping information), we need to store lists in a way in which we can find them easily, to name them, to classify their contents, and to relate them. Your personal information management may need to evolve in these directions. It is therefore wise to choose programs which permit you to create your own lists, to name them, to classify them either by a hierarchy of keywords or (better since more flexible) by more than one keyword, to search them and to enable you to link one item to another. Unfortunately, very few programs do all of these things!

Reminders

- You should aim to *Get Things Done*
- You should aim to *Keep Found Things Found*
- Take into full consideration
 - Personal work
 - Group work: work you do with others

- We hope that you'll enjoy this self-audit
 - Certainly aim to PROFIT from it!
- We hope you'll want to talk to us about your experiences, and thereby contribute to our research

Interim conclusions and future work

We have scoped the field of enquiry. We have identified the need to experiment with and evaluate information organisations that allow more convenient storage and retrieval than current mainstream approaches so as to be able to make well-founded recommendations. We note the importance of cognate questions, notably how business users can create their own systems (end user programming, situational applications).

Constraints

We have described a potentially very rich area of research in which at the moment there is not a great deal of published material. We recognise that there are large areas that deserve attention but that we will not have time to investigate. In particular, we note with approval the work of Penrose, Roger (1990) – which points to fundamental limitations on the usefulness of computer-based approaches in this and in other areas; and the large existing literature on human computer interface (HCI) issues, much of which has relevance to this enquiry but which we have deliberately excluded from consideration in this paper.

Initial Findings and Future Research: The initial experiment with students

We are still carrying out quantitative and qualitative evaluation of the mass of research data we acquired in this experiment. However, we can make some initial observations. Students aged 21-22 do not have a large volume of personal data to manage and most of them remain very unreflective in its management. Most of the students were native French speakers and favoured the use of software with a French language user interface. That considerably restricted their adventurousness in tool choice. The choice of using generic tools which they already knew (in part or in small part!) was widely made. The tool most commonly used was Excel, which with add-on programs is not an inappropriate choice. Many students did seek out a PIM package, most frequently EssentialPIM because it is available in French. The practical difficulties associated with computer-based personal information management mean

that most of them liked the ideas behind computer-mediated PIM but would not continue the experiment further. Very few had the foresight to choose tools which integrate PC and smartphone data; those that did reported higher levels of satisfaction. The value of the experiment as a whole is compromised by two major constraints. One is that the students only had six weeks from the assignment briefing to the hand-in. This is a short time in which to review information management, to choose, learn and fully appropriate a tool. The second constraint is that many students did not follow our advice to use their chosen software over a six-week period but did everything at the last possible minute.

Migrating the data to the web: cloud computing and Web-Based Applications

In recent months, practitioners have become very focussed on the potential of what is sometimes called “cloud computing”, that is, the developing probability that personal, small-group and corporate data and the programs used to manage them will “migrate” from local client and server computers to reside in part or in whole on server computers accessed via the web and the Internet. See for example Miller, Michael (2008). Miller believes that, as computer-mediated communication becomes faster, cheaper and more accessible (for example, in the home, across the wireless office, via the mobile web), users are increasingly freed from the necessity to own expensive programs stored on client computers with data stored only on that one computer. In Miller’s view, so-called “cloud computing” liberates users to collaborate with friends and family, share ideas with co-workers and friends, and to become more productive. The “cloud” consists of thousands of computers and servers, all linked and accessible via the Internet. With cloud computing, computing becomes web-based instead of being desktop-based. Users can access all their programs and documents from any computer that is connected to the Internet.

The significance of cloud computing and Web-Based Applications for personal information management

If information management increasingly resides in the cloud and simple web application builders become more usable, then the options for effective personal and small-group information management increase greatly. The ways in which a company can procure the business applications necessary for its operations include combinations of the following four basic approaches:

1. Bespoke (custom) development
2. Purchase and use of packaged applications
3. Systems integration – composing applications from different components found on the web; this requires technical skills
4. So-called “end user programming”, which can be done by business professionals.

Web-enabled, cloud-hosted and cloud-delivered end user development is likely to cause a considerable increase in the extent to which at least some knowledge workers can develop their own team-based information management.

Towards Personal Knowledge Management and knowledge creation

Personal information management is increasingly influenced by techniques and approaches emerging as the World Wide Web evolves.

Three generations of the World Wide Web

We can informally identify three generations in the web (the classification is informal because the language used is by no means standardised or agreed). The first to be widely adopted, following Tim O’Reilly’s suggestion (see Oreilly, Tim (2007)) was Web 2.0; by extension, Web 1.0 and Web 3.0 have been identified. The associated meanings vary, but we will characterise them as:

Generation	Meaning
Web 1.0	The web as interlinked or hyperlinked pages; a web of data or of information
Web 2.0	The web as society, in which individuals and groups identify and reveal themselves; and in which as they discover or reveal things and people, they wish to share that discovery with others: by means of blogs, wikis, folksonomies, social networking, shared bookmarks and the like. The means chosen to identify information and knowledge are often based on so-called “tagging”.
Web 3.0	The web as knowledge exchange and representation, enabled by <i>semantic web</i> techniques originally identified by Berners-Lee, Tim (1998).

We suggest the following characteristics of these three generations:

	Web 1.0	Web 2.0	Web 3.0
When	1990-2000	2000-2010	2010-2020
Headline	Data & information	People	Knowledge expressed as semantics
Data formats: language	HTML, Flash	XML, RSS	RDF, OWL
Delivery	HyperText Transmission Protocol	HyperText Transmission Protocol	Thin client; cloud computing
Data processing: applications or so-called “apps”	Monolithic; dominantly server-side	Modular; client or server; complex systems integration of client and server-based apps. Initial appearance of web-hosted applications and applications builders.	Loosely-coupled Lego-brick applications following Google and their CEO Eric Schmidt. Proliferation, general or even ubiquitous use of web-hosted applications and applications builders

Semantic Web: current state of the art

Large-scale research prototypes aimed at the corporate level are beginning to emerge. Their implementation and use is fraught with practical and conceptual difficulties. The best-known example to date is MIT’s Simile project (Simile (2008)). See also Gnowsis (2008).

Semantic web approaches applied to personal information

Two possibilities exist: either PIM products or services which incorporate semantic web techniques; or systems which apply semantic web techniques to pre-existing data stored on a specific computer. The latter approach is referred to as the semantic desktop.

Enhancing the usability and usefulness of the Web and its interconnected resources might be achieved by:

Servers which expose existing data systems using the RDF and SPARQL standards. Many converters to RDF exist from different applications. Relational databases are an important source. The semantic web server attaches to the existing system without affecting its operation.

Documents “marked up” with semantic information (an extension of the HTML <meta> tags used in today’s Web pages to supply information for Web search engines using web crawlers). This could be machine-understandable information about the human-understandable content of the document (such as the creator, title, description, etc., of the document) or it could be purely metadata representing a set of facts (such as resources and services elsewhere in the site). (Note that anything that can be identified with a Uniform Resource Identifier (URI) can be described, so the semantic web can reason about animals, people, places, ideas, etc.) Semantic mark-up is often generated automatically, rather than manually.

Common metadata vocabularies (ontologies) and maps between vocabularies that allow document creators to know how to mark up their documents so that agents can use the information in the supplied metadata (so that Author in the sense of ‘the Author of the page’ won’t be confused with Author in the sense of a book that is the subject of a book review).

Automated agents to perform tasks for users of the semantic web using this data. Web-based services (often with agents of their own) to supply information specifically to agents (for example, a Trust service that an agent could ask if some online store has a history of poor service or spamming).

A very important issue: whose ontology?

If we accept the necessity for imposing some sort of classification mechanism to achieve accuracy and precision in searching for information, the next question which inevitably arises is “whose ontology shall we adopt?” We can identify three broad and overlapping alternatives:

Standardisation by committee (or by professional body, or by employer): top-down imposition

This is frequently done within communities of experts, such as pharmacists or medical practitioners.

Emergent ontology - ontologies shared between workers in small, often virtual, groups: bottom-up conceptualisation

This situation is common in areas of fast-changing technology or practice. A common vocabulary and classification system “emerges” and almost imposes itself. Evolution, when it occurs, is ad hoc.

Specialist programs which recognise or implement user-defined ontology

E.g. Ideospace (2008).

“Web 2.0”, Social Networking and personal and group information management

Fichter (2004) presents “solutions to electronic problems in information management: tools to solve problems such as collecting, organizing, searching and sharing information online, simplify storing, keeping, searching and sharing Web resources”. Such applications tools are sometimes called social bookmark tools and have many features in common with the older generation of Web-based bookmark sites and personal information managers. Many want to accomplish tasks quickly and easily when a useful online resource is available. Such tasks include bookmarking the site with one click of the mouse and have the Internet domain name and page title automatically populated into the appropriate fields, ready to edit, jotting down a comment or description of the site, clipping out important excerpts, and filing it in categories that are created for suitability, among others. The new breed of social bookmarking applications offers more than just a universally accessible search and store facility for links. Such applications assume that once a site is found, not only do users want to share it with others but users want to discover other related sites and people who are interested in the same topics.

Semantic wikis

An emerging approach which arguably combines the power of social networking and formal knowledge representation is that of the so-called “semantic wiki”. Wikipedia – Semantic wiki (2008) suggested that a semantic wiki is “a wiki that has an underlying model of the knowledge described in its pages. Regular wikis have structured text and untyped hyperlinks (such as the links in this article). Semantic wikis allow the ability to capture or identify further information about the pages (metadata) and their relations... semantic wikis try to ... allow users to make their internal knowledge more explicit and more formal, so that the information in a wiki can be searched in better ways than just with keywords, offering queries similar to structural databases.”

Emerging commercial products

The first commercial products to exploit semantic web approaches are beginning to appear. Radar Networks has recently introduced its Twine service. Radar Networks is claiming to “pioneer the mainstream adoption of the Semantic Web, or what is sometimes called ‘Web 3.0’” (see Twine (2008)).

Implications

We observe three basic approaches to identifying and/or creating more effective personal information management in order to evaluate that effectiveness.

One is to create a unifying “super-app”: one program which does everything, bundling the (personal) world into a super PIM/GIM. Two major research prototypes have emerged which take this approach and use semantic web techniques (notably RDF and OWL). They are MIT’s Haystack (Haystack (2006)) and the Gnowsis project (Gnowsis (2008)).

A more conventional “super PIM” approach is being taken by a small Californian start-up company called NeoTech systems with their SQLNotes (2008) product (still on beta test at the time of writing; its name on appearance is expected to change to InfoQube). SQLNotes, should it ever stabilise and be completed, is close to a dream or ultimate “power user” PIM, being based as it is on the decade-old NetManage Ecco application’s approach Ecco (1997) but very much better integrated with Windows

and Office. Information can be stored in an outline, in a spreadsheet-like grid, or in rich text documents within a grid. SQLNotes even permits access to the relational tables which store its data. However, it may well fall victim to its own flexibility, because the flexibility is accompanied by a conceptual complexity which makes its usefulness difficult to grasp and its power difficult to manage.

The *second* is to take a federating approach in which minimal assembly or composing of emerging building blocks is undertaken: just sufficient to provide to a very small community of users, tools of sufficient usefulness to permit the hypotheses of this study to be investigated and evolved. This approach is explicitly espoused by Google and its CEO Eric Schmidt.

The *third approach* is consistent with the new phenomenon characterised in a recent conference as the “disappearing desktop” (see PIM (2008)). Increasingly capable client computers (typically smartphones rather than PCs, at least for a numerical majority of the world’s web users) will access semantic networks based on server computers. A server based approach is typified in the Radar Networks Twine product mentioned above in section 0. The biggest single architectural advantage of this approach is that it makes the mutual recognition of ontological tags very much easier. The corollary is that the approach might in practice favour the emergence of overbearing “common” tagging schemes which are not, in fact so much emergent as imposed.

References

- Allen, David (2001) *Getting Things Done: The Art of Stress-Free Productivity*. Penguin Books.
- Berners-Lee, T. (1998) 'Semantic web road map.' To be found at <http://www.w3.org/DesignIssues/Semantic.html> 1998
- Boardman, Richard (2004) "Improving Tool Support for Personal Information Management" - Abstract of PhD Thesis in Human-Computer Interaction. Intelligent and Interactive Systems Group, Dept. of Electronic and Electrical Engineering, Imperial College London, SW7 2TE, UK. Now at Google UE group.
- Cherbakov, L. & A. Bravery & B. D. Goodman & A. Pandya & J. Baggett (2007) 'Changing the corporate IT development model: Tapping the power of grassroots computing' *IBM Systems Journal* Volume 46, Number 4, 2007
- Denzin, Norman (1970) Strategies of multiple triangulation, in Denzin N (ed): *The Research Act*. New York, McGraw- Hill, 1970, pp 297-313.

- Drucker, Peter (1999) 'Knowledge Worker productivity: the biggest challenge.' In: The knowledge management yearbook 2000-2001.
- Ecco (1997) can be found at http://supportweb.netmanage.com/ts_viewnow/downloads/patchesUnsupporterd/ecco.asp
- Fichter, D. (2004) Tools for Finding Things Again. Online; Sep/Oct2004, Vol. 28 Issue 5, p52-56, 5p, 2 charts, 3bw
- Gartner Group (2007) 'Market Share: Enterprise E-Mail and Calendaring Software, Worldwide, 2004-2006.' 27 July 2007
- Gemmell, Jim & Gordon Bell & Roger Lueder (2006) MyLifeBits: a personal database for everything. Communications of the ACM Volume 49, Number 1 (2006), Pages 88-95
- Gnowsis (2008) is described at <http://www.gnowsis.org/> accessed 20-06-2008.
- Gregory M.R. & Norbis M. (2008a) 'Towards a Systematic Evaluation of Personal and Small Group Information and Knowledge Management'. Paper presented to 5th International Conference on Cybernetics and Information Technologies, Systems and Applications: CITSA 2008, in July 2008.
- Gregory M.R. & Norbis M. (2008b) 'The business of knowledge.' Paper given at the 8th International Conference on Knowledge, Culture and Change in Organisations, at the Cambridge University, United Kingdom between 2008/08/05 and 2008/08/08.
- Gregory M.R. & Norbis M. (2009a) 'Auditing Personal Information Management'. Paper submitted to the UKAIS 2009 conference.
- Gregory M.R. & Norbis M. (2009b) 'Evaluating Situational Applications Builders.' Paper submitted to a conference in summer 2009.
- Haystack (2006) is to be found at <http://simile.mit.edu/hayloft/index.html> checked 28/07/2008
- Ideaspace (2008) is to be found at <http://www.ideaspace.com/>
- Jones, William (2007) 'Keeping found things found: the study and practice of personal information management.' William Jones, University of Washington. Morgan Kauffman 2007
- KDE (2008) is described at <http://pim.kde.org/> accessed 20-06-2008.
- Kelly, D. (2006) Evaluating Personal Information Management Behaviours and Tools. Communications of the ACM; Jan2006, Vol. 49 Issue 1, p84-86, 3p
- Lotus Domino (2008) is described at <http://www-306.ibm.com/software/lotus/products/domino/> accessed 02-07-2008.
- Microsoft Exchange Server (2007) is described at <http://www.microsoft.com/exchange/default.msp> accessed 20-06-2008.
- Microsoft Office (2007) is described at <http://office.microsoft.com/fr-fr/products/FX100487411036.aspx?pid=CL100571081036> accessed 20-06-2008.
- Miller, Michael (2008) Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online. Que Books, 2008
- Oreilly, Tim (2007) What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software Tim O'Reilly Media Communications & Strategies, No. 1, p. 17, First Quarter 2007 Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1008839
- OWL (2009) <http://www.owlseek.com/whatis.html> checked 26/03/2009
- Penrose, Roger (1990) 'The Emperor's New Mind: Concerning Computers, Minds and the Laws of Physics.' Vintage, 1990

- PIM (2008) The disappearing desktop: PIM 2008 Conference on Human Factors in Computing Systems CHI '08 extended abstracts on Human factors in computing systems
- Simile (2008) is described at <http://simile.mit.edu/hayloft/index.html> accessed 20-06-2008. (2008) 'Simile is described at <http://simile.mit.edu/hayloft/index.html> accessed 20-06-2008.
- Strassman, P. (1999) 'Paradox revisited.' Computerworld, Vol. 33 Issue 36, pp. 40, 1999.
- Teevan, Jaime & William Jones & Benjamin B. Bederson (2006) Personal information management: Introduction. Communications of the ACM Volume 49, Number 1 (2006), Pages 40-43
- Twine (2008) is described at <http://www.radarnetworks.com/> accessed 28-04-2008
- Weir, D. & Hutchings, K. (2005) Cultural embeddedness and contextual constraints: knowledge sharing in Chinese and Arab cultures. doi.wiley.com
- Whittaker, Steve & Victoria Bellotti & Jacek Gwizdka (2006) Email in personal information management Communications of the ACM Volume 49 , Issue 1 (January 2006)
- Wikipedia - Semantic Wiki (2008) 'Semantic Wiki' found at http://en.wikipedia.org/wiki/Semantic_wiki