Manifesto

The mission of the Advanced Knowledge Technologies consortium is to identify where we can, and invent where we must, the next generation of technologies for organisations to create, manage and extract value from their knowledge assets, and to integrate these technologies to show a complete approach to the knowledge life cycle.

AKT is an ambitious project with the remit and resources to think 'out of the box'. The design, implementation, integration and deployment of state of the art knowledge technologies in authentic work contexts is the distinguishing mark of AKT.

AKT: An EPSRC Interdisciplinary Research Collaboration

The UK's Engineering and Physical Sciences Research Council (EPSRC) has established a series of Interdisciplinary Research Collaborations (IRCs) to help identify and solve the IT problems of the future. They are intended to bring together researchers from differing but complementary backgrounds to address such problems. Several IRCs have been funded to begin work in October 2000 (one partly funded by the Medical Research Council), and will run for six years. The EPSRC has attempted to stimulate the creativity of these high-calibre consortia by encouraging the most ambitious projects. Links with industry will ensure the relevance of the research, but imaginative solutions will also be the order of the day.

The EPSRC's IRCs represent an opportunity to address IT problems in an integrated fashion. The AKT consortium consists of academics from five premier British universities: Southampton, Edinburgh, Aberdeen, Sheffield and The Open University. The director is Professor Nigel Shadbolt, of the Department of Electronics and Computer Science in the University of Southampton. Seven other principal investigators ensure an impressive range of experience and talent.

While focusing on computer science, the consortium has in its possession a wide range of expertise, in disciplines as diverse as: cognitive science, multimedia, knowledge management, intra/internet technology, computational linguistics and natural language understanding, machine learning, databases, artificial intelligence, philosophy and logic, knowledge engineering, software agents, human-computer interaction, computer-supported collaborative work and hypertext. AKT weaves these into an integrated, tightly specified framework, exploiting the synergy between them to facilitate knowledge technology solutions. With a £7.5 million budget, AKT has a mandate to think radically about the future, and shape it.

As well as the academic expertise, AKT will benefit from a close relationship with an industrial community increasingly aware of the difficulties of knowledge management. A steering committee of major industrialists from all sectors of the economy, from blue chips to cutting-edge technology firms, will keep the AKTors abreast of the perception of knowledge problems 'on the ground'.

Visit **www.aktors.org** to learn more.

1. Overview

Knowledge Management (KM) is not only about Knowledge Technology. The organisational debris from failed attempts to impose new technical infrastructures that are either inappropriate to their work environments, or where people are not willing to share knowledge is ample evidence. However, even if we accept the slogan, “80% people, 20% technology”, large scale KM initiatives will invariably depend on some form of technology to assist communication and storage; moreover, certain forms of KM depend on new, emerging technologies, and technology can (as in many fields) also be a driver for innovation. AKT brings together knowledge technology specialists from both the computing and human sciences, with an acute recognition of the interplay between people, organisations and technology in effective KM.

In this White Paper, we will discuss six challenges of knowledge management, and show how any integrated KM approach will need to meet these challenges. The AKT research programme is built around these challenges, and the technologies that result will be essential for anyone who needs to deal with knowledge and to maximise its value. The research will be carried out with a strong focus on a number of real-world testbed scenarios, ensuring applicability. There is no presumption that an organisation will need to reconfigure itself completely just to run advanced knowledge systems; we aim to produce advanced and illustrative prototypes of software and methods that will slot easily into a business environment.

This is an ambitious programme, but with the AKT consortium’s resources and impressive track record in collaborative research, we are confident that we will achieve our goals.

2. From Information Overload to Actionable Knowledge

In our wired world, data and information (which might be defined as ‘raw’ data that has been structured and represented for the human senses) have never been so easy to access and store. Globally, it’s estimated that 1-2 exabytes of data is now being generated each year, almost all of it in purely digital form.[1] Not surprisingly, this info-bonanza is experienced all too frequently not as a boon, but as a burden. We are not just swimming but drowning in data. Managers, researchers or workers have to grope their way through the ‘infosmog’ before reaching their decisions, confident of one thing alone: they have not taken everything properly into account.

We find ourselves in this situation because whilst accessing and storing data and information is technically cheap and relatively straightforward, *interpreting* it requires intellectual investment: *attention, time, expertise and experience* are now the commodities in highest demand and shortest supply. Veterans and other experts are so valued because they *make sense* of information, seeing patterns, implications and connotations that others miss. Knowledge, like beauty, is thus in the eye of the beholder: ‘one man’s data is another man’s knowledge’.

So, in contrast to information, let’s be pragmatic and define *knowledge* as *information sufficiently interpreted to enable action.* Everything is information until it is interpreted and enables some form of action (even if only to decide that it is irrelevant). Even if accessible one click away, until it is interpreted – by a human or a machine – it is meaningless, and thus useless. When information is interpreted, then it can be matched with, and brought to bear upon, the particular problems your business or organisation is addressing. To turn your information into actionable knowledge, it is necessary to understand the connections between it and your business processes. This brings us to the kinds of pressing questions now facing

1 *How Much Information?*, University California, Berkeley [www.sims.berkeley.edu/research/projects/how-much-info/]

organisations: what information to use when, how to find it, who may know, who may already have rated its significance, how to present it to the right people in the right form at the right time, etc.

This is, of course, easy to say and hard to do. Most information is not currently stored in ways that assist interpretation, whether by people or machines. People can always interpret information given enough time and skill, but these are scarce, valuable resources. Put bluntly, *there's plenty of information technology in use, but barely any knowledge technology*.

If knowledge assets are to be managed effectively, then Advanced Knowledge Technologies that render information interpretable in order to enable effective action must be a part of the solution. AKT is very much about 'the economics of attention.' Our goal is to release people from coping with the information deluge, in order to focus on the really pressing, hard, interesting problems. AKT consortium technologies already enable machines to do some of the basic interpretive work: filtering, classifying and recommending information. We can already intelligently present hyperlinks to assist browsing, and can link people to relevant experts. We have a suite of tools for enriching content with conceptual information, automatically, semi-automatically, and where a human is required, manually.

3. Challenges for Managing Knowledge

Dealing with knowledge is easier said than done (witness the dreams, and failure, to codify expertise as rules in any but extremely well understood domains). Most often it lies within an organisation implicitly, out of sight, undervalued and underused. Often, it leaves the building each night at 5pm. When people move from department to department, or to a competitor, they take their knowledge with them. Expertise that has been built up over years can be lost overnight.

Managing the flow of knowledge around an organisation is a challenge. More precisely, it is *six* challenges. Failing to meet any one of these challenges can derail an organisation's ability to use its knowledge assets to its best advantage.

We see the six generic challenges for knowledge management as follows:

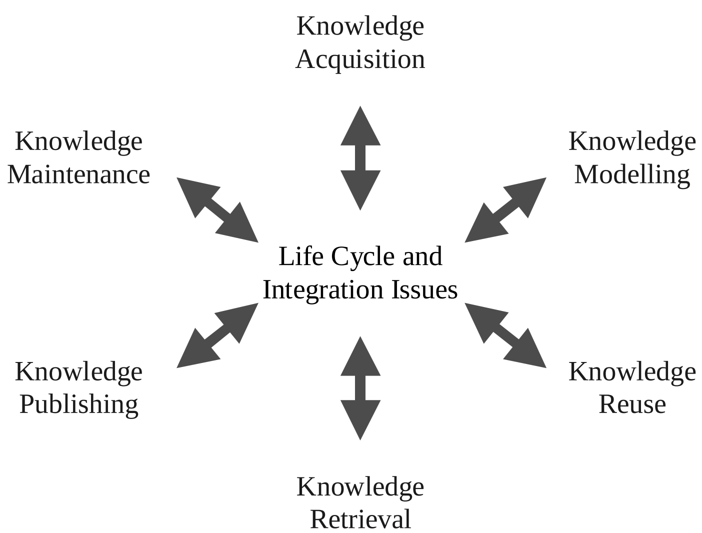
- **Acquiring it**: Although we often suffer from an overdose of data – infosmog – all too often the problem is that the knowledge available is insufficient or poorly-specified. The challenge here is to get hold of the information that is around, and turn it into knowledge by making it *usable*. This might involve, for instance, making tacit knowledge explicit, identifying gaps in the knowledge already held, acquiring and integrating knowledge from multiple sources (e.g. different experts, or distributed sources on the WWW), acquiring knowledge from unstructured media (e.g. natural language or diagrams).

- **Modelling it**: Modelling bridges the gap between the acquisition of knowledge and its use. Knowledge model structures must be able to represent knowledge so that it can be used for problem-solving. One important knowledge modelling idea is that of *ontologies*, which are specifications of the generic concepts, attributes, relations and axioms of a knowledge base or domain. Ontologies can act as placeholders and organising structures for acquired knowledge, while also providing a format for understanding how knowledge will be used.

- **Reusing it**: One of the most serious impediments to cost-effective use of knowledge is that often knowledge bases or systems are constructed afresh. It is unusual for problem-solving experience or domain content to be acquired and then reused, partly because knowledge tends to require different representations depending on the problem-solving that it is intended to do. Understanding the use and application of knowledge would enable more leverage to be gained from the knowledge already at hand, thereby increasing the returns on the investment in those knowledge assets.

- **Retrieving it**: When a knowledge repository gets very large, finding a particular piece of knowledge can become very difficult. There are two related problems to do with knowledge retrieval. First, there is the issue of finding knowledge again once it has been stored, understanding the structure of your archive in order to navigate through it efficiently. And second, there is the problem of retrieving the subset of content from the repository that is relevant to a particular problem. This second problem, the dynamic extraction of knowledge from a repository, may well set problems for a knowledge retrieval system that alter regularly and quickly during problem-solving.
- **Publishing it**: The challenge of publishing or disseminating can be described as getting the right knowledge, in the right form, in the right place, to the right person, at the right time. Different users will require knowledge presented and visualised in different ways, and the quality of such presentation is not merely a matter of preference, but can radically affect the value of the knowledge to the user. Getting presentation right will involve understanding the different perspectives of people with different agendas, while an understanding of knowledge content will help to ensure that important related pieces of knowledge get published at the appropriate time.
- **Maintaining it**: The last challenge is to keep the knowledge repository functional. This may involve the regular updating of content as content changes (e.g. as price lists are revised). But it may also involve a deeper analysis of the knowledge content. Some content has a considerable longevity, while other knowledge dates very quickly. If a repository of knowledge it to remain active over a period of time, it is essential to know which parts of the knowledge base must be discarded and when. Other problems involved in maintenance include verifying and validating the content, and certifying its safety.

Each of these six challenges requires a multidisciplinary strategy. They also implicitly define a seventh challenge: to produce an *integrated response* to the six challenges. “Integrated” in the sense of embedded into an organization’s existing infrastructure, knowledge flow and working practices. And “integrated” in the sense that we need a seamless way to manage knowledge throughout its lifecycle, where a system would facilitate, for example, both the acquisition of knowledge and its retrieval. The links between the challenges are related to issues such as, for instance, a common infrastructure, and a life cycle model that treats the various challenges as elements of a global process.



4. AKT: Advanced Knowledge Technologies

This is the background to the creation of the project AKT: Advanced Knowledge Technologies. AKT is an interdisciplinary research collaboration, funded by Britain's EPSRC, to address the six knowledge challenges in an integrated fashion, developing technology and an understanding of the dynamics of knowledge within an organisation or firm to facilitate the management of knowledge and the extraction of value from it.

The six challenges described above – which every knowledge manager will recognise – are the organising principle behind AKT. The major technology development workpackages within the project are each focused on one of these challenges. Of course, an integrated approach is essential, and other workpackages are devoted to ensuring integration between the approaches to each challenges, and to ensuring underlying, lightweight, flexible infrastructure support.

To this end, the second focus of the AKT project is on a number of testbed scenarios. These are real-world problems, which exist independently of the project, to which we will apply our ideas, ensuring relevance, practicality and applicability. The challenges will drive our theory; the testbeds will drive the applications of that theory.

It can be seen from the discussions of the knowledge context that technology is already part of the problem. The Internet is a technological creation, and it is technology that will be required to get the most out of the new developments of the semantic web. And information overload is caused chiefly by computational storage methods multiplying the amount of knowledge that can be stored enormously while diminishing the costs of the storage. Whether these are seen as technological problems or technological opportunities, it is clear that technology is an essential part of maximising the use of an organisation's knowledge in the knowledge context we have described.

AKT's Roots

The original impetus for AKT came from two workshops held under the aegis of the EPSRC on Software Assisted Knowledge Acquisition (SAKA). SAKA I was held at Abingdon in November 1996, under the coordination of Professor Derek Sleeman. It brought together leading academics and industrialists from a number of major blue chip (e.g. Unilever, BA, BT, CCN, Rolls-Royce) and high tech firms (e.g. Cogsys, Integral Solutions, Telictech, CCL) to discuss the requirements that industry would wish to place on the knowledge technologies of the future. SAKA II, held at Nottingham in April 1997 under the coordination of Professor Nigel Shadbolt, involved leading knowledge technology academics from the UK, Europe and the US in discussions on the industrial requirements that SAKA I had elicited.

These meetings identified a felt need within industry for Advanced Knowledge Technologies and sketched the intellectual agenda that evolved into AKT.

Knowledge Context: Knowledge as a Source of Competitive Advantage

Intellectual resources – knowledge – have fast become the major source for the creation of value. Market to book ratios of firms' values have become increasingly large as knowledge has overtaken labour or capital as the basis of competitive advantage in developed economies.

Pharmaceuticals, software, foods, telecommunications – examples abound of products and markets where raw materials are of negligible value compared to the know-how that goes into design, assembly and marketing. Unsurprising, then, that a recent survey in the *Economist* (18.11.00) placed effective knowledge management as one of the ten essential requirements of management in the Internet age.

Technology Context: The Semantic Web — and AKT

The World Wide Web is currently an enormous source of information, but its anarchic roots – the reason for its success of course – leave us also with a tangled spaghetti Web. The Web is designed for human consumption, and the resource-heavy activities of extraction and interpretation of information are overwhelmingly left for the human user. Web page after web page is downloaded, inefficiently examined, hurriedly assessed and stored, and probably never looked at again. Distributed, knowledge-intensive work, and the broader imperatives of the knowledge economy, mean that organisations will need to exploit Web/intranet resources as far as possible, but its focus on presentation for human perception increases the 'infosmog' for machine-support.

An evolution of the Web's infrastructure is required. Following the vision of W3C's Tim Berners-Lee, moves are afoot to make the concepts underpinning web content explicit (i.e. machine-readable and interpretable). Put another way, web pages in the future could contain 'annotations' that explain the content of that page in terms that a machine can understand. For instance:

- **Now:** to find an online course, you search for possible course providers, check on their status, download course details, check that they meet your requirements and that you fulfill theirs, fill in an application form and transfer the appropriate funds.
- **Future:** you set an agent to look for and make use of a service which accesses course providers, vets likely candidates using another service and finally retrieves course descriptions and requirements. Your agent can then present you with a list of candidates which meet your requirements and whose requirements you meet, fill in an application form and, finally, employ a service to negotiate payment terms and to ensure that credit accrued from other institutions is recognized.

AKT is ideally positioned to drive the requirements, and set the standards, for the Semantic Web to support organisational knowledge management. For instance:

- While has the Web has gathered unprecedented critical mass, technically it is extremely primitive as a hypertext system. AKT members are world leaders in **hypertext systems that have evolved several generations** on from static content and 'goto' links.
- It's fine to have machine-readable metadata, but **how to define and manage the concepts that this metadata must express?** Ontologies are a central component in the Semantic Web, an area in which AKT members have been working long before the Web emerged.
- AKT is already building **knowledge services** for the Web, with agents negotiating with each other, and reasoning about the content of ontologically marked up resources.
- Even if we can represent concepts explicitly, **who is going to do all this knowledge acquisition, formalization and encoding work?** How much consensus between disparate parties must there be to have something useful? The Semantic Web, like the Web, is not a technical system: it is a *socio-technical* system. AKT has expertise in the cognitive and collaborative demands of using formal notations and tools associated with knowledge-based hypermedia.

The current levels of research and business interest in the Semantic Web reflect the pressing need to evolve a **knowledge-level infrastructure**. It's early days yet, but AKT is already deeply involved in international efforts, and industrial collaboration with us is an ideal way to track developments, with the additional benefit of our specific orientation to knowledge management. Like any technology, the Semantic Web will need to be *embedded* into an organization's existing infrastructure, knowledge flow and working practices. See **AKT Technical Agenda** below, for more details, particularly on the knowledge lifecycle.

Background: *The Semantic Web, Tim Berners-Lee et al. Scientific American, May 2001. [www.scientificamerican.com/2001/0501issue/0501berners-lee.html]*

5. Knowledge Dynamics, Knowledge Services

As we have emphasised, to say that technology is crucial is far from saying that it is 100% of the solution to KM problems. Technology lives within other contexts than the knowledge context. Technology has to be used within organisations to meet the problems to which those organisations need to respond. Therefore it is essential that we come to understand the particular requirements of the firms and organisations that will wish to use advanced knowledge technologies, so that we can provide genuine solutions to real problems.

The aim of AKT is to develop a deeper understanding of the *dynamics of knowledge* in a technological, industrial context, the ways that knowledge might be said to 'flow' around organisations, and how the use of knowledge should be integrated into the Semantic Web. This will enable us to become a *knowledge service provider*. AKT's research will enable knowledge solutions to be tailored to particular knowledge problems, rather than requiring your whole knowledge management practice to be uprooted and replaced.

Put another way, AKT's research agenda is **not**: develop technologies to facilitate knowledge management. Rather, the agenda **is**: understand the *dynamics* of knowledge, and thereby understand the roots of knowledge management problems; understand the *properties* of knowledge, and thereby develop the most effective representations and technologies; understand the *capabilities* of knowledge technologies, and thereby match appropriate solutions with problems.

Understanding Knowledge Dynamics

What do we mean by the *dynamics of knowledge*? At this early stage, suffice to say that we aim to engage with a wide variety of perspectives on issues such as the nature of individual and collective organisational knowledge, how 'it' may be said to 'flow', or what it means to say that we have 'captured', 'stored', or 'reused' knowledge. Do different technologies encourage the dominance of particular metaphors for knowledge? For instance, what assumptions do we make when we use formal ontologies to 'model knowledge'? We draw inspiration from relatively recent work that has articulated the nature of the *"practices"* from which the term *community of practice* derives its name, a concept that has gained enormous currency in recent years within both the KM research and business communities. *Practices* are the stuff of everyday office life, emerging through the interplay of *tacit processes* and *formal artifacts* (see left).

> …Such a concept of practice includes both the explicit and the tacit. It includes what is said and what is left unsaid; what is represented and what is assumed. It includes language, tools, documents, images, symbols, well-defined roles, specified criteria, codified procedures, regulations, and contracts that various practices make explicit for a variety of purposes. But it also includes all the implicit relations, tacit conventions, subtle cues, untold rules of thumb, recognizable intuitions, specific perceptions, well-tuned sensitivities, embodied understandings, underlying assumptions, and shared world views. Most of these may never be articulated, yet they are unmistakable signs of membership in communities of practice and are crucial to the success of their enterprise.
>
> Etienne Wenger, *Communities of Practice* (CUP), 1998, p. 47

For instance, we see formal ontologies (regarded by many as a foundation for the Semantic Web) as symbolic tools within a community of practice. We have shown that when there is reasonable consensus, it opens the possibility for introducing knowledge services based on an ontology co-constructed by knowledge engineers with stakeholders. An ontology can reify a "shared world view", codifying "well-defined roles", "specified criteria" and "codified procedures" as described by Wenger. We regard symbolic representations such as ontologies as *boundary objects* whose role is to support communication and negotiation over meaning between communities of practice.

Understanding what we have termed 'knowledge dynamics' will draw on our expertise in cognitive science, philosophy and collaborative work processes, and as we build and integrate technologies, we will maintain an ongoing dialogue with other leading KM research and industrial groups for whom this is of particular interest. This should lead to a much deeper understanding of how knowledge technologies impact, and are moulded and sustained by, the everyday work practices of the communities using our tools.

Configurable Knowledge Services

In tandem with studying knowledge dynamics, the design and implementation of our technologies will be shaped by a number of principles:

- We will aim for maximal flexibility.
- We will incorporate the best aspects of standard or well-known software (if Lotus Notes will do the trick, then why force people to learn new languages and interfaces?).
- We will aim to influence standards.
- The Internet is crucial. We hope that much of the functionality we provide will be accessible using a standard web-browsing interface, and we will focus on how knowledge can best be represented, used and managed on the Semantic Web.
- We will not try to integrate our approaches by providing a single overarching system.
- We will not assume that users will be willing or able to scrap their existing systems, learn new representation languages, or reorganise their knowledge management teams. Of course, solving a problem or responding to an opportunity will require change to corporate practice. Nevertheless, our research is intended to help determine the *appropriate* amount of change, and so avoid the extremes of mere cosmetic tinkering on the one hand, and overreaction on the other.
- Where we do provide new methods, languages, tools or whatever, we will always seek a flexible approach to dissemination and product take-up, to provide all and only the knowledge services our users need.

Our project is focused and structured around the six knowledge challenges outlined above. However, that does not mean that we are forcing knowledge managers and other professionals to 'see' or conceptualise knowledge in a particular, rigid, way. We maintain that the six challenges will present themselves in different guises to anyone who is concerned with the problems of maximising the value of knowledge. (See the Appendix on how AKT's lifecycle models maps to other knowledge lifecycle models.)

> **Technology Context: Ubiquitous Computing**
>
> A trend that must be recognised is ubiquitous computing. The Internet is largely understood from the point of view of the currently dominant user interface, the PC. However, it is clear that Internet access will be much more widely available in future thanks to developments in communications and wireless technology. The networked computer is melting into the infrastructure: when you can access the Internet from mobile phones, handheld computers, or even your car, keys or credit card, it is clear that many of our ideas about the presentation and publication of knowledge and information using the Internet will need to evolve.

6. AKT: Technical Agenda

To an extent, the technical agenda of AKT is relatively unpredictable, given our remit to perform imaginative research and to "think the unthinkable". But the general pattern of the first few years of the project is clear, and will aim to pull together all our strengths (see box) and harness them to the problems and goals at hand.

Track Record: The AKT Consortium and the Knowledge Challenges

The research groups coming together in AKT have been active in various closely-related but complementary fields for some years. The full extent of their achievements can be seen in the individual partners' websites, but here's a thumbnail sketch.

Acquisition. See Aberdeen's innovative work on cooperative KA and knowledge refinement systems, or the breadth of work done by Shadbolt's group at Southampton on KA tool design and evaluation, culminating in the development of portable workbench PC PACK. Sheffield's GATE architecture makes it possible to integrate developments in computational linguistics and KA. The OU's Compendium approach supports real time capture of hybrid, group knowledge during meetings.

Modelling. The OU is a world centre for work here – just consider their modelling language OCML or their work on web-based ontology construction (e.g. WebOnto). Shadbolt's team's GDM methodology supported the construction of domain models, while Edinburgh's Enterprise ontology can be used to represent the organisational context.

Reuse. Southampton's MEMOIR uses distributed OO database technology for corporate information systems. Aberdeen has a semi-formal framework relating problem-solving, KA and knowledge refinement. The OU is involved in IBROW3 developing an intelligent broker to help configure knowledge-based items from web-based components. Edinburgh is deeply involved in a number of standardisation initiatives, such as PIF and I-X.

Retrieval. MAVIS and MAVIS 2 at Southampton have developed multimedia architectures for extracting video, image and sound. Aberdeen has systems to learn individual user selection preferences. The OU's document enrichment idea allows the creation of extra structure for documents which can be exploited for search. Shadbolt's group produced IMPS, an agent architecture for doing online KA and retrieval. Work at Edinburgh has moved in the converse direction, and tries to improve search by pruning knowledge structures.

Publishing. For producing visualisations of content, we can boast the Edinburgh Software Systems and Processes Group website, the OU's in-house online newspaper KMi Planet, and WWW-document publishing tool D3E, and Southampton's multimedia authoring and presentation tool Multicosm. Sheffield's IE tools can be used to retrieve knowledge dynamically from large texts, and summarisation tools can produce short customised texts as a result.

Maintenance. Aberdeen are acknowledged experts in the verification and validation of knowledge bases and KBSs. Shadbolt's group has produced APECKS, a collaborative ontology construction tool which highlights consensus and conflict, and has also investigated the issues raised by certification and recertification of KBs. Edinburgh are involved in the US High Performance Knowledge Bases project, concerned with building and testing tools to cope with the problems of scale.

Knowledge acquisition (KA) is a field which has reached a certain level of maturity. It began as part of the drive to build knowledge-based systems, and was a line of research devoted to developing methods and software tools to provide knowledge content for such systems. There are many tools and techniques available, and a number of integrated workbenches and methodologies on the market. It is not the intention of AKT to reinvent this particular wheel.

However, where AKT can make a difference is in a number of tricky, non-mainstream areas of knowledge representation from which the extraction or acquisition of knowledge is non-trivial or poorly-understood, such as group KA (the acquisition and integration into a single model of knowledge from a number of experts), incidental KA (the acquisition of knowledge as a by-product of other processes) and diagrammatic KA.

One area of particular interest is that of multimedia KA, i.e. coming to understand (finding models and structures for) knowledge expressed in non-text media. One example of this might be discovering the relationship to an organisation's business processes of a series of videoed 'war stories' or other shared experiences; another might be the marking up of mpeg files with semantic tags. The range of possibilities opened by the use of multimedia in a KA context is both an opportunity and a problem. The opportunity depends on making effective use of the extra expressive possibilities that using media other than text suddenly confer. The problem to be faced is that text is much the most tractable medium for knowledge representation (hence the relative progress made by text- and logic-based KA), and that new regularities and techniques will have to be learned and created.

Another area where we would expect to make progress is the integration of natural language information extraction (IE) techniques with standard KA methods, for example applying domain ontologies to facilitate the IE from texts. Such ontologies could be automatically generated for simple IE tasks.

A third area is that of artefact enrichment, in other words, methods, languages and logics for annotating and enriching the content of objects such as web pages, KA materials, hypertext fragments (e.g. in a protocol analysis tool) etc.

Finally, the field of KA can be transformed from its current state by a shift in the technologies available for it. There are two obvious directions of research to exploit here. First, there is the harnessing of the Internet. As noted above, there have been integrated sets of KA tools in workbenches, but such integration can be carried a step further by integrating the tools in a web-based environment, thereby allowing, e.g. the use of web-based libraries of knowledge model components, or the generation of knowledge models in XML.

The second direction of technological change for KA is that of software agents. Much KA, e.g. from information repositories on the WWW, might be automated by using intelligent agents, armed with structuring schemata such as ontologies of the domain in question.

In **knowledge modelling**, the central issue with which our vision is concerned is the semantic web (see box). The Internet and the WWW are clearly key focuses of AKT's future development and exploitation, and we will need to track and influence the development of the semantic web. We shall investigate the emerging suite of languages for expressing knowledge and exploiting the possibilities of the semantic web. The use of *ontologies*, powerful formalisms for expressing and structuring conceptual schemes for domains (as noted above), will be a central idea to look at here, but will not be the only focus of our work.

There will also be a drive to examine special-purpose modelling formalisms for particular important subsets of discourse. One example will be the modelling of best practice procedures; the dissemination and sharing of best practice across an organisation or firm is clearly going to be one of the most important applications of knowledge sharing. Another example, again central to our user base's concerns, is the modelling of argumentation and design rationales. Often with an artefact – which may be a piece of knowledge as well as a concrete object – knowing what design decisions, constraints and pressures led to the creation of an object of that particular form is key to understanding a domain.

Finally, we are also working on the next generation of general purpose tools to support knowledge markup (i.e. enabling users to enrich web content semantically). These tools assist in the creation of one or more semantic layers from different perspectives, using different ontologies, and amenable to multiple representations for different user groups.

In **knowledge reuse**, the aim will be to exploit the software engineering principle of reuse as much as possible. An obvious place to start will be the field of knowledge engineering, in which there are basically three reusable types of object. First, as discussed elsewhere, there are ontologies. Second, there are *problem-solving methods* (PSMs), which describe ways of solving particular problems, and specify the types of knowledge that will be required for the method to work. PSMs are generic objects – hence reusable – and when knowledge of the requisite type is acquired and used by the method, the result is a domain-specific description of the problem-solving in the domain. Thirdly, there is also the possibility of reusing knowledge bases (KBs) themselves, and we shall be looking at ways of adapting existing KBs for use in other contexts.

Of course, not all reusability issues are based around knowledge engineering. For example, it is possible to identify user profiles, such as navigation strategies for getting round a complex website. Reuse of such strategies is very likely to be a valuable extension of our web technologies.

In **knowledge retrieval**, the aim would be to develop user-friendly tools for retrieving knowledge from repositories. One obvious place to begin is to try to exploit natural language. The use of natural language in retrieval has two aspects. First, there is the use of natural language as the basis for the interface to our knowledge services. For instance, ontologies could be used to interpret knowledge queries in natural language forms from users. Second, there is the extraction and retrieval of information from texts and other unstructured sources.

The other focus of our effort to meet this challenge is focused around the technologies behind the interface. Our expertise in a number of areas, such as search engines, multimedia thesauri, capability decisions and statistical sampling will be used to increase the efficiency of retrieval technologies, and to provide technologies that cope with the new demands of the semantic web.

In particular, we are very alive to the technological possibilities of the enriched markup of web pages. Given the central importance of the Internet, web pages will clearly be a vital storage and dissemination medium for knowledge. The use of semantic markup technologies such as XML or RDF, driven by our proven knowledge technologies such as ontologies and dynamic hypertext link generation, will enable efficient and accurate retrieval of web pages from the web itself, or from highly structured and possibly very complex intranets.

For the challenge of **knowledge publishing**, there are two promising areas for the beginning of our work. Firstly, there is the construction of visualisations. These are essential for making knowledge usable, however well managed it is before it reaches the user. The aim would be to take the knowledge that is stored in knowledge models and synthesise web pages from such models according to the requirements of the user. Agent technology could be exploited to publish online; another line of enquiry would be the use of natural language summarisation tools linked to automatic web publishing tools.

Secondly, there is the personalised presentation of content; methods for doing this may be elicited from computational linguistics and natural language generation technology, or multimedia authoring tools and distributed link services.

In **knowledge maintenance**, the challenge is to ensure the continuing usability of knowledge models and other artefacts after creation and their original exploitation; clearly this is a prerequisite of reuse. We will be applying our experience on essential post-development maintenance issues, such as verification & validation, certification and auditing. Tool support, e.g. debugging tools, will also be an issue we will be addressing.

An interesting part of the knowledge maintenance challenge is the idea of understanding the quality of the knowledge, which of course is an essential factor in any decision to continue maintaining or scrapping some knowledge repository. Placing a value on knowledge assets is one such theme we shall be examining; the economics of intangible assets such as knowledge is an interesting area for study. As it stands, few knowledge asset valuation methodologies if any focus on the content or context of the assets themselves; this is likely to be a key issue with respect to understanding the place of knowledge (and knowledge technologies) in its real-world context. A further intriguing issue that follows from this is whether it could be useful to ‘forget’ knowledge, and what ‘forgetting’ might mean in such a context.

Of course, an important meta-issue is the **integration** of the various approaches to the challenges. There are a number of strategies in place to do this. Of course, as we have pointed out already, it would be a grave mistake to assume that technology was the major part of any solution to knowledge management problems. Technology can make a massive contribution, but we must be prepared to fit into the frameworks of our potential users, rather than expecting them to adopt our views.

Nevertheless, it is clear that we should have a view of the lifecycle of knowledge, and ensure that we can link our approach based on the six challenges to others' understanding of the flow of knowledge in their institution. We have already discussed how the six challenges are ubiquitous within knowledge management (and in the Appendix show how they can be mapped onto some other well known models). As a further structuring principle, we recommend the use of the knowledge engineering and management methodology CommonKADS (see Further Reading) as a standard for documentation and development. This does not mean that we advocate following the method precisely or completely, but it does constitute a clear, understandable and thorough amalgamation of software engineering and knowledge engineering. To that extent, it is a welcome advance in the field.

In terms of infrastructure, the aim is for AKT to create a light infrastructure giving maximal freedom and Internet-accessible API. A layered approach will be used to achieve interoperability. A 'knowledge bus' will mediate knowledge-level interoperability between applications, and work is underway to examine precisely what expressive power such a bus will need. The current version of the infrastructure, AKT-0, is web-based, using HTTP and RDF(S)to communicate knowledge about relevant concepts and instances between a web-browser and an ontology service. As the project continues, the scope and scale of this infrastructure will increase dramatically and appropriately.

Further Reading

Major AKT publications and releases will appear at www.aktors.org/publications/

There are useful collections of papers in *IEEE Intelligent Systems*, on **knowledge management and the Internet** (May/June 2000) and the Semantic Web (March/April 2001). An **AI perspective** on knowledge management can be found in Reid G. Smith and Adam Farquhar's article 'The Road Ahead for Knowledge Management' in the *AI Magazine* of Winter 2000. The influential **CommonKADS methodology** for knowledge management is presented in Guus Schreiber et al *Knowledge Engineering and Management* (MIT Press 2000). The possibilities for **ontologies** can be found in a number of articles, though we might single out Mike Uschold's paper 'Knowledge level modelling' in the *Knowledge Engineering Review*, volume 13 (1998) as an introductory survey, and for an important application relevant to AKT, see Enrico Motta et al. (Open Univ.), 'Ontology-Driven **Document Enrichment**' in the *International Journal of Human-Computer Studies* 52, (6), 1071-1109 (2000).

On the **knowledge economy**, key references are Charles Leadbeater, *Living on Thin Air* (Viking 1999), Ikujiro Nonaka and Hirotaka Takeuchi, *The Knowledge-Creating Company* (OUP 1995) and K-E Sveiby, *The New Organizational Wealth* (Berrett-Koehler 1997). A review of **E-management** issues, including knowledge management can be found in Frances Cairncross's 'Inside the machine', a survey in *The Economist* of November 18th, 2000.

7. AKT: Applications Agenda

It is an explicit aim of the EPSRC's Interdisciplinary Research Collaboration (IRC) programme to encourage imaginative and far-sighted research. Nevertheless, the AKT consortium wishes to see its own work taken up and used as cutting edge methods for knowledge management and leverage in industry, government, science, etc. The aim is for the project to form the centre of an entire research programme; we wish to attract researchers in knowledge management and knowledge technologies in both the academic and industrial communities. We hope to bring interested parties together for related consortium-led research, partly by acting as a magnet for the best thinkers and thinking in the field, and partly by providing a conceptual framework in which the links between disparate research programmes are made clear.

To this end, active support for our research programme has been sought throughout the process of setting up AKT. At the time of our initial application, seventeen companies, from blue chips to hi-tech, had pledged some level of commitment to AKT.

Active Navigation	Baker Hughes
BG Technology	Boeing
BP Amoco	British Airways
Clifford Chance	Epistemics
IMS	ISX
Open University Learning Technologies and Teaching Office	Parametric Technology
QinetiQ	Rolls-Royce
SELLIC	Teknowledge
Unilever	

We are grateful for the support provided by all these organisations. Throughout the AKT project, we intend to add to this list. We are confident that as our research is disseminated through the research community, the quality of our approach to the six challenges will become clear.

There are three main ways of following AKT, depending on the level of interest.

1. As academics, we are disseminating our research through the public domain: journals, conferences and workshops. Our website www.aktors.org documents the progress of the project, and we are already making an impact at major conferences.
2. A number of interested organisations will have a direct link with the consortium via the AKTors club, in which they will declare a specific interest in the AKT research, with an identified point of contact.
3. Throughout the six-year run of AKT, a steering committee made up of senior and committed industrialists will oversee the research produced by AKT. The committee will not direct our research, and neither will the companies represented on it provide money. In this sense, AKT is independent. However, the committee will comment in detail on the quality and relevance of the research we have performed. In this way, we will ensure that AKT's products are targeted to the actual problems of knowledge management as understood on the ground.

As well as this, we are applying our technology to a series of *testbeds*, real-world problems for knowledge management that will play a triple role in our research programme.

1. They will help prove the concepts associated with our research. Success here will show that our approach can work in real-world situations – and could also help to show where there are problems.
2. They will facilitate the integration of our approaches to the six challenges.
3. They will engage our users with the project.

The testbeds include:

1. Knowledge auditing – This is the problem of identifying areas within an organisation where knowledge crucial to its business processes is missing, out of date or otherwise inaccurate, and areas where the organisation has a particular knowledge strength. Knowledge audits will help organisations (a) acquire the knowledge that they are lacking, and (b) preserve their competitive advantage by exploiting areas of strength.

2. Industrial design processes – In large manufacturers, development of new products from conception to market may take years. Redesign or replacement inevitably puts strain on such firms' knowledge resources. Lessons painfully learned in one iteration may be lost as the people involved move on within the firm or elsewhere. Storage and maintenance of experience, for example on corporate intranets, can help preserve the knowledge; the management problems this raises are exactly the sort of area AKT is interested in. Rolls-Royce will work with us on this testbed.
3. Publishing – With large repositories of unstructured or partially structured text, perhaps on heterogeneous topics, the difficulties of searching for particular relevant items are well-known. Using a large corpus of scientific papers, we are applying natural language processing techniques to extract specific types of information.
4. Scientific research management – Perhaps the most pure example of knowledge work is that of scientific research. We have the perfect testbed under our noses – we will use our own products to coordinate our own research.

It is expected that other testbeds will be developed from the interests of AKT club members.

Our independence and our remit from the EPSRC ensure that our research will be visionary, and will go beyond current approaches to transform the field. Our links with industry will ensure that our products are practical and applicable to real-world situations.

8. AKT: Philosophical Agenda

For two and a half millennia, the study of knowledge was driven by the question "how do we know what we know?" The aim of the philosophy of knowledge – epistemology – was to prove that our reasoning was sound, and that something known is somehow distinct from something merely believed, against the arguments of a sceptical observer. This, of course, is a perfectly legitimate intellectual quest.

However, in recent years the study of knowledge has focused on a rather more prosaic aspect of knowledge. Knowledge has an economic value; it is a resource. Knowledge can provide a person, a firm or a country with a competitive advantage. It can lead to innovation, to smarter methods of production, and to informed management and understanding of risk. The *justification* of knowledge has become less of an imperative than the *leveraging* of it to create value.

In this context, our understanding of knowledge has changed dramatically. Knowledge, in this view, need not be a psychological state; organisations, computers and processes as well as people may hold knowledge. It can be bought and sold, used and misused.

Nevertheless, we certainly do not know all the properties of knowledge. What is a piece of knowledge worth to its owners? What processes are reliable producers of knowledge? Can the completeness of a body of knowledge be assessed? How does a piece of knowledge alter, if at all, when its representation alters, and how tight is the connection between knowledge and its representation? Can we know when a piece of knowledge has reached the end of its useful life, and if so, how? Where can tacit knowledge be found, and how? How can different understandings of some domain be merged to provide a single model? There are countless epistemological problems waiting to be answered.

AKT, by virtue of the breadth of its scope and the depth of its ambition, will be obliged to address these questions, and others just as fundamental. To an extent, the technologies we develop will change the way that knowledge as a resource or a commodity is understood. By advancing coherently, rather than piecemeal, we have a chance to develop an intelligible view of what knowledge is. In some sense, we are (re)defining knowledge instrumentally by specifying or circumscribing what it can do.

9. AKT: Political Agenda

The knowledge economy is here to stay. Companies are already negotiating the seismic shifts of e-commerce, and politicians of all creeds are clear that for any country to get ahead, the best way is to exploit the productivity that knowledge can produce. No-one is going to advocate a return to a muscle-bound economy, where low wages and low education attract low-grade unskilled manufacturing and labouring jobs.

We can expect to see more e-government. Government infrastructures the world over are 'going digital' – but will they reap the benefits that a knowledge-level infrastructure can provide? Yet the limits of what can be done with knowledge will be set, not by the imagination (or perhaps the lack of it) of our rulers, but rather by the capabilities of the tools that government has on its collective digital desktop, and the *quality* as opposed to *quantity* of information and expertise it can readily access in a coherent manner.

To that extent, Advanced Knowledge Technologies are key not only to the transformation of Britain into a fully-fledged knowledge economy, but also to the goals and policies of the politicians and businessmen who want to be part of that transformation. Through our steering committee and wider links with the non-scientific world, we intend to make available the 'next generation pens and pencils' of the 21st Century knowledge society.

10. AKT and other Knowledge Management Philosophies

There are a number of different views of knowledge in organisations. This is not the place for an exhaustive survey, but as you may be familiar with other models, we'll just skim through three in order to show how they map to AKT's six knowledge challenges.

Nonaka and Takeuchi model

For a simple, and well-known, opener, consider Nonaka and Takeuchi's groundbreaking work *The Knowledge-Creating Company*. In that book, they see knowledge creation in Japanese firms as a series of repeat interactions between tacit and explicit knowledge, with four possible transfer permutations, as shown in the diagram.

From / *To*	Tacit Knowledge	Explicit Knowledge
Tacit Knowledge	(Socialization) **Sympathized Knowledge**	(Externalization) **Conceptual Knowledge**
Explicit Knowledge	(Internalization) **Operational Knowledge**	(Combination) **Systemic Knowledge**

- **Tacit to explicit**. This type of transformation takes knowledge from tacit contexts (e.g. expertise of individual workers), and renders it explicit (e.g. in a database). This sort of making explicit is a central task for *knowledge acquisition*, where many tools and methods have been developed to provide exactly such knowledge capture. Furthermore, to make knowledge explicit, it should be represented in a usable, (human- or machine-readable) format, where we meet the challenge of *knowledge modelling*.

- **Explicit to explicit**. This might involve standard knowledge sharing and transfer, such as the collation and storage of sales data in large databases. To find the important knowledge to transfer from a repository, we encounter the challenge of *knowledge retrieval*. Generic libraries may also be used to provide a basis for context-dependent knowledge to be expressed, which will involve the challenge of *knowledge reuse* (of the generic), as well as *knowledge acquisition* (of the domain-specific). Explicit to explicit transfer will also require knowledge repositories to be kept abreast of the latest position in the domain – indeed keeping knowledge up to date may involve exactly this type of knowledge transfer (e.g. adding new rules and deleting old ones from a rule base). This is the challenge of *knowledge maintenance*.

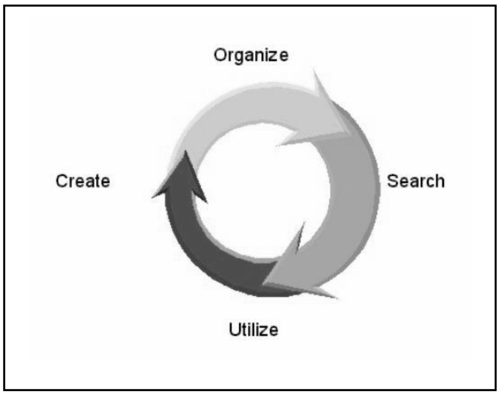
- **Explicit to tacit**. This sort of transfer might involve using systems that contain rich arrays of contextual information for knowledge sharing with individual operatives, such as using case based scenarios of previous corporate experience as opposed to rule-based procedures. This sort of example is a clear case of *knowledge reuse*. Furthermore, given that the important contextual analogies will not always be clear cut, there will be the important challenge of *knowledge retrieval*. The knowledge must also arrive on the user's desk or computer screen in a form which will facilitate his extracting the value from it, an issue of *knowledge publishing*.

- **Tacit to tacit**. This is the sort of transfer that goes on during training, and again the important lessons of experience may be half-hidden in the context rather than made explicit. It is by finding the analogies with his or her current situation that the trainee extracts the value from a narrative account. *Knowledge retrieval* is an issue here, to the extent that the tacit source needs to be located, and *knowledge publishing* will also be a key challenge.

Nonaka and Takeuchi are more concerned with the flow and creation of knowledge within a company, but even so we can see how each of their transfer permutations is essentially and directly concerned with the six knowledge challenges. Anyone who seeks to implement a Nonaka-Takeuchi-style structure to assist the dynamics of innovation within a firm will need to address each one of our six challenges. Their case studies could just as easily be conceptualised as a series of coordinated attempts to meet the challenges. And we have only discussed knowledge *transfer* here. Knowledge *acquisition*, *publishing* and *maintenance* in particular might all be background or input processes as well.

John Thomas/IBM model

As a second example, we might look at a model used by John Thomas of IBM[2], who sees four important processes as follows.



Here the four processes map onto our challenges as follows.

- **Create**. The creation of knowledge is an instance of *knowledge acquisition*.

- **Organize**. Organising knowledge involves storing the knowledge in a suitable way, i.e. *knowledge modelling*. Once the knowledge is stored in the first place, organising it becomes an ongoing problem of *knowledge maintenance*.

- **Search**. Finding knowledge obviously involves meeting the challenges of *knowledge retrieval* and *knowledge reuse*.

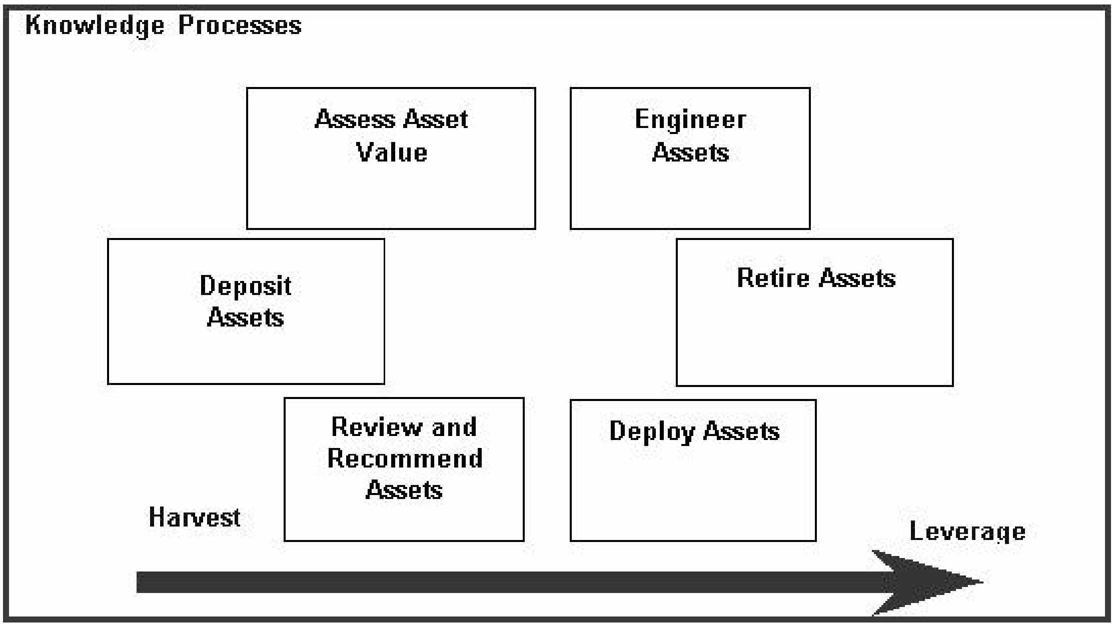
2 www.research.ibm.com/knowsoc/presentation

- **Utilize**. Deploying knowledge will require its being put into the right form, which is a *knowledge publishing* problem.

Computer Science Corporation model

Theory and practice, of course, may differ radically. For a third example, then, we will use the Knowledge Program of CSC (the Computer Science Corporation), a *Fortune 500* company with over 60,000 employees working on over 700 locations worldwide. Their knowledge management program was described recently by Pierre Dagneau, CSC's Knowledge Program Manager, in a seminar on knowledge management, and we take our account from his presentation.[3] Note that CSC have no formal connection with AKT; we have taken this paper as an example because it is an especially clear presentation of an especially successful knowledge management programme.

Consider CSC's view of the essential knowledge processes, shown below.



In this view of the essential knowledge processes, we can see an important overall structure, moving from *harvesting* knowledge to *leveraging* it. Again, there is a different take on knowledge from the six challenge idea that we have developed, but it is pretty clear that there are strong links between CSC's understanding of the knowledge within their organisation and its dynamics, and AKT's view of the challenges to be met by any knowledge management methodology. Consider the six processes that CSC undertake.

- **Deposit assets**. The initial creation of knowledge assets is of course an instance of *knowledge acquisition*. To ensure that the knowledge appears as an asset, i.e. as a usable resource as opposed to junk data, the challenge of *knowledge modelling* must be met.
- **Assess asset value**. The appraisal and reappraisal of the value of an asset to a firm in a particular context is an important issue of *knowledge maintenance*.
- **Review and recommend assets**. The selection of particular assets for use in a context is an instance of the challenge of *knowledge retrieval*.
- **Engineer assets**. The engineering of assets to meet particular situations rather cuts across the AKT account, but depending on the context, is related to a number of challenges. If the assets are being engineered immediately after deposit, we have *knowledge modelling*. If they are generic assets being made domain-specific, then we have *knowledge reuse*, and in certain circumstances *knowledge acquisition*. If they are being amended to get

3 aifbhermes.aifb.uni-karlsruhe.de/dagstuhl-km-2000/presentations/KM%20Seminar%20Germany.ppt

them in the right form for some particular use or user, then we are looking at *knowledge publishing*.

- **Deploy assets**. Here to select the assets that are the right ones to deploy at some particular context, we must meet the challenge of *knowledge retrieval*. To leverage successful experiences – a key goal of CSC's knowledge management program – the challenge of *knowledge reuse* is important. Furthermore, to ensure that the assets appear in the optimal form, particularly when the recipient will be a human user, the challenge of *knowledge publishing* must be met.
- **Retire assets**. The assessment of the moment when an asset has lost its value to a firm is an extremely difficult task, part of the challenge of *knowledge maintenance*.

Our conclusion is clear. There are a number of different ways to conceptualise the flow of knowledge around an organisation or firm; some of these will be very specific to a firm. This can make it difficult for a technology initiative such as AKT to ensure that its products will be relevant to an organisation. If the intention of the project is to impose a certain structure on the firms it deals with, then the result will inevitably be failure. But in AKT, we have isolated a number of challenges that any knowledge management method will need to meet, however the method conceptualises knowledge processes. These challenges are the structuring principle of AKT's research, and so any firm can use AKT's products by simply establishing which of the challenges it needs to meet. The research we perform will enable us to supply a complete solution for those who need one, or partial solutions tailored to particular firms' cultures when that is appropriate.

11. Conclusion

AKT is a project of extraordinary ambition, with the remit and resources to 'think the unthinkable'. AKT aims to create technologies to take knowledge through the whole of its lifecycle from acquisition to obsolescence, technologies that will be integrated and yet flexible enough to apply to the specific problems of individual users in their organisations. Our systems will be grounded in industrial and other organisational contexts in order to understand how the dynamics of knowledge shape – and are shaped by – our tools. We recognise that our assumptions about knowledge technologies may be challenged as our tools are adopted, and appropriated by knowledge communities.

The integration and deployment of state of the art knowledge technologies in authentic work contexts is the distinguishing mark of AKT, and we invite you to participate as industrial partners in advancing this research and development programme.