ENGINEERING THE INFORMATION AGE

WHERE NEXT: INFORMATION UTILITIES AND INFORMATION APPLIANCES

J M Taylor

Rapporteur: Thomas Rischbeck

(1)

1















	Units of Storage	gineering th ormation Ag
1 Gigabyte	= 1 billion bytes = 1000 Megabytes = 1KMbyte	
	= (1 thousand million bytes)	
1 Terabyte	= 1 trillion bytes = 1000 Gigabytes = 1KGbyte	
	= (1 thousand billion bytes)	
1 Petabyte	= 1 quadrillion bytes = 1000 Terabytes = 1MGbyte	
	= (1 million billion bytes)	
1 Exabyte	= 1 quintillion bytes = 1000 Petabytes	
	= 1 KMGbytes = 1 GGbyte	
	= (1 billion billion bytes)	presVeepre sadd/010

<u>Internet</u> Scale, G	t and Cyberspace Engineering the rowth, Complexity			
PC's on Internet:	1996 - 30m; 1998 - 100m; 2000 - 230m			
Host computers:	1998 - 30m; 1999 - 43m; 2000 - 100m			
WWW Sites:	1997 - 1.7m; July 1998 - 10m			
URLs:	Dec 1996 - 72m; Sept 1997 - 200m			
Storage per Server:	1998 - 50Gb; 2000 - 500Gb+			
Storage per PC:	1999 - 4Gb; 2000 - 20Gb+			
Total Internet Storage:	1998 - 2000m Gbytes			
	2000 - 50,000m Gbytes			
	- 50KMGbytes			
	-Increase 25x in 2 years			
	= <u>50 exabytes</u>			





÷

Engineering the Information A **Giant Information systems** The Global Storage Explosion We need information search and management technology which will: scale to exabyte-order global distributed information resources handle image & video data handle distributed persistent hypermedia linkages We need information infrastructure technology to manage Giant Information Systems global telecomm+Internet+Web is >I billion computers 1 in 4 people owning 10 information appliances is >10 billion computers Ref: slidepresVeepresadd/010















ŝ







	Digital Photography	Engine
One high	n quality digital photograph :	1
 2m pix levels (els each with 3 colours and 256 b (8 bits)	rightness
One pho	tograph needs 50Mbits of storage	;
1 Megaby	yte after compression	
Digital w	edding album requires 100 Megat	oytes
(approx 2	20% of a CD-ROM)	
100 albu	ms over 15 years:	
=	1.4 Gbytes per year	
) =	3 CD-ROMs per year	
=	1 PC hard disc drive per year	

IMAGE AS DATA *Cheap imagers in personal products *Everything networked *Cheap viewers & printers *1000 million cameras on the WWW *Home, work, lifestyle, leisure *Wasteful - like the phone! *From saving time to spending time



















7.4.0

Collaboration Treasure				
Comments	Motive	Relationship	Group	Date
 high value cultural challenges 	evaluate BN potential	~18 month contract	DERA	96/97
long term	encourage relevant basic research	equipment grants	U. Calabria, Italy	96-
 networking multi-party leadership challenges 	exploration of synergistic areas	Foresight novel optics project	DTA	97-
	computational	CASE award	U. Bristol	98-
	modeling			
	specific measurement technique	Short term contract	U.Exeter	98-

10











0









DISCUSSION

Rapporteur: Thomas Rischbeck

Due to time shortage all discussion was deferred to after the talk. The discussion then started with Professor Cockton enquiring about the rough plans of UK research councils in the area of information organisation and management in the light of his involvement with similar programmes in Finland. Dr Taylor expressed concerns about nobody actually mounting corresponding proposals for basic strategic research and pointed out the necessity for fundamental research.

Further discussion then focused on the ever-recurring topic of information management. Professor Jones raised the opinion that much storage space is actually filled up with multiple copies of the same data. Dr Taylor spoke of an explosion of informal digital information from different kinds of sources, e.g. digital cameras or digital video. New standards of information organisation and also new concepts are needed. Those might include "forgetfulness" to filter out the "0.1%" of data we really want to keep.

Dr Taylor repeatedly appealed to the research community actively to engage in the development of underlying science in order to make new technologies easy to use, nondisruptive, well-behaved, safe, trustworthy and affordable. Despite the need to achieve this set of properties, Professor Mamdani said that one should also consider what is "already going on" and achieve at an appropriate level of coping with current technology. Key delimiters for next generation scale-up are science and fundamental research. However, as researchers and technology forecasters we are faced with the responsibility to assess the impact of avalanching technologies, which is almost impossible. Insofar, coping with current technology is an important aspect of our role.

In addition, Professor Mamdani explained the "don't worry" attitude towards information management. He gave the example of the home attic which is organised by the technique of throwing away unused things. Similarly, accumulated data could be cleared out by getting rid of "rubbish". If information is not visible any more it is probably not important and can be discarded.

Professor Randell then went on to issues of security and dependence. He recalled Professor Neumann's seminar on Tuesday/Wednesday and in his absence gave a voice to his opinions. Professor Neumann would have queried fundamental problems in the base infrastructure, like the internet and deployed operating systems. On the one hand commercial pressures, like short-time to market imply a lack of robustness and lead to the "laughable security of PC operating systems." On the other hand, government pressures, like restriction of cryptography, exist. Both parties need to rethink their view as in the future they can bring about a solution for getting the quality of infrastructure that is needed.

Dr Taylor agreed with a governmental responsibility for new laws and their enforcement to protect citizens from crime and fraud in cyberspace. Security of enduser systems is important. Part of this is the release of cryptography technology and in parallel the prevention and prosecution of fraud. He was sceptical about doing it slowly to get everything right. For the moment only "good enough" fraud detection is needed. Professor Cockton gave the example of Y2K compliance which was enforced by insurance clauses rather than governmental pressure.

Returning to issues of data management, Professor Mamdani suggested "rusting memory". Information in a system could decay automatically; if you really want to keep it, explicit action would be necessary. This property could be "built into data itself". Professor Randell added that in fact an essential property of data is that it is best used by data. Mr Rowley contributed the implementation idea of expiry date and owner attached with each data item. If the expiry date comes, a couple of warnings could be issued to the owner, before final removal from storage. The owner's responsibility for individual data items is a key concept - "orphan" data is of no value.

Professor Jones mentioned that there is not only the issue of finding information with unknown location. Another problem which contributes to the explosion of information originates from the greatly lowered threshold of publication in times of the internet. Also, Professor Tedd pointed out that organisation of the world-wide information (or even of a restricted subarea), though badly needed, won't be attainable. The reason being that "the world is not organised" which opposes the creation of a top-down catalogue of all accessible content. Dr Taylor countered that his vision of organising information is given by a collection of networks with decentralised "intelligence". Scholarship over the next 50 years could provide different threads through all the material.

Professor Martin reinforced the idea that the process of organising information implies the need for a valuation model. Market value which manifests itself with the statement "if they don't pay for it, it's not worth keeping" is a possible candidate. This is not desirable, because aspects, like social value, heritage value or cultural value should be taken into consideration, as well. Certainly, valuation cannot be done with universal validity. A side-look at sociology gives hope here: There is no universal community or society, "but we get along" as a self-organising society.

Professor Nygaard mentioned a religious and ethical niche problem of information creation and deletion. He questioned whether if an author dies, her "web page stays alive?" Is there an obligation of preserving this heritage or can it be deleted? Should the web pages of outstanding personalities be kept longer on storage before deletion?

Mr Rowley expressed concern that the creation of the so-called "docuvers", the single document that links to every other information is the death of IPR. The technology to create exact copies of the original is at every body's disposal, cheap and easy to use. This is a completely new situation compared to the effort required to copy a book for example, which makes it impossible to prosecute copyright infringement. Technology, which provides evidence for copyright infringements, e.g. watermarking must be underpinned by appropriate laws as Professor Mamdani said. At the moment, we face an impotence of national governments in defining and regulating those issues. National differences in jurisdiction are a problem.

Professor Mamdani pleaded for a more flexible interpretation of the term intellectual property than the one currently applied by lawyers. Not only the current owner should be considered, but the whole "value chain" in which information is generated and exchanged. There are many different ways in which generators, designers, editors, organisers or distributors can share revenue.

Dr Taylor agreed that new ways of how people can add value to intellectual property have to be identified. Again, he stressed the point that careful thinking and pragmatics are necessary before these issues can be dealt with properly. To protect intellectual property, responsive action can then be taken by advising appropriate jurisdiction, possibly including changes to the constitution.

Mr Yapp then pointed out the limits of current rasterisation technology. Many artifacts, so he said, cannot be digitised without destroying them. For others the digital representation is very different from the original, e.g. a scan of a Van Gogh painting has no depth information. Dr Taylor said that of course, digitised versions are not a replacement for the experience of the original version in many cases but mcrcly a high-fidelity version of it with greatly improved accessibility for millions of people. 3D scanning technology will not be available in the near future. However, as Professor

Randell mentioned some things are actually more legible through digital imaging, as for example the discipline of digital restoration proves.

Summarising the week, Professor Cockton said there were technological factors and occasionally social shaping factors mentioned. In his opinion, the coinciding design process was not emphasised enough. Dr Taylor answered that design professionals need to think radically different about the new developments. I.e., how do you add to a giant evolving system and at the same time protect your subsystem. Ad hoc systems and end user systems must be considered from the point of view of what they deliver to the user. Research is required to identify enabling standards and pieces of infrastructure that have to get into place beforehand.

Chairman Professor Randell thanked the speaker and the audience.

General Discussion

There followed a more general discussion of both talks. Professor Martin commenced with a criticism of the current practice of standardisation as exemplified by DAVIC (Digital Audio Visual Council) or the Intelligent Agent standardisation process mentioned by Professor Mamdani. Neither included assumptions about specific configurations and value chains built into the technical architecture. Although there has been some development, no distinction can be made between a political process and a standardisation process. The distinction between political and commercial assumptions must be made explicit, as it is an undeniable fact that industrial protection is necessary for all standards. As Professor Mamdani said, the overall goal of the standard making process is to reach a common agreement on which way to go based on the technological offers and proposals handed in by participating companies; also risk management is an important aspect. However, Professor Martin believed that often the "sort of business a company is in" determines their degree of influence in shaping the new standard. Likewise, as Dr Kay noted, lobbyism is another important factor instead of the desired confluence of participants

Returning to IPR issues, Dr Kay was reminded of a point of Dr Taylor's talk saying that one of the reasons for the success of the current IPR regime is "that we have the right to break the rules." Technological attempts to establish a new regime have failed to date as their viability is limited by restrictiveness. Although the current regime does not imply a guarantee "to get away with having broken the rules" as Professor Rousset de Pina expressed, often a "very flexible interpretation" (Dr Kay) is possible. Professor Mandani noted that this is made possible by a body of case law which is not available for new regimes. For the issues discussed this will be available sooner or later. The law profession is considerably well informed and a lot of effort has been put into content-related case law by both judges and law firms.

Problems in the legal framework according to Dr Kay are the current emphasis on copying rather than theft. In analogy to physical goods an owner is deprived from getting value of his (intellectual) property. Moreover, grey areas exist, e.g. the act of caching a website is technically against the law.