REVIEW

Memories for life: a review of the science and technology

Kieron O'Hara[1], Richard Morris[2], Nigel Shadbolt[1], Graham J. Hitch[3], Wendy Hall[1],* and Neil Beagrie[4]

[1]*Intelligence, Agents, Multimedia Group, School of Electronics and Computer Science, University of Southampton, Highfield, Southampton SO17 IBJ, UK*
[2]*Division of Neuroscience, College of Medicine and Veterinary Medicine, 1, George Square, Edinburgh EH8 9JZ, UK*
[3]*Department of Psychology, University of York, York YO10 5DD, UK*
[4]*The British Library, 96, Euston Road, London NW1 2DB, UK*

This paper discusses scientific, social and technological aspects of memory. Recent developments in our understanding of memory processes and mechanisms, and their digital implementation, have placed the encoding, storage, management and retrieval of information at the forefront of several fields of research. At the same time, the divisions between the biological, physical and the digital worlds seem to be dissolving. Hence, opportunities for interdisciplinary research into memory are being created, between the life sciences, social sciences and physical sciences. Such research may benefit from immediate application into information management technology as a testbed. The paper describes one initiative, memories for life, as a potential common problem space for the various interested disciplines.

Keywords: memory; information storage; information management; interdisciplinarity

1. INTRODUCTION

The interface between the physical world and the digital world seems to be blurring, and becoming less determinate (Abowd *et al.* 2002). It has long been recognized that the interaction between technology and human society can have far-reaching psychological effects (Ong 1982). Thinkers as early as Socrates and Plato focused on memory as one faculty of mind for which technologies of storage could change individuals' psychological makeup, by, so to speak, externalizing or 'outsourcing' mental function. In recent years, the development of such commonplace innovations as email, ubiquitous computing (including the Internet), virtual reality and advanced prosthetics have brought home the requirement for an increase in the scientific and social understanding of cognitive function, in order to design and evaluate appropriate technological devices.

Memory is by no means the only relevant area for interaction at the interface of the mind and the digital, but it is a very exciting one, as evinced by the impressive convergence evident from research reviews commissioned by the United Kingdom's Cognitive Systems Foresight programme (Morris *et al.* 2006*b*).

The ability to co-opt electronic media for the storage of personally relevant information gives rise to the notion of memories for life (M4L), currently being discussed as a 'grand challenge' for computing (Fitzgibbon & Reiter 2003; Shadbolt 2003*b*), to define and solve the problems caused by people storing increasingly large quantities of information about themselves.

Memory for life is a research problem, and a problem space—but what problem, and why now? The use of electronic media for supporting human information storage and recall needs defines an area. We have always had artificial aids to memory of course; the twenty-first century twist is that suddenly we are presented with the possibility of memories for *life*.

Paper survives, but not predictably. Our knowledge of Ancient Egypt, for instance, stems partly from the accidental survival of certain papyri from various rubbish tips (cf. Hunt & Edgar 1932); we have no idea whether these agreements, letters, wills, accounts and charms are representative or not of social and business life in Hellenistic Egypt. We can doubt whether the Egyptians themselves, who of course cared deeply about posterity, would have selected these papyri had they been commissioned to set up a time capsule. But the use of digital and electronic media moves information storage on from paper; survival can now be managed.

It is now possible to store digital versions of life's memories. As Alan Dix playfully noted, it takes $100 \, \text{kbits s}^{-1}$ to get high-quality audio and video. If we imagine someone with a camera strapped to his or her head for 70 years (2.2×10^9 s), that is something of the order of 27.5 terabytes of storage required, or about four hundred and fifty 60 GB iPods. And if Moore's Law continues to hold over those 70 years (admittedly a large assumption!), it would be possible to store a continuous record of a life on a grain of sand (Dix 2002).

Of course, memory storage is unlikely to be continuous over a lifetime, but it could be very rich. The variety of information captured varies between the home and the workplace, and ranges from documents and emails to digital photographs, video and so on. The 'information overload' problem, long recognized in the computing industry, is becoming a major issue in the workplace for everyone and for the private citizen too, as retrieving and selectively deleting (forgetting) these data become ever more of a challenge. And the challenge is now directly related to our conception of ourselves, as the information we collect is increasingly generated by us, not by outside observing bodies. The information will have some sort of role in defining our identities, over a lifetime.

In this space, interdisciplinarity is crucial. The issues of storage, retrieval and forgetting have analogies across a range of sciences, and yet these analogies are by no means fully understood. For instance, in cognitive psychology, the problem of understanding selective attention—the processing by which an abundance of sensory information impinging upon us is filtered to enable a manageable flow of information for the brain to handle—has been studied for many years (e.g. Broadbent 1958). But how many of the insights that have come from this work have had an impact on the design of operating systems used in desktop PCs?

A large number of disciplines may contribute to M4L; in this paper, we focus on a small but central subset of them. At a minimum, the M4L research programme will require input from mechanistic studies of the brain (neuroscience), the human mind (cognitive psychology), the structures and limitations imposed by human society and human social behaviour (sociology), information technology (computer science) and management (knowledge management). Of course, this is not an exhaustive list of relevant disciplines of those who could contribute.

But how can we move from analogy and metaphor (different ways of understanding 'memory') to concrete two-way interaction between disciplines? M4L may or may not end up as a recognized subdiscipline, but what seems inevitable is that the various disciplines, all looking at memory, understood differently, with different methods and jargons, will move towards each other as scientists begin to understand each other's languages and methodologies. The first stage will be converging understanding of the problem space (convergence of language and method); a second stage will be learning from results in different disciplines (for instance, computing memory 'borrowing' structures from psychology, such as working memory and long-term memory). The third stage of interaction will be closing feedback loops and genuinely collaborating. We are, hopefully, well on the way to achieving stage one. And the prizes for going beyond this stage are great.

Analogous to the switch from orality to literacy centuries ago, the digital revolution gives rise to momentous opportunities. For instance, the existence of large quantities of multimedia information recorded from a life may benefit from research to integrate information from distinct media to create a narrative of events that may be of value to a person's descendents. Intelligent querying of large data stores, mapping of deep structures in such stores and development of user models are some examples of important technical challenges that present themselves immediately, issues being confronted by the research teams developing the Semantic Web (Berners-Lee *et al.* 2001). Analysing stored information to model a person's lifestyle (with potential commercial implications in the insurance and security industries), behaviour and health (the 'virtual general practitioner'), or intelligent Web pages that can adapt themselves to a person's linguistic and other competence are examples of more ambitious applications that we might look forward to in the coming decades. Developing prosthetic memories for those with impaired cognitive function may seem too far fetched to be worthy of discussion, yet projects with precisely this in mind are currently underway.

Our aim here is to discuss recent progress in these exemplar fields, and suggest how this disparate work might be pulled together to address a common problem space, which we might call the grand challenge of M4L. The structure of the paper is as follows. We begin with a brief review of the different disciplines' conceptions of memory, partly developed by theoreticians, partly as a function of methodology (§2), including relatively extended reviews of recent developments in the science of human memory (§2.1), our understanding of memory as a social construct, and as embedded in material culture (§2.2), and developments in digital storage of information (§2.3). As our conceptions of memory evolve, various social issues become relevant. Section 3 examines the context of memory in digital form, and reviews some of the issues following from the need to transmit memories in other than haphazard ways, and the perhaps unexpected difficulties there are in preserving supposedly permanent digital memories for longer than a few years. And when digital information is preserved, transferable and, crucially, searchable, its ethical handling become even more important; some of the issues here are reviewed in §4.

Having reviewed the scientific, technological, social and ethical context, we will then be in a position to describe the grand challenge of M4L in more detail (§5), setting out a technological challenge and showing how it is amenable to an interdisciplinary response, before a final set of concluding remarks (§6).

2. CONCEPTIONS OF MEMORY

The term 'memory' is used in numerous ways and has several technical definitions. In everyday discourse, memory is generally used to refer to the act of bringing to mind information that is retained from the past. We

all speak of remembering having seen someone yesterday, or that Paris is the capital of France. There are important differences between these two instances of memory, which we shall come to, but both are propositional. Memory as an 'act' is commonly distinguished from the mere execution of a learned skill and, in this way, folk psychology embodies Ryle's (1949, pp. 26–60) distinction between knowing that (propositional knowledge) and knowing how (skills)—a distinction incorporated into artificial intelligence as early as the discussions by the Dartmouth Conference Group in the late 1950s. Were we to ask someone whether they could remember how to ride a bicycle, we would not be surprised if they replied 'I don't know—let's get out the bike and see'. Skills are not propositional in the same way as declarative forms of memory; they are forms of procedural knowledge that are embedded into the behavioural expression systems of which they are a part.

More technically, memory can refer to the mental faculty of retaining information about stimuli of some sort when those stimuli are no longer present. An organism (or robot) that can do this is said to have 'a memory'. Second, it may refer to the contents of that storage system rather than the system itself. This is closer to its everyday usage—that phenomenological sense of having a memory *of* something. Third, there is the much broader reference to the family of processes and underlying mechanisms that implement the various forms of memory, and different levels of analysis at which these may be studied. Their study is the substance of large areas of neuroscience and cognitive psychology.

Is there a key logical concept that marks off memory from other cognitive capacities, such as perception and attention? The generally accepted view is that there is in the case of long-term memory, which entails some conjunction of recording events and the passage of time. Specifically, something happens at time t_1 that causes a change in the organism, such that something else at time t_2 is influenced by the prior event at t_1. What has taken a long time to appreciate, an appreciation that is a product of science in the twentieth century, is recognition that this 'influence' can take many forms (such as propositional versus procedural). One of the challenges for the twenty-first century is to understand how these different forms of memory are implemented and to use this burgeoning knowledge to develop effective engineering devices that can emulate and externalize these forms of memory to advantage. In contrast, it is increasingly realized that attention and perception are intimately concerned with a further form of memory known as working memory, which is concerned with the special function of maintaining and transforming temporary information.

The disciplines that need to be brought together by the M4L vision are inevitably influenced in their interpretation of what memory might be by the methodologies that accompany them. For instance, neuropsychologists are now endeavouring to map psychological processes onto neuroanatomical structures and networks using the analysis of selected neurological patients, and both functional magnetic resonance and other imaging techniques. Those at a more fine-grained level of analysis focus on the biochemical pathways that mediate changes at the level of the neuron and its sub-cellular components, including the synapse (Ahmed *et al.* 2006; Morris *et al.* 2006*a*). The logical concept of memory entails that it should be detectable by some change in a brain state that is doing the storage (although in practice, this may prove very difficult to identify—the so-called 'needle in a haystack' problem). One focus in the neuroscience of memory is on investigating what such states might be, with the strength of synaptic connections between neurons being very important here (Martin & Morris 2002; LeDoux 2002). Another interdisciplinary focus is the use of computer modelling to implement biologically realistic models of synaptic modification and topological rearrangement in synthetic nervous systems (Elliott & Shadbolt 2003).

The investigation of networks of neurons is complex enough requiring technology ranging from multiphoton confocal microscopy to image living synapses through to ensemble single-unit recording to get a handle on neuronal networks. On top of that, there are also questions about brain development, of how the brain is shaped by genetic instruction, of how the environment has its effect. This inevitable focus on these complex systems has meant that neuroscientists have, until recently, worked with a highly individualistic conception of the person, where the environment is encoded largely through sensory inputs, and systems are small networks of neurons that can store or retrieve information (cf. LeDoux 2002, pp. 31–32). This perspective is changing with new research programmes focusing on the cognitive neuroscience of social interactions between people, concerned with how higher-level mental functions are implemented in the brain (Blakemore & Decety 2001). It builds on the early success of cognitive neuropsychology in accounting for cognitive impairments resulting from brain damage or disease, but goes further to investigate the neural basis of cognition in the healthy brain using new technologies, such as functional neuroimaging (Posner & DiGirolamo 2000).

Areas such as knowledge management, in effect, treat organizations as relatively determinate structures embodying processes of knowledge production, manipulation, retrieval or storage (Nonaka & Takeuchi 1995, pp. 56–94). Corporate memory is therefore the focus of information retrieval strategies (Brooking 1998), and memory is then seen as the capacity of an organization to generate the key information for its information processing without acquiring it afresh. A corporate memory is therefore likely to be heterogeneous in form, taking in paper-based resources, computer files and actual memory or know-how in people's heads (Douglas 1987). It will also be closely associated with a corporate context, which may or may not be explicitly modelled (O'Hara & Shadbolt 1997; Schreiber *et al.* 2000, pp. 25–67). Pressing issues are the organization of the corporate memory, its representation and its retrieval (both in terms of how to fish the right information out of the repository, and of which repository to search in). An issue, then, that

immediately arises, of great relevance to our own concerns, is the extent to which technology can and should be applied. Projects such as Advanced Knowledge Technologies (AKT 2001—http://www.aktors. org) seek to discover ways of exploiting new technologies, such as the Semantic Web, to ease the creation and transfer of knowledge through its organizational life cycle (Shadbolt *et al.* 2004*a,b*).

Sociology differs by having a much less functional view of the surrounding environment. Hence, the workings of social memory become relatively difficult to uncover against what is in effect a highly indeterminate backdrop. Memory has been seen as a dialogue between present and past, as the past has to some extent to be articulated in order for its effects to be seen in the present (Huyssen 1995). That articulation of the past may be expressly in order to remember, to see the past as past (as with various mechanisms, such as commemorative artefacts like statues, storytelling practices, photographs or texts); or may be involved in the reproduction of useful performance in the present, inscribing the past in the present as present (as with habit or other repetitive learned behaviour, known in sociological theory as *habitus*; Bourdieu 1977). Shared memories also have a quasi-subjective property, which implies a contrasting methodology from that used; for example, to investigate more objective evidence in history, although drawing the dividing lines here is non-trivial (Misztal 2003, pp. 99–108).

Meanwhile, computing focuses on the provision of information storage capacity. Here, the attention includes keeping costs low and reliability up. There is much less interest in the quality of the knowledge that is being recalled, this being the responsibility of the user. Research bifurcates between the development of hardware and software. What has drawn the eye in recent years is the continuing relevance of Moore's Law, that storage capacity and computing power of chips have doubled every 18 months and will continue to do so, which has endured remarkably ever since Moore made his observation in 1965. In pursuit of this aim, engineering researchers are currently examining new types of hardware, as the limits of the standard silicon chips are in sight. Fast non-volatile memory is beginning to appear on so-called magnetic RAM, chalcogenides that change shape when an electrical charge is applied to them and carbon nanotubes. However, as the amounts of information being stored continue to increase exponentially, the requirement for sensible and intuitive organization methods to facilitate retrieval becomes ever heavier. Such retrieval methods may simply involve advanced knowledge-based reasoning, but may also include neurobiologically inspired methods, such as content-based retrieval. As with neuroscience, the methodological decision here will have a strong effect on which other disciplines it will be easier to integrate this work with (O'Hara *et al.* 2006).

The growing subfield of nature-inspired computing (Shadbolt 2004*a*) involves the harnessing of processes (or analogues of processes) found in the natural world, such as phylogeny, ontogeny and epigenesis, to discover biologically realistic methods of processing information that are very different from the computational architectures; they can implement very high-level operations without top-down planning. From the early days of nature-inspired computing, situated, embodied agents were being built which exploited features of the environment to support information processing (e.g. by not storing information, but using sensors to rediscover it anew). In effect, the environment acted as the memory of such agents (Steels & Brooks 1995).

Computational modelling of memory processes has been discussed as a potential area that could work to help bring theoretical accounts of memory together (Morris *et al.* 2006*a*). For instance, nature-inspired computing, network computing and neuroinformatics could all play a role in bringing together diverse researchers to try to understand aspects of brain function. The argument of this paper is that M4L, as a cluster of research projects, will help dissolve certain communication and comprehension problems between researchers into the brain, human psychology, society and the digital sharing and archiving of information about the past.

Finally, we must not forget an important part of memory—forgetting (Schacter 2001). The loss of stored information is again conceived very differently across the disciplines. In the human sciences, forgetting—though sometimes a dysfunction—is often a boon, a form of mental housekeeping that usefully gets rid of information that is out of date, unlikely to be required, or traumatic. Schacter argues that forgetting is an inevitable consequence of a mental system that ordinarily works very well and that the various manifestations of 'normal' forgetting are extremely useful tools for researchers that are studying how the biological system is organized. Still, forgetting can be troublesome, such as when one forgets the name of someone one has met the day before; or even debilitating to a point where a person's capacity to live independently is compromised. On the other hand, in computing, forgetting is almost always a failure of one sort or another; information once stored on a hard disc should stay there. Knowledge management stands somewhere in between, with the added wrinkle that forgetting can sometimes be anticipated (e.g. when an expert decides to leave the firm, or retire), and it may then be an economic decision whether to bribe the expert to stay, to employ a new person and perhaps even ensure some overlap between the contracts of the new and the old, or to engage in some other knowledge acquisition exercise (Cowan *et al.* 1999). In sociology, the forgetting of aspects of the past strongly affects the political myths in a society, and so the investigation of forgetting is connected with power structures, and who is able to drive social memories underground (O'Hara & Stevens 2006). Socially and politically, forgetting stands alongside forgiving; it may be the latter option that strikes the better balance between avoiding present-day conflict while respecting those who have suffered in the past (Margalit 2002).

While we do not discuss forgetting in detail in this paper, for reasons of space, we recognize that it is a very important issue in the context of M4L. Two research questions immediately leap to mind guided by the

prospects of externalizing memory using technology or, almost the opposite, emulating the failings of human memory in technology in pursuit of devices that are more user-friendly. First, can technology reliably be harnessed to lower the burdens of memory on the human subject, by greater externalization of the information we need from memory in our daily lives? As we move from the 'yellow pages' of the Gutenberg world to the Google of the twenty-first century, carrying with us the permanently connected PDAs of our personal digital environments, Morris *et al.* (2006*a*) have already been moved to wonder, as Socrates did in Plato's *Phaedrus*, whether 'the need for an endogenous memory [will] become a thing of the past'. This is unlikely, although many are already in the workplace not bothering to file information that is better accessed anew through Web searches (e.g. airline timetables). Conversely, confronting the assumption in computer science that forgetting is always a 'bad thing', Morris *et al.* (2006*a*) have also wondered whether certain forms of software should not more explicitly emulate the patterns of forgetting of the human mind. That is, given constraints on the techniques of intelligent search, and requirements for the trustworthiness of large heterogeneous data stores such as the World Wide Web, should forgetting techniques be imported into computer science as well? And in particular, should such techniques be biologically inspired (O'Hara *et al.* 2006)?

Our argument so far has been that integrating the various disciplines relevant for ventures such as M4L will be a formidable task. But the advantage of the M4L challenge is that it does provide the functional centripetal force of a common problem space. We now offer a brief survey of the relevant disciplines, to show where discoveries and progress are being made that will help address the problems of M4L, and incidentally to demonstrate where integration is being achieved already. The surveys here are of necessity highly selective, and show only a fraction of the relevant work in the various fields.

2.1. Psychological and neural conceptions of memory

Very impressive synergies have been revealed by the intersection of experimental psychology and neuroscience with respect to studies of human memory. A key and very influential discovery is that there is no single memory system in the mind, no device which does all the work of what we call 'memory'. Humans (and animals) have multiple memory systems—distinct systems for processing, storing and retrieving information of different kinds that integrate smoothly enough to give the illusion of a single faculty. We also have an array of memory processes, such as encoding and retrieval, and these reflect different brain states and, with these, different states of mind.

Various psychological taxonomies of the multiple types of memory exist. All of these divide memory with respect to both capacity and persistence, with short-term or 'working-memory' systems having limited capacity and persistence but high-fidelity, serving as a

central workspace for bringing together and transforming information from other memory systems, closely linked to attention. In contrast, long-term memory serves as the ultimate repository, mainly passive, of vast quantities of propositional information and skills. Long-term memory has, for example, been taxonomically divided by Tulving (2002) into perceptual–representational systems, semantic memory, episodic memory and procedural skills. This classification fares reasonably well, but is not without its problems (Morris *et al.* 2006*a*). An important omission relates to emotional or 'value' memory, in which specific stimuli may evoke feelings of pleasure or sadness. Another important distinction to bear in mind is that between implicit and explicit memory. Only the latter involves active conscious awareness of the contents of the memory system. Implicit memory, in contrast, can still be propositional, but is unavailable to conscious awareness—a distinction that is difficult to grasp, but will prove very important in the engineering of new memory devices.

The neuroscience dimension comes in at numerous levels, as already noted. The active memory system that we use for manipulating and 'working on' verbal and visuo-spatial information—Baddeley's (1992) working memory—involves sub-systems that are now thought to be located in specific parts of the brain. Brain imaging studies have revealed that some of these are located in the frontal lobe, others in posterior regions, and that they interconnect in a network-like manner as a function of the task being performed (Fletcher & Henson 2001). The perceptual–representational components of long-term memory are presumed to be intrinsic to the sensory–perceptual components of the various sensory systems, and include front-end modules for identifying colour and motion, specialized modules, such as the face-processing system of the superior temporal gyrus (Perrett *et al.* 1992, 1998) and the fascinating mirror-neuron system of the parietal lobe (Rizzolatti *et al.* 2001); mirror-neurons being a set of neurons that simultaneously encode actions (e.g. a monkey grasping an object) and the watching of actions (e.g. a monkey watching another monkey grasping an object). Semantic and episodic memory are both thought to involve numerous cortical modules as well, and there is currently much discussion about the specific role of the hippocampus of the medial temporal lobe in the formation and consolidation of new episodic memories. Some hold that the hippocampus is equally involved in the formation of all declarative memories, episodic and semantic (e.g. Bayley & Squire 2003), while others believe the cortex alone is capable of encoding and storing new semantic information, but requires the spatio-temporal contextual information provided by the hippocampus for new episodic encoding.

As discussed by Morris *et al.* (2006*a*), other levels of analysis in the neurosciences take the study of memory beyond the level of 'where in the brain' to the study of the physiological patterns of activity required for new memory encoding, consolidation or retrieval; thence to the sub-cellular sites at which physical changes happen in the neurons, particularly alterations in synaptic

efficacy; and on to the activation of biochemical signal-transduction pathways and genes, whose protein products are essential for normal memory function in diverse circuits. One contentious issue is whether the cellular mechanisms of neuronal and synaptic plasticity have long been conserved, such that the mechanisms in the human brain are little different to those expressed in the lowliest sea-slug, such as *Aplysia*. If true, this idea has a beautiful simplicity about it, for it implies that there are only a few underlying mechanisms by which neurons can alter their connectivity (e.g. pre- and postsynaptic mechanisms). It also implies that these evolved early and that they have been re-used, like letters of the alphabet, in newly evolved circuits of the vertebrate brain that enable more complex forms of information processing. The alternative is to suppose that evolution is more likely to operate at all levels, with certain connectional mechanisms re-used to be sure, but new ones developed at the gene, protein, signal-transduction pathway, synapse, cellular and circuit level.

From the perspective of M4L, all of these distinct forms of memory are relevant to the development of novel computational software that might be used for organizing and storing information about a person's life in a manner that would later be accessible. We shall highlight two topics that have been extensively studied in cognitive psychology and neuroscience, respectively, that are likely to be relevant to workers in other disciplines: (i) the distinction between semantic and episodic memory and (ii) the problem of memory consolidation.

Semantic memory refers to the storehouse of facts that we know about the world—such as our knowledge that Paris is the capital of France. Episodic memory, in contrast, involves memory for individual events, with the act of remembering involving mental time travel to the time and place where this event happened in one's mind's eye (a capacity that some suspect to be uniquely human; Tulving 2002). A neuropsychological patient who knew that he or she was married but was unable to remember anything about where or when this event happened, or who was there at the time, would be said to be suffering from a (severe) deficit in episodic but not semantic memory. The importance of this distinction for M4L is that computing software designers will need to take on board that the human memory operates seamlessly with these two forms of memory for a very important reason. Put simply, to operate effectively, the system needs to forget all the spatio-temporal tags of episodic events that accompany information that gradually finds its way into semantic memory if memory is not to be incapacitated by handling a large volume of irrelevant information. It may be interesting to remember one's first visit to Paris, or the person with whom one walked hand-in-hand down the banks of the Seine, but the issue at hand in a busy life is very rarely going to require this additional information when all that matters for the decision to be taken is knowing that Paris is the capital of France. The same issue arises as children acquire general knowledge as they grow up and go through the educational system. So far so good. The scientific problem is how does one best go about

constructing semantic networks? How does the brain do it? And why are we typically unable to recall episodic memories from the very earliest years of our life? Might this 'infantile amnesia' be an adaptive feature of the human memory system? These are topics of widespread interest with efforts ranging from experimental studies through to computational modelling. And how might we go about doing it on the Web? In sharp contrast to the mental reflex of certain computer scientists that forgetting is a 'bad thing', the message from psychology and neuroscience is that forgetting is vital for effective function. The decreasing cost of silicon memory devices could be leading computer scientists down an unwise path.

The problem of memory consolidation is a parallel problem. Consolidation refers to the acquisition by memory traces of some measure of permanence. Information may get into long-term memory and be stored effectively, such that a person can recall relevant information the next day—well beyond the domain of working memory—but it may not last. Either the memory traces themselves do not last, or the associations may be insufficiently connected to other information in memory to be readily accessible. The process that enables recently formed long-term memories to last is generally referred to as 'memory consolidation' and it is thought to be a selective process. In part, memory consolidation is one of those topics in neuroscience that could be said to fall into the 'merely interesting' category, in that a current focus in brain science is on the specific signal-transduction pathways that neurons use to ensure the permanence of the physical traces that mediate memory. However, there is a growing body of evidence to suggest that the physiological basis of long-term memory has implications for higher levels of analysis that other disciplines may wish to take on board. Specifically, physiological studies indicate that some information seems to enter long-term memory automatically, but whether it is retained depends upon whether that information is subsequently subject to consolidation.

This is a useful trick for the nervous system to have, because the 'decision' about whether to retain or lose information (effectively forever) need not be taken at the moment the remembered event itself actually happens. It can be temporally discontiguous with the encoding operation. Events may be happening too quickly for the consolidation mechanisms to keep up, other later events may occur that render earlier information important, irrelevant or misleading, it may be valuable to wait until the body is in a state of quiescence (e.g. sleep), and so on. Here again, it may be possible in a digital device to make very fast 'decisions' about storage and to interleave new information with old so quickly that this apparent design fault of the human brain is finessed by superior silicon technology. However, while time waits for no man, computers must wait for time. It is very often the unfolding events of the day that determine what earlier information is worth retaining, and the interleaving process for storing information can also be one that necessarily involves re-activating memories and then integrating new information on the basis of some conceptual reasoning

process, or even replacing old false information; as Anderson (1993) has argued, the neural processes underlying memory are adapted to the statistical properties of the environment. It is unlikely that these processes can be speeded up much faster than the human mind—they need to run in real time. M4L scenarios are very likely to involve the development of software that knows about the circadian cycle of the human day and uses this knowledge in interesting and novel ways.

2.2. Social and material memory

Memory is increasingly under examination in the social sciences and humanities, where, analogous to the study of the memory of the human individual, the main foci are the functions that are performed by social memory, and the (physical and institutional) mechanisms that underlie those functions (Misztal 2003). The functional connections between social memory and wider social and political processes can be very far-reaching; so much is in agreement. However, it is equally true to say that there is considerable disagreement over how 'social memory' or 'collective memory' should be characterized, which makes the area particularly attractive for interdisciplinary study. Some issues pertaining to social memory include the following.

(i) *Commemorative activities.* Changes in politics, for example the move from party- or ideology-based politics to identity politics, in the late twentieth century has put the issue of memory centre-stage; for example, in the creation of myths of success or martyrdom of ethnic groups, of heroes in struggles (e.g. the suffragettes), and so on (Ashplant *et al.* 2000). Some thinkers have worried that the creation of such traditions is governed largely by relatively powerful interest groups (Hobsbawm & Ranger 1983); one hope of the M4L process is that such processes may be democratized by sufficiently powerful technology (O'Hara & Stevens 2006).

(ii) *Identity.* Following John Locke, a strong tradition in philosophy has focused on memory as a major constituent of identity (Warnock 1987); similarly, there is a strong link drawn in cognitive psychology between identity and autobiographical memory. Much of the M4L proposal (Fitzgibbon & Reiter 2003) looks at the cementing of the ability of individuals to construct narratives of their past out of diverse materials (and also, since these records have a persistence that other materials for memory do not, to avoid mythmaking). Such narratives, backed by technology, could have very strong effects in the development of group identities, which often have been seen as important in resistance of communities to political or economic pressure (Castells 1997).

(iii) *Trauma.* Trauma, via the process of forgetting explored in early psychoanalysis, has long been an important object of study in social science. On the scale of a society, organized forgetting

and positive distortion of the past have often been a response to a traumatic history (Paez *et al.* 1997), although many others have been sceptical about the extension of the psycholanalytic metaphor to a society as a whole (Margalit 2002, pp. 4–6). Trauma theories have a linear historical model of temporality which may help bring social memory and history closer together (Radstone 2000). Given that the materials that the M4L programme would be bringing to the study of memory are equally suited to the creation of history as the fashioning of memory, it is possible that M4L could continue that process of *rapprochement.* In particular, an interesting topic would be the discovery of social analogues of psychogenic versus organic amnesia, and using this sort of metaphor to investigate possibilities of 'treatment'.

(iv) *Conflict.* It perhaps goes without saying that memories of past evils, and their conceptualization, are the starting point for many of today's conflicts. To take only the most prominent present-day example, the grievances behind al-Qaeda's assault on Western liberalism are, to judge by their own statements, motivated less by modern-day conflicts in places such as Palestine, Chechnya and Iraq, controversial as they may be, and more by the defeat of the Moors by the Christians in Spain in 1492 (Roy 2004). Furthermore, the Spanish name for this event, the *reconquista,* or reconquest is an interesting example of a loaded choice of term. Actually, the Visigoths themselves invaded Spain under Theodoric II in 456AD, and suffered their first defeat by the Arabs under Roderic in 711AD (Collins 1983). The Arabs' 700 odd years in Spain were a much longer span than the Visigoths they defeated, and indeed more than the intervening period since the *reconquista,* even though the term 'reconquista' makes the Moorish period seem like a small inconvenient episode.

(v) *Justice.* The discussion of memory and justice is important mood music behind debates such as that over globalization, as first world and developing countries struggle to deal with the legacy of the colonial era. Germany has made serious attempts to address its totalitarian and genocidal past (Habermas 1997; McAdams 2001), as has South Africa (Krog 1998). There will need to be similar attempts in years to come, for example in the former Yugoslavia; there will certainly be plenty of materials (news footage, video diaries) for fashioning memories of the civil wars that simultaneously give voice to marginalized communities without demonizing former aggressors.

(vi) *Transhumanism.* It has long been noticed by philosophical and cultural critics that the increasing exploitation of technology by humans could well have important and far-reaching effects on human nature itself; indeed, concerns such as this date back at least as far as Socrates

(Havelock 1963). The nature of this effect has been hotly debated, optimists including the feminist Donna Haraway (1985) and the computer scientist Kevin Warwick (2002), pessimists including post-modernists, such as Lyotard (1988/1991), neuroscientist Susan Greenfield (2003) and political scientist Francis Fukuyama (2002).

In addition to social memory, the study of material culture and how individuals utilize physical artefacts as memory aids also has considerable relevance to those interested in developing digital memory and for the computing science challenges of M4L.

For centuries, individuals have used physical artefacts as external memory and reference aids. Over time, these have ranged from personal journals and correspondence, to photographs and photographic albums, to whole personal libraries of books, serials, clippings and off-prints. Annotation of these items and organization of personal workspace to house them can provide factual and emotive references, such as date, place, context and personal meaning or significance. Existing practice in annotation (Marshall 1998) or organization of workspace (Henderson 2004) are often not well supported in digital environments and there is much we can learn from their study in developing information management practices for digital memory and personal collections.

These personal collections and memory aids are often of intense importance to individuals, their descendents, and sometimes to a wider society. They have often been the foundation or a core component of museum, library and archive collections, and of the cultural record. Their study therefore has been an important component of traditional archive, library and information science. As we move from externally supported memory based on physical artefacts to a hybrid digital and physical environment, and then increasingly shift towards a digital memory for life, many new research issues arise around personal collections, personal libraries and information management (Beagrie 2005). These include how to physically secure such material sometimes over decades; how to protect privacy; how to organize and extract information and to use it effectively; and for material intended to be shared, how to effectively present and control access by different groups of users.

2.3. Digital memory

Computer memory is now so cheap that there appear to be few limits to the amount of information—and therefore the imperative to rationalize that information—that can be stored. The need to develop computer systems whose storage will encompass such quantities of time is increasing as societies continue to develop a prolific and promiscuous attitude towards information. For instance, proposals for identity cards will require, *inter alia*, databases that remain extant for the maximum human lifetime, something in excess of a century—in other words twice as long as the entire history of electronic computing. The issues with regard to, for example, representational formats (which generally change every few years) are major ones.

Many aspects of computer science will impinge upon the M4L challenge, including databases, artificial intelligence, human–computer interaction and operating systems. Security will be an issue of increasing importance (Schneider 1999). One area of growing importance, both with regard to computing issues and also knowledge management issues, is that of *knowledge technologies* (Shadbolt *et al.* 2004*a*,*b*). Knowledge technologies are systems intended to manipulate knowledge through its life cycle, from acquisition, through retrieval and publishing, to the essential maintenance of knowledge repositories. Hence, knowledge is information that can be put to use, for example, can be fed into a problem-solving process, or can be extracted from large stores of unstructured data.

Hence, knowledge technologies include: ontologies, shared conceptualizations of domains that facilitate communication between people, agents or systems about those domains (Gruber 1993; Fensel 2003); systems for enabling or automating the annotation of knowledge sources, such metadata allowing reasoning about knowledge sources to take place (Vargas-Vera *et al.* 2002; Handschuh *et al.* 2002); human language technologies, which enable structured knowledge to be extracted from unstructured text, or alternatively readable text to be created from machine- but not human-readable information resources (Cunningham *et al.* 2002; Ciravegna & Wilks 2003); and systems that enable information to be transferred around an organization, for instance capturing knowledge (and its context) generated in meetings, in order to distribute it to other stakeholders (Buckingham Shum *et al.* 2002; Eisenstadt & Dzbor 2002; Tate *et al.* 2002).

Such technologies often build on the extra intelligence capabilities generated by the Semantic Web, seen as a potential successor to the World Wide Web (Berners-Lee *et al.* 2001; Fensel *et al.* 2003). Given such a Web environment, individuals would have access to stores of knowledge with a wide range of Web services processing such content (Shadbolt *et al.* 2004*b*), accessed through intelligent brokers (Motta *et al.* 2003).

The WWW provides a number of opportunities for technologies relevant to the M4L area. For instance, the links on the current WWW are provided by the author of the page; this gives a type of associative linking from one page to the next, but the person in charge of the associations is the author, not the reader. Compare the famous scene in Proust's *Remembrance of Things Past*, in which the narrator, upon eating a coconut madeleine dipped in tea is immediately reminded of boyhood visits to his Aunt Léonie in Combray where he used to have the same treat. The associative links provided by the WWW in its current form are different; if we draw an analogy with Proust, the current WWW acts as if one eats a cake and then proceeds to have the memories of the *baker*! Hence, there is a great deal of research being carried out into the provision of links by the reader's own system, based on a user model of the reader (Hall 2000).

Associative memory generally will be of importance. For instance, in a neural net, the pattern of the unit's activation stores the data. Hence, a memory, which may have a number of attributes, is stored as a pattern of activations. The memory, unlike many symbolic systems, is not stored in an index system; it can, in fact, be retrieved by using any of the attributes of the memory as retrieval keys. Memory of this type is called *content addressable memory*, and has many advantages; not only flexibility of recall, but also the ability to work round corrupt data (e.g. Hodge & Austin 2001). An interesting open question here is how to store symbolic data in such a way as to support associative recall, in other words, so that it can be retrievable through its attributes (Gardenfors 2000; O'Hara *et al*. 2006).

A related issue is that of multimedia storage and retrieval. Search and retrieval of text is relatively straightforward, but without textual cues, other media present problems of framing queries, organization of memory stores, etc. Textual annotations may be added to video or graphics, say, but such annotations can never cover the full range of associations possible with a picture; added to which, having to annotate presents the user with a tedious overhead of effort.

There are many issues to be resolved here. For instance, there are different storage requirements (e.g. with respect to capacity) with dense media, such as video or audio. Searchable hierarchical structures provide one method of cutting the size of search spaces. But such structures must also support a number of other functions; for example, hypermedia systems need models able to support the complex real-time selection, retrieval and display of heterogeneous sources.

Over time, we can also expect the development of new media types, such as haptic representations (touch) or olfactory ones—computer systems are well on the road to a full range of sensory input capabilities. Integrating media types across different modalities, understanding how to index, retrieve and integrate information that originate in heterogeneous modalities, will be a central research challenge for the next 20 years. One possible method for this could be derived from multiresolution modelling, i.e. developing models that operate at many different levels of abstraction which can be reconstructed on demand (O'Hara *et al*. 2006).

Finally, as we saw in §2.1, human memory is actually made up of several different specialist systems. Integration between human and digital memory could be facilitated via an understanding of digital memory in these terms. For example, does having a limited capacity working memory impose useful constraints on the management of large amounts of permanent information? Does having separate but interlinked memory systems for verbal and visuo-spatial information suggest a useful way of organizing digitally coded language and images? Does consolidation provide a fruitful model for how irrelevant information gets removed (or ignored), and can the formation of semantic networks from episodic memories provide clues as to how 'irrelevant' is defined relative to context? How do we characterize and select the focus of attention (Oberauer 2002; O'Hara *et al*. 2006)?

3. TRANSMISSION AND LONGEVITY OF MEMORY

Digital memory over a human lifetime or beyond will require persistence and accessibility over many decades and transmission of that memory over many generations of hardware and software. This is by no means a trivial concern and involves organizational, legal and technical challenges.

It will not be possible to rely on the benign neglect or accidental survival, which has often been a significant factor in transmission of material memory. In the right conditions, papyrus or paper can survive by accident for centuries or in the case of the Dead Sea Scrolls for thousands of years. It takes hundreds of years for languages and handwriting to evolve to the point where only a few specialists can read them.

In contrast, digital information will never survive and remain accessible by accident: it requires ongoing active management (Beagrie & Jones 2001). The information and the ability to read it can be lost in a few years. Storage media such as paper tape, floppy disks, CD-ROM, DVD evolve and fall out of use rapidly. Digital storage media have relatively short archival lifespans compared to other media. Research on digital data loss has suggested that a substantial amount of personal data is not backed up and that, on average, while less than 2% of desktops are likely to experience an episode of data loss each year, the corresponding rate for laptops is greater than 10% because of the higher incidence of theft (Smith 2003). For any memory collection intended for access and use over a decade or more, the incremental accumulation of risk year by year will become unacceptable. Inability to access memory held in obsolete file formats will also be problematic. Its mitigation will need to become more inherent and automated in systems. One possible development to achieve this may be an increasing move towards information being held in online managed services, with personal devices acting solely as access points.

The experience of archives, libraries and museums (often referred to as 'memory institutions' because of their role in social and cultural memory) could enrich and interact with computing science research on M4L and the development of memory systems for individuals. In particular, research on digital preservation (the long-term preservation and accessibility of digital information) in these domains will be relevant to persistence and transmission of digital memory. In the mass consumer market, current interests in family history may drive interest in transmission of memories held in personal digital collections. In academic research, these personal histories will also be of interest, but it is suggested that the personal digital collections of leading creative writers and artists, politicians and scientists may first engage research libraries and archives in this area (Beagrie 2005).

Digital systems are currently poorly adapted to what might be called individuals' discontinuity of interest. There is a focus on the immediate needs of users and little in the way of digital equivalents of physical storage spaces, in which material can be laid down and later re-discovered, forgotten or discarded. Some

personal interests in collections change or may lie dormant over time. For example, in family history, one of the largest and most rapidly growing personal pastimes, use of personal collections and material may lie dormant for many decades. Individuals with no interest in historic material or potential future applications early in life are highly likely to be interested in them at a later stage of their lives.

4. TRUST, ETHICS, IDENTITY AND FORGETTING

The creation of large, long-lived stores of information, the development of techniques for efficient searching of them and the potential for wide access to these stores together raise the obvious problem of trust (Shadbolt 2003a). As Reiter noted in a magazine interview, we need to protect privacy when information about one person is featured in another's memories. How would we treat the claims, often legitimate, to that information from the police or security services (New Scientist 2003)?

Trust issues are already pressing where new technologies are enabling new spaces for interaction; for example, on the Internet, there is already an imperative to ensure that information is accurate and properly curated (O'Hara 2004, pp. 112–118). Work is beginning on the creation of trust mechanisms for Internet systems, and trust, it is fair to say, is a hot topic in Internet circles (O'Hara & Shadbolt 2005). In particular, relatively little is known about trust as a second-order property, whether it should be placed rationally, and what the trade-off is with security (O'Hara 2005).

A key difficulty is that of the blurring of identities that the Internet makes possible. In many ways, and for many users, this is the beauty of the Web; however, the requirements of security and such applications, as e-commerce, require relatively stable identities. This sort of trade-off is very hard to resolve (Lessig 1999; O'Hara 2004, pp. 99–112). The M4L context will place a very large number of constraints on this debate, not always consistently. Privacy will be essential for the promotion of trust; on the other hand, many of the uses for large memory repositories will be leisure-based, and may well require flexibility to allow large numbers of inexperienced users.

And to reiterate a debate in sociology and philosophy alluded to in §2.2, how will the externalization of memory, or the adding to our personal store a large repository of externalized digital memories, affect our identities as people, our human nature? Would such accessorizing be empowering? Or merely evidence of decline? On the level of society, will the heaping up of information about the past replace history and promote populist narratives (Lukacs 2005, pp. 192–199)? Or could it be an important aid to forgiveness without the moral compromise of forgetting shameful episodes in the past (Margalit 2002, pp. 183–209)? What will the relation be between the memory of the whole and that of the individual? Does the collective have the right to subpoena individuals' memories? And how will equal access, either as user or contributor, be secured (O'Hara & Stevens 2006)?

Potentially as important as preservation are the issues of retention, selection and disposal, and the role of forgetting in the transmission of digital memory. Dodge and Kitchin (2005) have highlighted some of the potential ethical issues of M4L in an era of pervasive electronic recording of all human activity. They have argued that forgetting should be an integral part of the process of designing and implementing M4L systems, as the information could be open to exploitation for commercial benefit or to abuse of civil liberties. They propose that digital memory could mirror some of the characteristics of forgetting in human memory listed by Schacter (2001) to overcome these issues by ensuring a sufficient degree of imperfection, loss and error.

However, does digital memory needs to be less effective in all circumstances or to mirror all the random imperfections of human memory? It can be argued that many of the concerns raised could be addressed by existing memory institution techniques for access control and retention schedules being applied to digital memory and their transmission to others. Professional practice in memory institutions employs a number of procedures for access control to safeguard privacy, confidential or sensitive information which can be applied to either material or digital memory. These can include redaction (blackouts), time-activated release and anomalization. Retention policies and procedures, i.e. what to keep, for how long and where, can also be used. Although it is technically possible to capture and retain in its entirety a continuous data stream for digital memory, there may be several reasons including privacy to apply an approach of selective retention. A formal transparent process for retention will contribute to trust in the record, which will be central to many potential applications of M4L and to its value as an extension of human memory.

5. MEMORIES FOR LIFE: A GRAND CHALLENGE

The scientific and technological threads described above are beginning to be drawn together. In particular, neuroscience and cognitive psychology are influencing each other's research agendas through cognitive neuroscience; biologically inspired computing is becoming increasingly mainstream; knowledge technologies are exploiting insights from both computer science and knowledge management; and current strands of social science research such as neuroeconomics are suggestive of important interdisciplinary arenas.

Right in the middle of this interdisciplinary mix is the grand challenge of M4L. As computers become increasingly able to store a lifetime's worth of memories, in various forms—digital photographs, emails, documents, accounts, blogs, video diaries—the question of managing such stores is becoming serious. Such management questions are typical of the corporate world, but are rapidly entering the private space too.

In particular, if we take a conception of such information stores as part of *memory* proper, outsourced to machines to be sure, but still adjuncts to the human act of recall (in the way that written texts became; Havelock 1963), then the challenge of M4L takes shape. The interdisciplinary opportunity becomes

the need to understand the operation of human memory, and its interaction with the environment, in order to augment it with technological support. The result of such a research programme, it is hoped, would be better human use of artificial memory storage and retrieval systems, and the smoother integration of such systems into real lives.

The statement of the M4L research challenge (Fitzgibbon & Reiter 2003) sets out many opportunities for bringing these various sciences and technologies to bear on human problems. The commercial implications of bringing together science and technology in a comprehensive way to deal with information through its life cycle in both individual and corporate contexts are likely to be dramatic. Furthermore, given that memory is, in general, a more serious problem for more marginalized communities (for instance, on the level of the individual, the elderly; on the social level, ethnic minorities), this is a real opportunity to apply technology sympathetically to such groups.

Fitzgibbon & Reiter mention a number of challenges that might fall under the M4L heading. Multimedia searching is an obvious challenge, as a hard and complex problem that will be central to most conceivable applications, as noted above. A possible move here would be to research the possibility of search of multimedia repositories using examples or exemplars.

A research problem bringing together a wide range of interests, is that of 'stories from a life'. Given a set of resources, how could one construct narratives around them, integrating the different modalities that may be present into a single seamless account—an account, moreover, that may differ depending on the person to whom it is presented? This is vital if we are to go beyond the storage and retrieval of vast amounts of *unanalysed* information. How do we understand the context of stories? How do we know when two different representations are of the same memory? How do we know the assumptions that a listener will make about what underlies a narrative? How do we know when a narrative is achieving its purpose (e.g. retelling some key event in a group's history)? Work such as this will be able to draw on applications for developing and displaying narratives in differing contexts from a variety of sources (Alani *et al.* 2003).

A third example from Fitzgibbon & Reiter is that of providing prosthetic memories for those with memory dysfunctions. Is it possible to analyse memories and create a schedule for a typical day? Will it be possible to monitor a person's activities and help them through their day, alerting carers when there is a significant deviation? How can we use our neuropsychological knowledge of memory dysfunction to specify the types of memory function to model (and can computing methods replicate such memory types accurately enough in real time)? Just how much memory could be successfully and seamlessly externalized? Or, in a more positive vein, will it be possible to develop models to help create lifelong 'companions' or personal agents to support, for example, one's personal e-learning profile, or to monitor and record one's health record?

The examples adduced here are relatively heterogeneous. Research in this area needs to be careful about

how tenuous are the metaphors that govern it (cf. O'Hara *et al.* 2006, pp. 240–241), and how they are stretched by usage. For example, forgetting in a human is an unconscious thing, some sort of second-order phenomenon the net result of which is an inability to recall some knowledge or memory; in a machine it is a first-order phenomenon of deletion. Does that mean that if we talk about forgetting in the cross-disciplinary M4L context we are in danger of incoherence?

In a comment on Fitzgibbon & Reiter (2003), Sparck Jones (2004) notes that there are a number of basic distinctions that need to be observed in order to place some coherence on the research area of M4L. One distinction is 'internal' versus 'external', in other words the commonsense distinction between something 'in' my mind and something public. Sparck Jones argues that currently there is little or no technological access to whatever is in my mind, and that the focus of M4L is necessarily to the external. On the other hand, the sort of anti-private language arguments made by the later Wittgenstein may alert us to the requirement to recast this distinction (Wittgenstein 1953). After all, on Wittgenstein's reading, there is nothing on the 'internal' side of this distinction anyway. A more important distinction is that between the preservation of memories for me, or for others, though even there methods for extracting and preserving what seem to be very personal memories sometimes have repercussions for outside observers (for instance, oral history tries to preserve very subjective experiences using idiosyncratic representations, but the aim is to try to recreate for history the experience of whatever is being recorded; and also in this context cf. Wilks 2004).

Based on an examination of these distinctions, Sparck Jones suggests five types of M4L project.

(i) *SuperMe.* A body of data that is an electronic enhancement of my memory, data that can be invoked to amplify events that I actually remember (e.g. photographs, Powerpoint slides, diary entries). The effect is somewhat prosthetic.
(ii) *Deposit.* A data repository of items that is meaningful to me.
(iii) *Persona.* A data declaration by me for consumption by others, a sort of public history of me developed by me that I make available to others.
(iv) *Assembly.* Somewhat like Persona, but not necessarily (entirely) under my control. An example of an Assembly might be a doctor's medical record of me.
(v) *Collective.* A body of data associated with different, though connected, people, such as a corporate memory, a society's archive or museum, or even the World Wide Web.

There are, no doubt, other interpretations of M4L logically available. For instance, one could imagine a body of data meaningful to a subject, yet not under control of the subject (a cross, then, between Deposit and Assembly); this might be a set of records and cues intended to help secure the coherence of daily life for someone with a failing memory or a chaotic lifestyle. As Sparck Jones points out, these different types of M4L

make many different assumptions about my understanding and manipulation of the data, of the privacy and trust issues, of representation requirements and so on, and the examples of Fitzgibbon & Reiter (2003) are distributed across these different types. On the other hand, part of the M4L research challenge is to understand methods of storage, retrieval and sensemaking that will be important in preserving these large and very possibly growing mountains of data, whose meaningfulness depends to a very large extent on the context (made up of the human memory with which they are associated, the rest of the data in the repository and the requirements of the repository's users), which we can expect to be highly heterogeneous in form, and which may need to be retrieved by many different users. Furthermore, we should not assume that the five types of M4L that Sparck Jones has set out will be mutually exclusive; it may be that the same data repository is used for multiple purposes, or, perhaps more likely, as a component in a number of different systems alongside other repositories. The issue, as Sparck Jones suggests, is whether generic data storage and retrieval techniques exist that could be used in a series of applications such as those suggested above, in such a way that they go beyond technological fixes, but genuinely import insights from medical, biological and social sciences. Focusing now on the 'for life' part of M4L, such generic techniques should be capable of supporting repositories or memories for far longer periods of time than we currently typically envisage. After all, at the time of writing the entire history of computing is somewhat less than a normal human lifetime.

Technically, the challenges are enormous. For example, open data structures would allow assembled data to outlast the systems that generated and stored them, which could be a boon in this context. Indexing strategies must be flexible enough to allow new sets of unpredicted questions to be asked. How do we develop interfaces that maximize access for the non-computer literate?

One immediate issue of great importance is the development of a set of working data that can help define a problem space, act as a testbed for particular approaches, while also facilitating comparisons (Fitzgibbon & Reiter 2003). Such data should not have been processed significantly, of course, for this would import too many assumptions about formats and purposes to allow for the pursuit of genericity, or to support genuine research about technological possibilities far into the future. The media of the data may be very heterogeneous, but some of it needs to be pretty raw.

Such are the complexities of the technological issues here—not merely in technological development, though those are hard enough, but also with respect to situating technology within recognizably human and social contexts—that M4L has been selected as one of the United Kingdom Computing Research Committee's (UKCRC) Grand Challenges. The UKCRC is a joint committee of the British Computing Society and the Institute of Electrical Engineers, and has sponsored a series of challenges the pursuit of which will enhance and focus future computing research (http://www.ukcrc.org.uk/grand_challenges/index.cfm). The

criteria for selection as a grand challenge are that (i) the challenge has international scope, (ii) its ambition exceeds that of a single research group or single grant, (iii) it should promise revolutionary rather than evolutionary advance and (iv) there should be a consensus within the scientific community that the research goal will be interesting and fruitful (cf. Fitzgibbon & Reiter 2003). Furthermore, the British Engineering and Physical Sciences Research Council has sponsored a network called M4L in order to take the short-term steps to begin to engineer a multi-disciplinary community of scientists willing to work together to attempt to implement the M4L challenge (http://www.memoriesforlife.org/). The M4L network has set up a working group to investigate the possibilities for securing testbed data.

Forecasts about future technological developments are fraught with pitfalls, needless to say, and Fitzgibbon & Reiter's account is meant to be no more than suggestive of opportunity. But the examples given, and the general understanding of the significance of the blurring of the physical and the digital worlds, make it clear that there is a problem space emerging where all the disciplines discussed in this paper, and doubtless many more, will be able to contribute.

It is, of course, a moot question as to how much genuine interdisciplinary collaboration could emerge in pursuing a vision of the augmentation of human memory using open-ended technologies. We have already reviewed the state of the art in various disciplines' study of memory (see also Morris *et al.* 2006*b*), and seen how close a genuine problem space is.

For instance, different types of human memory, and their neurological implementations, are suggestive to computer science, not only for the purposes of imitation, but also to find artificial memory functions that complement human function. Sociological research is beginning to tell us how memory fits into social behaviour, how social memories are constructed and which memories it is therefore important to preserve. In more formal social structures, knowledge management and organizational science are telling us more or less the same things, while providing an important influence on technological development. It is impossible to predict which developments will be the most far-reaching, but we might venture to hope that many developments will improve the lot of both individuals and communities.

6. CONCLUSION

In a celebrated essay on the new electronic media, Marshall McLuhan wrote in 1962:

> Our private senses are not closed systems but are endlessly translated into each other in that experience which we call consciousness. Our extended senses, tools, technologies, through the ages, have been closed systems incapable of interplay or collective awareness. Now, in the electric age, the very instantaneous nature of co-existence among our technological instruments has created a crisis quite new in human history. Our extended faculties and senses now constitute a single field of experience which demands that they become

collectively conscious. Our technologies, like our private senses, now demand an interplay and ratio that makes *rational* co-existence possible. As long as our technologies were as slow as the wheel or the alphabet or money, the fact that they were separate, closed systems was socially and psychically supportable. This is not true now when sight and sound and movement are simultaneous and global in extent.

(McLuhan 1962, p. 5, emphasis in original)

Over 40 years later, the seamless interplay that McLuhan demanded between our technologies is still barely visible. McLuhan's predictions of the spread, and increased importance, of electronic media have of course been borne out, but the open systems that he called for remain on the drawing board. Part of the problem is that technology has all too often been inserted in social situations without a thought about how people could interact with it, as many commentators, not only those critical of science and technology, have pointed out.

The M4L challenge has the potential to bring together the fascinating results from various disciplines that we described above. At a minimum, it promises to provide a common problem space to promote the interdisciplinary integration that would in itself be a benefit. At best, the understanding of the social context, the possibilities and limits of human capacities and the sensitive application of powerful technologies could revolutionize both our understanding of memory and our effectiveness as agents.

The role of learned societies, such as the Royal Society, or the British Computer Society, in assessing these questions, and in communicating scientific and technological advance to stakeholders, is clearly key here. The M4L programme will need scientific infrastructural support, as interdisciplinary collaboration can be hard, both in terms of achieving commonality of view between collaborators, and with respect to funding. Naturally, funding bodies and university departments are organized along disciplinary lines. Near-market products can tap into resources from the private sector, but *ex hypothesi* M4L is not at that stage yet.

This paper, whose authors are drawn from the fields of computer science, cognitive psychology, neuroscience and information science, has developed over a long period as a result of a series of reports written for the UK Government's Cognitive Systems Foresight programme (Morris *et al.* 2006*b*), which enabled the authors to identify potential synergies and areas where research from different disciplines would dovetail. Such juxtaposition of scientists is important for spotting opportunities; it is to be hoped that a challenge such as M4L could provide a focus for funding bodies, scientists and industry to work together effectively in pursuit of the science of memory and the technology of information storage and retrieval.

REFERENCES

Abowd, G. D., Mynatt, E. D. & Rodden, T. 2002 The human experience. *IEEE Pervas. Comput.* **1**, 48–57. (doi:10.1109/MPRV.2002.993144)

Ahmed, B., Driver, J., Friston, K., Matus, A., Morris, R. & Rolls, E. 2006 Advanced neuroscience technologies. In *Cognitive systems: information processing meets brain science* (ed. R. Morris, L. Tarassenko & M. Kenward), pp. 263–282. Amsterdam: Elsevier Press.

AKT 2001 *The AKT manifesto*. The Advanced Knowledge Technologies project, http://www.aktors.org/publications/.

Alani, H., Kim, S., Millard, D. E., Weal, M. J., Hall, W., Lewis, P. H. & Shadbolt, N. R. 2003 Automatic ontology-based knowledge extraction from web documents. *IEEE Intell. Syst.* **18**, 14–21. (doi:10.1109/MIS.2003.1179189)

Anderson, J. R. 1993 *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

Ashplant, T. G., Dawson, G. & Roper, M. (eds) 2000 *The politics of war memory and commemoration*. London: Routledge.

Baddeley, A. D. 1992 Working memory: the interface between memory and cognition. *J. Cogn. Neurosci.* **4**, 281–288.

Bayley, P. J. & Squire, L. R. 2003 The medial temporal lobe and declarative memory. In *Cognition and emotion in the brain* (ed. T. Ono, G. Matsumoto, R. R. Llinas, A. Berthoz, R. Norgren, H. Nishijo & R. Tamura), pp. 245–259. Amsterdam: Elsevier Press.

Beagrie, N. 2005 Plenty of room at the bottom? Personal digital libraries and collections. *D-Lib Mag.* **11**. (doi:10.1045/june2005-beagrie)

Beagrie, N. & Jones, M. 2001 *Preservation management of digital materials: a handbook*. London: British Library.

Berners-Lee, T., Hendler, J. & Lassila, O. 2001 The semantic web. *Scientific American*, May 2001.

Blakemore, S.-J. & Decety, J. 2001 From the perception of action to the understanding of intention. *Nat. Rev. Neurosci.* **2**, 561–567.

Bourdieu, P. 1977 *Outline of a theory of practice* (transl. by R. Nice). Cambridge, UK: Cambridge University Press.

Broadbent, D. E. 1958 *Perception and communication*. London: Pergamon Press.

Brooking, A. 1998 *Corporate memory: strategies for knowledge management*. London: Thomson Business Press.

Buckingham Shum, S., De Roure, D., Eisenstadt, M., Shadbolt, N. & Tate, A. 2002 CoAKTinG: collaborative advanced knowledge technologies in the grid. In *Proc. Second Workshop on Advanced Collaborative Environments, 11th IEEE Int. Symp. on High Performance Distributed Computing (HPDC-11), Edinburgh, Scotland, 24 July 2002.*

Castells, M. 1997 *The power of identity*. Oxford, UK: Polity Press.

Ciravegna, F. & Wilks, Y. 2003 Designing adaptive information extraction for the Semantic Web in Amilcare. In *Annotation for the Semantic Web* (ed. S. Handschuh & S. Staab). Amsterdam: IOS Press.

Collins, R. 1983 *Early medieval Spain: unity in diversity, 400–1000*. Basingstoke, UK: Macmillan Education.

Cowan, R., David, P. A. & Foray, D. 1999 The explicit economics of knowledge codification and tacitness.

Presented at the Third TIPIK Workshop, BETA, University of Louis Pasteur, Strasbourg, 24 April 1999, http://www-econ.stanford.edu/faculty/workp/swp99027.pdf.

Cunningham, H., Maynard, D., Bontcheva, K. & Tablan, V. 2002 GATE: a framework and graphical development environment for robust NLP tools and applications. In *Proc. 40th Anniversary Meeting of the Association for Computational Linguistics (ACL02), Philadelphia, July 2002,* http://gate.ac.uk/sale/acl02/acl-main.pdf.

Dix, A. 2002 The ultimate interface and the sums of life? *Interfaces* **50**, 16.

Dodge, M. & Kitchin, R. 2005 *The ethics of forgetting in an age of pervasive computing.* Centre for Advanced Spatial Analysis (CASA) Working paper 92, London: University College London, http://www.casa.ucl.ac.uk/working_papers/paper92.pdf.

Douglas, M. 1987 *How institutions think.* London: Routledge & Kegan Paul.

Eisenstadt, M. & Dzbor, M. 2002 Buddyspace: enhanced presence management for collaborative learning, working, gaming and beyond. In *Proc. JabberConf Europe, Munich, Germany.*

Elliott, T. & Shadbolt, N. R. 2003 Developmental robotics: manifesto and application. *Phil. Trans. R. Soc. A* **361**, 2187–2206. (doi:10.1098/rsta.2003.1250)

Fensel, D. 2003 *Ontologies: silver bullet for knowledge management and electronic commerce,* 2nd edn. Berlin: Springer.

Fensel, D., Hendler, J., Lieberman, H. & Wahlster, W. (eds) 2003 *Spinning the semantic web: bringing the World Wide Web to its full potential.* Boston, MA: MIT Press.

Fitzgibbon, A. & Reiter, E. 2003 *Memories for life: managing information over a human lifetime.* Cambridge, UK: Computer Laboratory, University of Cambridge. http://www.nesc.ac.uk/esi/events/Grand_Challenges/proposals/Memories.pdf.

Fletcher, P. C. & Henson, R. N. A. 2001 Frontal lobes and human memory: insights from functional neuroimaging. *Brain* **124**, 849–881. (doi:10.1093/brain/124.5.849)

Fukuyama, F. 2002 *Our posthuman future: consequences of the biotechnology revolution.* London: Profile Books.

Gardenfors, P. 2000 *Conceptual spaces: the geometry of thought.* Cambridge, MA: MIT Press.

Greenfield, S. 2003 *Tomorrow's people: how 21st-century technology is changing the way we think and feel.* London: Allen Lane.

Gruber, T. R. 1993 A translation approach to portable ontologies. *Knowl. Acquis.* **5**, 199–220. (doi:10.1006/knac.1993.1008)

Habermas, J. 1997 *A Berlin Republic: writing on Germany* (transl. by S. Rendell). Lincoln, NE: University of Nebraska Press.

Hall, W. 2000 The button strikes back. *New Rev. Hypermedia Multimedia* **6**, 5–17.

Handschuh, S., Staab, S. & Ciravegna, F. 2002 S-CREAM: semi-automatic CREAtion of Metadata. In *Knowledge engineering and knowledge management: ontologies and the semantic web* (ed. A. Gómez-Pérez & V. R. Benjamins), pp. 358–372. Berlin: Springer.

Haraway, D. 1985 A manifesto for cyborgs: science, technology and socialist feminism in the 1980s. *Socialist Rev.* **80**, 65–108.

Havelock, E. A. 1963 *Preface to Plato.* Oxford, UK: Basil Blackwell.

Henderson, S. 2004 How do people organize their desktops? In *Conf. on Human Factors in Computing Systems CHI 2004 Extended Abstracts on Human Factors in Computing Systems,* pp. 1047–1048. New York, NY: ACM Press.

Hobsbawm, E. & Ranger, T. (eds) 1983 *The invention of tradition.* New York, NY: Cambridge University Press.

Hodge, V. J. & Austin, J. 2001 An integrated neural IR system. In *Proc. Ninth European Symp. on Artificial Neural Networks, Bruges, Belgium, 25–27 April 2001,* pp. 265–270.

Hunt, A. S. & Edgar, C. C. 1932 *Select papyri. Private documents,* vol. 1. Cambridge, MA: Loeb Classical Library.

Huyssen, A. 1995 *Twilight memories.* London: Routledge.

Krog, A. 1998 *Country of my skull.* Parktown, SA: Random House.

LeDoux, J. 2002 *Synaptic self: how our brains become who we are.* New York, UK: Viking Penguin.

Lessig, L. 1999 *Code and other laws of cyberspace.* New York, NY: Basic Books.

Lukacs, J. 2005 *Democracy and populism: fear and hatred.* New Haven, CT: Yale University Press.

Lyotard, J.-F. 1988/1991 *The Inhuman.* Stanford, CA: Stanford University Press.

Margalit, A. 2002 *The ethics of memory.* Cambridge, MA: Harvard University Press.

Marshall, C. 1998 Towards an ecology of hypertext annotation. In *Proc. ACM Hypertext '98,* pp. 40–49. New York, NY: ACM Press.

Martin, S. J. & Morris, R. G. M. 2002 New life in an old idea: the synaptic plasticity and memory hypothesis revisited. *Hippocampus* **12**, 609–636. (doi:10.1002/hipo.10107)

McAdams, A. J. 2001 *Judging the past in unified Germany.* Cambridge, UK: Cambridge University Press.

McLuhan, M. 1962 *The Gutenberg Galaxy.* London: Routledge & Kegan Paul.

Misztal, B. A. 2003 *Theories of social remembering.* Maidenhead, UK: Open University Press.

Morris, R., Hitch, G., Graham, K. & Bussey, T. 2006*a* Learning and memory. In *Cognitive systems: information processing meets brain science* (ed. R. Morris, L. Tarassenko & M. Kenward), pp. 193–235. Amsterdam: Elsevier Press.

Morris, R., Tarassenko, L. & Kenward, M. (eds) 2006*b* *Cognitive systems: information processing meets brain science.* Amsterdam: Elsevier Press.

Motta, E., Domingue, J., Cabral, L. & Gaspari, M. 2003 IRS-II: a framework and infrastructure for semantic web services. In *Proc. Second Int. Semantic Web Conf. (ISWC'03), Sundial Resort, Sanibel Island, FL, USA, 20th–23rd October 2003.*

New Scientist 2003 The story of your life … on a laptop. *New Scient.,* 4 October 2003, p. 28.

Nonaka, I. & Takeuchi, H. 1995 *The knowledge-creating company: how Japanese companies create the dynamics of innovation.* New York, NY: Oxford University Press.

Oberauer, K. 2002 Access to information in working memory: exploring the focus of attention. *J. Exp. Psychol.: Learn. Mem. Cogn.* **28**, 411–421. (doi:10.1037/0278-7393.28.3.411)

O'Hara, K. 2004 *Trust: from Socrates to spin.* Cambridge, UK: Icon Books.

O'Hara, K. 2005 The ethics of cyber trust. In *Trust and crime in information societies* (ed. R. Mansell & B. S. Collins), pp. 442–449. Cheltenham, UK: Edward Elgar.

O'Hara, K. & Shadbolt, N. 1997 Interpreting generic structures: expert systems, expertise and context. In *Expertise in context* (ed. P. J. Feltovich, K. M. Ford & R. R. Hoffman), pp. 449–472. Menlo Park, CA: AAAI Press/MIT Press.

O'Hara, K. & Shadbolt, N. 2005 Knowledge technologies and the semantic web. In *Trust and crime in information societies* (ed. R. Mansell & B. S. Collins), pp. 113–164. Cheltenham, UK: Edward Elgar.

O'Hara, K. & Stevens, D. 2006 *inequality.com: power, poverty and the digital divide.* Oxford, UK: Oneworld.

O'Hara, K., Hall, W., van Rijsbergen, K. & Shadbolt, N. 2006 Memory, reasoning and learning. In *Cognitive systems: information processing meets brain science* (ed. R. Morris, L. Tarassenko & M. Kenward), pp. 236–260. Amsterdam: Elsevier Press.

Ong, W. J. 1982 *Orality and literacy: the technologizing of the word.* London: Methuen.

Paez, D., Basabe, N. & Gonzales, J. L. 1997 A cross-cultural approach to remembering. In *Collective memory of political events: social psychological perspectives* (ed. J. W. Pennebaker, D. Paez & B. Rimé), pp. 147–174. Mahwah, NJ: Lawrence Erlbaum.

Perrett, D. I., Hietanen, J. K., Oram, M. W. & Benson, P. J. 1992 Organisation and functions of cells responsive to faces in the temporal cortex. *Phil. Trans. R. Soc. B* **335**, 23–30.

Perrett, D. I., Lee, K., Penton-Voak, I., Burt, D. M., Rowland, D., Yoshikawa, S., Henzi, S. P., Castles, D. & Akamatsu, S. 1998 Sexual dimorphism and facial attractiveness. *Nature* **394**, 884–886. (doi:10.1038/29772)

Posner, M. I. & DiGirolamo, G. J. 2000 Cognitive neuroscience: origins and promise. *Psychol. Bull.* **126**, 873–889. (doi:10.1037/0033-2909.126.6.873)

Radstone, S. 2000 Working with memory: an introduction. In *Memory and methodology* (ed. S. Radstone), pp. 1–24. Oxford, UK: Berg.

Rizzolatti, G., Fogassi, L. & Gallese, V. 2001 Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* **2**, 661–670. (doi:10.1038/35090060)

Roy, O. 2004 *Globalized Islam: the search for a new Ummah.* New York, NY: Columbia University Press.

Ryle, G. 1949 *The concept of mind.* Harmondsworth, UK: Penguin.

Schacter, D. L. 2001 *The seven sins of memory: how the mind forgets and remembers.* New York, NY: Houghton Mifflin.

Schneider, F. B. (ed.) 1999 *Trust in cyberspace.* Washington, DC: National Academy Press.

Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N., Van de Velde, W. & Wielinga, B. 2000 *Knowledge engineering and management: the Common-KADS methodology.* Cambridge, MA: MIT Press.

Shadbolt, N. 2003a Someone to watch over you. *IEEE Intell. Syst.* **18**, 2–3. (doi:10.1109/MIS.2003.1200718)

Shadbolt, N. 2003b In memoriam. *IEEE Intell. Syst.* **18**, 2–3. (doi:10.1109/MIS.2003.1200718)

Shadbolt, N. 2004a Nature-inspired computing. *IEEE Intell. Syst.* **19**, 2–3.

Shadbolt, N. 2004b Moving forward: advanced knowledge technologies shaping a fast-developing world. In *Advanced knowledge technologies selected papers 2004* (ed. N. Shadbolt & K. O'Hara), pp. vii–xix. Southampton, UK: University of Southampton.

Shadbolt, N. *et al.* 2004a Advanced knowledge technologies at the midterm: tools and methods for the semantic web. In *Advanced knowledge technologies selected papers 2004* (ed. N. Shadbolt & K. O'Hara), pp. 1–60. Southampton, UK: University of Southampton.

Shadbolt, N., Gibbins, N., Glaser, H., Harris, S. & Schraefel, M. M. C. 2004b CS AKTive Space: or how we stopped worrying and learned to love the semantic web. *IEEE Intell. Syst.* **19**, 41–47. (doi:10.1109/MIS.2004.8)

Smith, D. M. 2003 The cost of lost data. *Graziadio Business Report: Journal of Contemporary Business Practice*, vol. 6. Los Angeles, CA: Pepperdine University. http://gbr.pepperdine.edu/033/dataloss.html.

Sparck Jones, K. 2004 *Comments on grand challenge document (GCD): memories for life: managing information over a human lifetime.* Cambridge, UK: Computer Laboratory, University of Cambridge. http://www.nesc.ac.uk/esi/events/Grand_Challenges/gcconf04/submissions/28.pdf.

Steels, L. & Brooks, R. A. (eds) 1995 *The artificial life route to artificial intelligence: building embodied situated agents.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Tate, A., Dalton, J. & Stader, J. 2002 $I-P^2$: intelligent process panels to support coalition operations. In *Proc. Second Int. Conf. on Knowledge Systems for Coalition Operations (KSCO-2002), Toulouse, France, 24th April 2002.*

Tulving, E. 2002 Episodic memory: from mind to brain. *Annu. Rev. Psychol.* **53**, 1–25. (doi:10.1146/annurev.psych.53.100901.135114)

Vargas-Vera, M., Motta, E., Domingue, J., Lanzoni, M., Stutt, A. & Ciravegna, F. 2002 MnM: ontology driven semi-automatic and automatic support for semantic markup. In *Knowledge engineering and knowledge management: ontologies and the semantic web* (ed. A. Gómez-Pérez & V. R. Benjamins), pp. 379–391. Berlin: Springer.

Warnock, M. 1987 *Memory.* London: Faber & Faber.

Warwick, K. 2002 *I, Cyborg.* London: Century.

Wilks, Y. 2004 *Companions: a response to a grand challenge (GC3) with a revival of something traditional.* http://www.nesc.ac.uk/esi/events/Grand_Challenges/gcconf04/submissions/33.pdf.

Wittgenstein, L. 1953 *Philosophical investigations.* Oxford, UK: Basil Blackwell.