Open Research Online

The Open University's repository of research publications
and other research outputs

Features for killer apps from a semantic web perspective

Book Section

How to cite:

O'Hara, Kieron; Alani, Harith; Kalfoglou, Yannis and Shadbolt, Nigel (2008). Features for killer apps from a semantic web perspective. In: Li, Eldon Y. and Yuan, Soe-Tsyr eds. Agent Systems in Electronic Business. Information Science Reference, pp. 265–288.

For guidance on citations see FAQs.

Version: Accepted Manuscript

Link(s) to article on publisher's website:
http://www.igi-global.com/Bookstore/TitleDetails.aspx?TitleId=59

oro.open.ac.uk

Features for Killer Apps from a Semantic Web Perspective

Kieron O'Hara, Harith Alani, Yannis Kalfoglou, Nigel Shadbolt

Intelligence, Agents, Multimedia
School of Electronics and Computer Science
University of Southampton, UK
{kmo, h.alani, y.kalfoglou, nrs}@ecs.soton.ac.uk

Abstract. Killer apps are highly transformative technologies that create new markets and widespread patterns of behaviour. Information Technology generally, and the Web in particular, has benefited from killer apps to create new networks of users and increase its value. The Semantic Web community on the other hand is still awaiting a killer app that proves the superiority of its technologies. There are certain features that distinguishes killer apps from other ordinary applications. This chapter examines those features in the context of the Semantic Web, in the hope that a better understanding of the characteristics of killer apps might encourage their consideration when developing Semantic Web applications.

1   Introduction

The Semantic Web (SW) is gaining momentum, as more researchers gravitate towards it, more of its technologies are being used, and as more standards emerge and are accepted. There are various visions of where the technology might go, what tasks it might help with, and how information should be structured and stored for maximum applicability [3][30][39][7]. What is certainly clear is that no-one who wishes seriously to address the problems of knowledge management in the twenty-first century can ignore the SW.

In many respects, the growth of the SW mirrors the growth of the World Wide Web (WWW) in its early stages, as the manifest advantages of its expressivity became clear to academic users. However, once the original phase of academically-led growth of the WWW was over, to the surprise of many commentators, the web began its exponential growth, and its integration with many aspects of ordinary life. Technologies emerged to enable users to, for example, transfer funds securely from a credit card to a vendors account, download large files with real time video or audio, or find arbitrary websites on the basis of their content.

It is important to note that the growth of the WWW had three separate but linked components, which led to three distinct sets of incentives. In the first place, there were quick wins from putting documents on the Web. Prior to the WWW there was a culture of privacy about documents, and the proselytisers of the Web had to convince a lot of people that documents should be published and made available to all. This was perceived as a risk by document owners, and involved breaking down preconceptions

about ownership, privacy, confidentiality and commercial advantage. To do this, when a person or an organisation posted documents, there had to be immediate and tangible gains from the individual act of publication — an increase in one's social circle, an expansion of business, or a wider set of business opportunities. Without such immediate and individual gains, independent of any future network effects, fewer documents would have appeared on the Web.

Secondly, however, those network effects also had to come into play. Because network effects are part of the context, and largely independent of individual decisions, they can sometimes be overlooked. But the creation of large business markets online happened because more and more people started an online existence. The evolution of social networks or multiplayer game scenarios depends on the critical mass of people spending a certain amount of time indoors next to the computer. But the network effects really do kick in, and an online presence for a business or a person is now so much more rewarding both financially and socially because so much of so many lives take place online. *In extremis* the failure to engage with the WWW can now mean the serious loss of business.

Thirdly, the tools had to be available for the WWW to take off. If creating html pages was at all difficult, involving steep learning curves, flaky or expensive software, and advanced design skills, then the WWW could not have integrated so easily with the rest of the environment. And if publishing documents online led to a backward step in information processing — if the documents, by being published, were somehow removed from an organisation's standard information management practices — then such an organisation might end up worse off than before, which would have strangled the global Web at birth, and it would have remained an academic tool. For instance, an organisation's practice in posting documents on the WWW might well compromise its version management control; one could imagine a situation that a draft document, posted on the Web, might be edited by two different people in parallel, and then two incompatible versions would circulate in parallel. The right tools, and the right management practices, needed to be in place in advance, to prevent such hiccups occurring; certainly not all of the potential pitfalls of posting pages would be easily predictable in advance.

The SW aims to do for data what the WWW did for documents. Most realistic visions of the future of a successful SW include a version of the WWW's exponential growth. The SW infrastructure should be put in place to enable such growth. With a clean, scalable and unconstraining infrastructure, it should be possible for users to undertake all those tasks that seem to be required for the SW to follow the WWW into the stratosphere, such as publishing their RDF, converting legacy content, annotating, writing ontologies, etc. There must be immediate gains, network effects, and the tools and management practices to prevent backward movement. Without this, the culture of privacy and data protection which now surrounds the storage of data will prevent the SW taking off. This culture needs to be adapted to the new technologies without compromising the important aspects of privacy [7].

However, that something is possible does not entail that it is inevitable. So the question arises of how developers and users might be persuaded to come to the SW. This type of growth of a network has often been observed in the business literature. Many

technologies depend for their usability on a large number of fellow users; in this context Metcalfes Law [21] states that the utility of a network is proportional to the square of the number of users.

Technologies which have this effect are called *killer apps*. Exactly what is a killer app is to a large extent in the eye of the beholder; in the WWW context, killer apps might include the Mosaic browser, Amazon, Google, eBay or Hotmail. Hotmail attracted over 30 million members in less than three years after its 1996 launch; eBay went from nothing to generating 20% of all person-person package deliveries in the US in less than 2 years. Of course, the WWW was a useful enough technology to make its own way in the world, but without the killer apps it might not have broken out of the academic/nerdy ghetto. By extension, it is a hope of the SW community that the SW might take off on the back of a killer app of its own.

The dramatic development of the WWW brought with it a lot of interest from the business community, and the phenomenon of killer apps has come under much scrutiny [16][27]. Attempts have been made to observe the spread of killer apps, and to generalise from such observations; the tight development cycles of WWW technology have helped such observations.

In this chapter, our aim is to consider the potential for development of the SW in the light of the killer app literature from the business community. Of course, it is impossible to forecast where the killer app for the SW will come from. But examination of the literature might provide some pointers as to what properties such an application might have, and what types of behaviour it might need to encourage.

2   Killer Apps and the Semantic Web

Killer apps emerge in the intersection between technology, society and business. They are technological in the broad sense of being artificial solutions to perceived problems, or artificial methods to exploit perceived opportunities (which is not to say that they need to have been developed specifically with such problems or opportunities in mind). Mere innovation is not enough. Indeed, a killer app need not be at the cutting edge of technological development at all. The killer app must meet a need, and be usable in some context, such as work or leisure or commerce. It must open up some kind of opportunity to bring together a critical mass of users.

To do this, killer apps have a number of features which have been catalogued by commentators. In this section, we will examine and reinterpret such features in the context of the SW. We reiterate that these features may not all be necessary, and they certainly are not sufficient; however they can act as an interesting framework to our thought on this topic.

The main point, of course, about a killer app is that it enables a superior level of service to be provided. And equally clearly, the SW provides an important opportunity to do this, as has been argued from the beginning [3]. There are obvious opportunities for any knowledge-based task or enterprise to improve its performance once knowledge sources are integrated and more intelligent information processing is automated.

2.1 The bottom line: Cost vs Benefit

However, merely providing the opportunity is not enough. Cost-benefit analysis is essential [16]. There are several aspects to costs. Obviously, there are financial costs; will people have to pay for the killer apps on the SW? Maybe not; there are many examples of totally free Internet applications, such as Web browsers, search engines, and chat messengers. Such applications often generate large revenues through online advertising. According to the Interactive Advertising Bureau UK[1] and PriceWaterHouseCoopers[2], the market size of online advertising in the UK for 2004 was £653.3m, growing more than 60% in one year. Free products may be very important in this context [27], and indeed killer apps are often cheaper than comparable alternative products [11].

But such costs are not the only ones incurred. There are also important resource issues raised by any plan to embrace the SW.

Conversion cost: As well as investing in technologies of certain kinds, organisations and people will have to convert much of their legacy data, and structure newly-acquired data, in particular ways. This immediately requires resources to support the development of ontologies, the formatting of data in RDF, the annotation of legacy data, etc., not to mention potential costs of exposing data in RDF to the wider world (particularly where market structures reward secrecy). Furthermore, the costs of developing smart formalisms that are representationally adequate (the fun bit) are dwarfed by the population of informational structures with sufficient knowledge of enough depth to provide utility in a real-world application [18]. Note also that such a process will require ascent of some very steep learning curves.

Much of the legacy data will be sitting around in relational databases, and this very large quantity of relational data is a major target of the SW initiative [7]. The mappings from relational data to RDF are fairly straightforward, and so the hope is that these conversions can lead to quick wins, particularly when an organisation has a lot of data distributed across its various divisions. However, the danger here — an important potential cost — is that an organisation takes data out of its usual information management streams. Seduced by the vision of the SW, it puts all its RDF in one place and sends it out into the world, without the tools to deal with it. Even though the older tools and structures were not perfect, they will of course have provided useful service, and a firm which takes its data out of its old-fashioned environment can find itself subtracting value from the data, not adding it. Joining the SW community without the proper information management regime in place is a recipe for disappointment and disillusion.

Maintenance cost: In a very dynamic domain, it may be that ontologies have to be updated rapidly [28][9]. The properties of ontologies are not as well-understood as they might be; areas such as mapping ontologies onto others, merging ontologies and updating ontologies are the focus of major research efforts. It is currently unknown as to how much such maintenance effort would cost over time.

Organisational restructuring costs: Information processing is integrated into an organisation in subtle ways, and organisations often subconsciously structure them-

---

[1] www.iabuk.net

[2] http://www.pwc.com/

selves around their information processing models [17], a fact implicitly accepted by the knowledge engineering community [36]. Surveys of organisations, for example, reveal that ontologies are used in relatively primitive ways; indeed, in the corporate context, the term 'ontology' is a generic, rarely defined catch-all term. Some are no more than strict hierarchies, some are more complex structures allowing loops and multiple instantiations, still others are in effect (sometimes multilingual) corporate vocabularies, while others are complex structures holding metadata [31]. Whatever their level of sophistication, corporate ontologies, support the systematisation of large quantities of knowledge, far from the traditional AI view of their being highly detailed specifications of well-ordered domains. Ontologies may refer to an internal view of the organisation (marketing, R&D, human resources, etc) or an external one (types of supplier and supplies, product types, etc). A recent survey showed that only a relatively small number (under a quarter) of corporate ontologies were derived from industry standards. The big issue for many firms is not representational adequacy but rather the mechanics of integration with existing systems [18].

With a complex organisation, working in a dynamic environment, it may appear a daunting task to develop an ontology to cover it. However, ontologies can be fairly lightweight, and their main aim is to facilitate communication and translation, not to produce a philosophically accurate model of the world. A series of small, overlapping ontologies, with *ad hoc* translators between them, will allow SW technologies to add value to data without too large an immediate cost. For instance, an application to fill in one's tax return will use data from one's bank about one's bank account. That data will be represented against the background of the bank's ontology, but the tax application need not commit to the whole of the bank's ontology, nor even translate all the concepts therein. It will suffice to translate the few concepts relevant to the problem where the two ontologies overlap. And this, in such technology, is what happens.

Tim Berners-Lee has argued, informally, that worries about ontology development are overblown. As the context of use and maintenance grows, and the number of users increases, the committees required to develop and maintain ontologies do not grow as quickly. The effort of developing large ontologies is obviously greater than the effort of developing smaller ones, but the effort *per user* will decrease. As long as ontology builders do not over-elaborate, it should be cost-effective for most organisations to devote resources to ontologies [6][7].

Transaction costs: On the other hand, it is also true that if the SW does alter information gathering and processing costs, then the result will inevitably be some alteration of firms' management structures. The result will be leaner firms with fewer management layers, and possibly different ways of processing, storing and maintaining information. Such firms may provide opportunities for new SW technologies to exploit, and a gap in the market from which a killer app may emerge.

It has long been argued that the size and structure of firms cannot be explained simply by the price mechanism in open competitive markets [13]. The allocation of resources is made using two mechanisms - first (between firms and consumers) by distributed markets and coordinated by price, but also (within firms) by the use of authority within a hierarchy (i.e. people get ordered to do things). The question then is how this relates to a firm's structure - when a firm needs some service, does it procure it from

outside and pay a market price, or does it get it done in-house, using workers under some contractual obligation, and why?

It is generally thought that such organisational questions are determined by the *transaction costs* within a firm [45][46], in other words, the costs of actually making a transaction, costs such as information processing and acquisition (for example, the costs of researching the market to find the best price), information asymmetries (outside firms know more about the nature of their product than purchasers of those products), uncertainty, incomplete contracting and so on. The promise of the SW is that many of the information gathering costs will be ameliorated, and at least some of the asymmetries will be evened up. The general result of this is likely to be a continuation of trends that we have seen in economies since the widespread introduction of IT, which is the removal of middle management ("downsizing"), and the outsourcing of many functions to independent suppliers. In the SW context, of course, many of those independent suppliers could well be automatic agents, or providers of web services. If the SW contains enough information about a market, then we might well expect to see quite transformative conditions, and several market opportunities. The killer app for the business aspects of the SW may well be something that replaces the coordinative function of middle management.

But we should add a caveat here: the marginal costs of information gathering will be ameliorated, but equally there will, as noted above, be possibly hefty sunk costs up front, as firms buy or develop ontologies, convert legacy data to RDF, lose trade secrets as they publish material, and so on. These initial costs may prove an extensive barrier to change.

Reducing costs: Here we see the importance of the increase in size of the user base. For example, the costs of developing and maintaining ontologies are high, but can be shared. Lightweight ontologies are likely to become more important [40]; not only are they cheaper to build and maintain, but they are more likely to be available off the shelf [31]. Furthermore, they are more likely to be easily understandable, mappable, maintainable, etc. The development of such lightweight multi-purpose ontologies will be promoted as the market for them gets bigger.

Similar points can be made about ontology development tools. Better tools to search for, build or adapt ontologies will spur their use or reuse, and again such tools will appear with the demand for them. And in such an environment, once an ontology has been developed the sunk costs can be offset by licensing the use of that ontology by other organisations working in that domain (if, that is, the value of retaining the ontology as a trade secret does not exceed the potential licensing income). The costs, in such a networked environment, will come down for everybody too over the period of use; if a single firm took on the costs of developing an ontology for a domain (and licensed that ontology to its competitors, thereby avoiding the reduplication of effort over the marketplace as a whole), that firm could also take on, individually, the maintenance costs. Organisations that specialised in ontology maintenance and training for users could spring up, given sufficient demand for their services.

Increasing Benefit: Similarly to data restructuring; there has to be some discernible benefit for organisations putting their data in RDF, and these benefits will become more

apparent the more published data in RDF there is. So the issue here, which a killer app might help with, is that there seems to be little or no advantage for an individual firm in moving first. A firm that publishes its data in RDF early incurs costs early and takes a risk, but gets little benefit; if it removes its data from its traditional information management stream, it might even reduce the value of that data. A firm that publishes RDF late incurs the costs later, take on little risk and gets more benefits. Nevertheless, being first in a new market is a distinct advantage [16][11], but late entrants can also succeed if they outperform existing services [19] (e.g. Google). So there is a Prisoners Dilemma to be sorted out.

Berners-Lee argues that the killer app for the SW is the *integration* [4]. Once distributed data sources are integrated, the sky becomes the limit. This of course could be true, but it will be hard to convince data providers to publish in RDF and join the SW movement without concrete examples of how they could benefit from doing so. This is probably supported by Berners-Lee suggestion that we need to "Justify on short/medium term gain, not network effect" [5]. The network effect is a long-term goal, and to achieve this goal we need to show short-term gains. Integration alone might not be seen as a gain on its own, especially when considering costs and privacy issues. On the other hand, one target application area for the SW is in adding value to the large amount of relational data sitting around in isolation in databases across organisations or sectors. Being able to manipulate data from heterogeneous sources, in effect cross-referencing data from different databases and being able to infer information about objects represented in different ways (or even to determine when what are apparently two objects are identical), are very powerful capacities.

In a survey for business use cases for the SW, researchers of the EU *KnowledgeWeb*[3] emphasised the importance of proper targeting for SW tasks [33] to avoid applying SW technologies to where they do not offer any clear benefit, which may discourage industry-wide adoption.

The survey concluded that the areas which seem to benefit more from this sort of technology are *data integration* and *semantic search*. It was argued that these areas could be accommodated with technologies for *knowledge extraction*, *ontology mapping* and *ontology development*. Similarly, Uschold and Gruninger [40] argue that ontologies are useful for better *information access*, *knowledge reuse*, *semantic-search*, and *interoperability*. They also list a number of assumptions to be made to progress towards a fully automated semantic integration of independent data sources. Fensel et al [20] describe the beneficial role of ontologies in general *knowledge management* and *eCommerce* applications. They also list a number of obstacles that need to be overcome to achieve those benefits, such as *scalable ontology mapping*, *instantiation*, and *version control*. Other obstacles, such as *trust*, *agent co-ordination*, *referential integrity*, and *robust reasoning* have also been discussed [26].

2.2 Leveraging Metcalfes Law

The relevance of Metcalfes Law, that the utility of a network is proportional to the square of the number of its members [21], is clear in the context of this examination

---

[3] `http://knowledgeweb.semanticweb.org/`

of the nature of the costs; it is often cited in other contexts as an explanatory variable for killer apps [16][27][19]. There are two stages to the process of growing a network; first get the networks growth accelerating, and second preserve the network once it is in place, in order to create a *community of practice*. The economics of network effects show that a network of users of some technology or other good has three equilibrium positions. The first is where the number of users is zero. The second is a point where the network is satiated. Any further growth of the network will merely drag in users whose use of it is marginal, while any diminution will be felt as a loss of opportunity by the defecting network members. These are stable equilibria. But between these points is an unstable equilibrium, where users and costs, demand and supply, are balanced. However an increase in the size of the network, even a small one, will increase the benefits to users dramatically and bring new users in, while a diminution in the size of the network, even a small one, will decrease the benefits to network members dramatically and drive existing users away. It is essential for any network to get beyond the unstable equilibrium in the middle, to promote growth [41].

Communities of practice: A community of practice [43] is a group of people sharing some work- or leisure-related practice; they need not all belong to the same organisation, but they do need some congruence of role. The community of practice that springs up around such a practice acts as a kind of support network for practitioners. It provides a language (or informal ontology) for people to communicate with, a corporate memory, and a means of spreading best practice. Such informal communities are self-selecting.

This self-selection, and informality, makes a community very hard to develop, because the community is a second-order development. So, we might take the example of Friend of a Friend (*FOAF*)[4]. FOAF is a basic ontology that allows a user to express simple personal information (email, address, name, etc) as well as information about people they know. Many SW enthusiasts considered FOAF to be *cool* and *fun* and started publishing their FOAF ontologies. Currently there are millions of FOAF RDF triples scattered over the Web, perhaps far more than any other type of SW annotations.

Social Network Applications: Surprisingly, there exist many Web applications that allow users to represent networks of friendships, such as Friendster[5], Okrut[6], LinkedIn[7], TheFacebook[8], SongBuddy[9], to name just a few. However, FOAF has simply become a more convenient format for representing, publishing, and sharing this sort of information. Even though none of the applications above are entirely based on FOAF, some of them have already begun reading and exporting their data in FOAF format. FOAF is certainly helping spread RDF, albeit in a way limited to part of the SW community, and could therefore be regarded as a facilitator or a medium for possible killer apps that could make use of available FOAF files and provide some useful service.

---

[4] http://www.foaf-project.org/

[5] http://www.friendster.com/

[6] http://www.orkut.com/

[7] https://www.linkedin.com/

[8] http://www.thefacebook.com/

[9] http://www.songbuddy.com/

Sustaining Network Growth: However, one interesting obstacle in the way of FOAF creating the nexus of users that will launch the SW is that a network is generally self-selecting and second order. A network is a network *for* something. There is absolutely nothing wrong with the FOAF method of providing a format for describing oneself. One obvious benefit of FOAF is that, as a pretty simple and flat ontology, it provides a relatively painless way of ascending the learning curve for non-users of SW technology. However, to sustain the network growth, there will still need for something underlying such networks, some practice, shared goal, or other practical purpose. Furthermore, it is also important to ascertain the invariants of the community or network experience, and ensure that the technologies and practices that produce growth preserve those invariants [7].

It may well be that a potentially more fruitful approach would be to support existing communities and try to expand SW use within them, so that little Semantic Webs emerge from them, as SW technologies and techniques reach saturation point within them [26]. And because communities of practice overlap, and converge on various boundary objects and other linking practices and artefacts [43], it may be that SW practice might spread beyond these islands of best SW practice. There are many obvious aids to such a development strategy; for instance, good-quality ontologies could be hand-crafted for particular domains. But also, it turns out that a number of the best SW tools at the moment also support this "filling out" technique. For instance CS AKTive Space [37] specifically enables people to find out about the state of the discipline of computer science in the UK context a limited but useful domain. Flink [32] generates FOAF networks for SW researchers. CS AKTive Space and Flink are winners of the 2003 and 2004 Semantic Web Challenges respectively.

Open systems and social aspects: One other useful aspect of Flink is that it integrates FOAF profiles with ordinary HTML pages, and therefore sets up an explicit link between the SW and the WWW. Direct interaction with other existing systems increases the value of a system by acquiring additional value from those systems [27]. One good example is Protégé; an ontology editor from Stanford [34]. By being open source and extendable, Protégé allowed many existing systems and tools to be linked or integrated with it, thus increasing its use and value. For this reason, and for being free, Protégé has quickly become one of the most popular ontology editing tools available today.

This openness is of course built into the very conception of the SW; the integration of large quantities of data, and the possibility of inference across them, is where much of the power stems from. As with the WWW, this does require a major programme of voluntary publication; for example, the possibility of e-commerce, and of the use of the WWW as a marketplace or as a version of the High Street stems from being able simply and conveniently to compare prices across retailers. The SW would add value (or reduce information processing costs) still more by allowing agents to do the same thing and more [25].

And as with the WWW, if this process takes off, then more and more vendors would have to publish their data in RDF, even if they are initially reluctant. The argument in favour of such coercion is that everyone benefits eventually, and that early movers (those who make data such as price lists or stock lists available) not only gain, but force laggards to follow suit.

Privacy and Trust: However, transparency and the removal of restrictions to publication are not undiluted goods. It may be that certain pieces of information benefit some organisations only as long as they withhold them from public view (trade secrets). Or that issues such as privacy and anonymity will rear their heads here. Or even that differing intellectual property regimes and practices will lead to competitive advantage being lost in some economies.

In particular, integrating large quantities of information across the Internet and reasoning across them raises potential problems from a number of directions [7]. Firstly, it is the integration of information that threatens to allow harmful inference; information is quite often only harmful when seen in the right (or, rather, wrong) context. The increasing numbers of mashups are potentially very dangerous here, as there is no ownership or chain of responsibility for information and its use. And the SW is the tool *par excellence* for putting previously harmless data in sensitive contexts. And secondly, publication of information (for instance, FOAF information) in a friendly and local context can quickly get out of ones control. There have been many warnings already about the placing of information on personal yet public spaces such as MySpace; there is anecdotal evidence that interviewees are being Googled to discover aspects of their private lives, to supplement the information about their public faces as presented on their CVs. Again, the SW is an important technology in this context.

It is often argued that standard data protection legislation is adequate for the new online contexts, but that policing is the problem. As it stands, traditional restrictions on the gathering of information are becoming decreasingly relevant as information crosses borders so easily; more plausible is policing restrictions on how information can be used once collected. The SW will require the development of formalisms that promote transparency of information use and accountability of information users. It should be possible to state one's preferences about the use of data in which one features (who, for example, can see the data, who can use it, and who can pass it on — and to whom). Sometimes it will be necessary to override privacy preferences (for example, for the purposes of criminal investigation). It should be possible to receive metadata about how one's data have been used. And if there is a problem, and one's policies have been transgressed, it should be possible to hold someone to account. These requirements demand both social institutions and technical innovation. Various initiatives of the World Wide Web Consortium (W3C) address some aspects of these problems. The Platform for Privacy Preferences (P3P) allows different agents a common view of privacy preferences, and is aimed at enhancing user control of data by simplifying the expression of privacy preferences. However, P3P contains no enforcement methods if preferences are violated [15]. The Policy-Aware Web initiative tries to provide rule-based formalisms to express preferences and increase transparency and accountability in the open and distributed environment of the WWW; SW technology can be important here, by allowing rules and even proofs based on those rules to be exchanged between agents [42].

Furthermore, formalising or externalising knowledge, for example in the creation of ontologies, can have a number of effects. First of all, knowledge that is codified can become more 'leaky', i.e. it is more likely to leave an organisation. Secondly, it will tend to reduce the competitive advantage, and therefore income, of certain experts. Thirdly, much depends on whether a consensus exists about the knowledge in the first place.

As [9] argues, in a complex world where politics renders certain ideas untenable, incomplete knowledge requires whole perspectives to be understood, expertise must be combined in the light of discussion and argument, and information must be interpreted on the fly, the consensual nature of ontologies (agreed conceptualisations of the world) can be unrealistic to insist on. The more lightweight the ontology, the easier it will be to adapt different viewpoints to it. Wilks argues [44] that there are strong links between formal ontologies and the natural language terms which they borrow, and as with natural languages, insisting on a formal and verifiable common understanding of shared vocabulary (or isomorphic structures in formal ontologies) is too strong a constraint. Ontologies have to be flexible enough to allow interaction to take place.

Such flexibility, which will promote ease of use and therefore also help foster a user community, does mean that certain issues, such as referential integrity, will need to be faced [8], with some sufficiently lightweight methods for dealing with such problems of ambiguity as they arise.

2.3   Creativity and risk

Killer app development cannot follow from careful planning alone. As we have noted already, there is no algorithm for creating a killer app. They tend to emerge from simple and inventive ideas; they get much of their transformative power by destroying hitherto reliable income streams for established firms. Christensen [11] points out that most killer apps are developed by small teams and start-ups. Examples include Google, eBay, and Amazon, which were all created by few dedicated individuals. Giant industrial firms are normally reluctant to support risky projects, because they are generally the ones profiting from the very income streams that are at risk [16].

However, even though most semantic web applications have so far been built in research labs and small groups and companies, there is clear interest expressed by the big players as well. So, for example, Hewlett Packard has produced Jena [10], a Java library for building SW applications, and IBM has developed WebFountain [23], a heavy platform for large scale analysis of textual web documents using natural language analysis and taxonomies for annotations. Adobe has perhaps gone further than many; Acrobat v5 now allows users to embed RDF metadata within their PDF documents and to add annotations from within Web browsers which can be stored and shared via document servers.

The SW provides a context for killer app development, a context based on the ability to integrate information from a wide variety of sources and interrogate it. This creates a number of aspects for the potential for killer apps. First of all, SW technologies might essentially be expected to enable the retrieval of data in a more efficient way that possible with the current WWW which is often seen as a large chaotic library. In such a world, the killer app might be some kind of retrieval technique to supplant Google, and therefore must do it without extra cost, without painful effort, and it should be the case that enough major players realise that Google is failing in some respect.

On the other hand, it may be that the SW might take off in an original and unpredictable direction. The clean infrastructure that the W3C ensures is in place could act as a platform for imaginative methods of collating and sifting through the giant quantities of information that is becoming available. This might result in a move away from

the webpage paradigm, away from the distinction between content providers and consumers, as for example with efforts like CS AKTive Space [37], or a move towards a giant, relatively uniform knowledge base (of the CYC variety) that could cope with all those complexities of context that foiled traditional AI approaches [2]. The ultimate vision of the SW that prevails should affect not only the standards developed for it [30], but also where we might look for killer apps.

2.4 Personalisation

Personalisation has been a common thread in the development of killer apps. Customers tend to become more loyal to services they can customise to their liking [16]. Many of today's killer apps have some level of personalisation; Amazon for example makes recommendations based on what the customer buys or looks at; Auto Trader[10] and Rightmove[11] save customers' searches and notify them via emails when a new result to their query is available; personalised web services attract more customers (if done properly!) and provide better tailored services [19].

Personalisation is often the key to providing the higher service quality than the opposition; the service itself need not be provided in any better ways, but the personalised aspect gives it the extra that is needed to see off the alternatives. Such a connection could be indirect; for example, an Amazon-style recommender system, linked with an advertising platform, could help find alternative revenue streams and therefore drive down the cost to the consumer. The appearance of consumer choice in particular has two effects. First of all, the consumer may actually do a lot of the work for the service provider him- or herself; for example, by negotiating a series of yes/no questions to an automated call centre, or by diligently scrolling through a set of FAQs, the consumer may well diagnose his/her own problems. And secondly, the consumers preferences get met so much more quickly [29].

It goes without saying that personalisation is a hot topic on the SW, as well-annotated knowledge sources can be matched against RDF statements about individual consumers, to create recommendations or targeted products. Indeed, the individual need not even do this voluntarily; as more information becomes available, for example through the FOAF network, or other RDF from third parties, users can become very known quantities indeed (raising all the privacy concerns we have noted elsewhere). There may well be major advantages to be had in systems that can feed information discreetly into recommender systems [35][14].

Nevertheless, it is the personalisation aspect that has much potential for the SW, as long as the provision of enough information for the system to work interestingly is not too painful.

2.5 Semantic Web Applications

There have been few sustained attempts to try to promote SW applications. For instance, the important work of the W3C naturally is focused on the standards that will create

---

[10] http://www.autotrader.co.uk
[11] http://www.rightmove.co.uk

the clean platform that is a necessary but sadly not sufficient condition for the SW to take off. But one of the most interesting and inspired is the series of Semantic Web challenges, which we will discuss briefly in the next section.

3   The Semantic Web Challenge

The annual International Semantic Web Challenge[12](SWC), has been a deserved success, sparking interest and not a little excitement. It has also served to focus the community. Applications should illustrate the possibilities of the Semantic Web. The applications should integrate, combine, and deduce information from various sources to assist users in performing specific tasks. Of course, to the extent that it does focus the community, the SWC will naturally influence the development of the SW.

Submissions to the SWC have to meet a number of minimum requirements, viz:

– First, the information sources used
  • should be geographically distributed,
  • should have diverse ownerships (i.e. there is no control of evolution)
  • should be heterogeneous (syntactically, structurally, and semantically), and
  • should contain real world data, i.e. are more than toy examples.
– Second, it is required that all applications assume an open world, i.e. assume that the information is never complete.
– Finally, the applications should use some formal description of the meaning of the data.

Secondly, there are desiderata that act as tiebreakers.

– The application uses data sources for other purposes or in another way than originally intended
– Using the contents of multi-media documents
– Accessibility in multiple languages
– Accessibility via devices other than the PC
– Other applications than pure information retrieval
– Combination of static and dynamic knowledge (e.g. combination of static ontologies and dynamic work-flows)
– The results should be as accurate as possible (e.g. use a ranking of results according to validity)
– The application should be scalable (in terms of the amount of data used and in terms of distributed components working together)

In the light of our discussion above, these are interesting criteria. Many of them are straightforwardly aimed at ensuring that the characteristic possibilities of the SW are realised in the applications.

---

[12] http://challenge.semanticweb.org/, where the criteria quoted below are to be found.

3.1 Semantic Web Challenge winners

Let us have a close look at the winners of the SWC over the past three years: CS AK-Tive Space (CAS), 2003 winner; Flink, 2004 winner; and ConFoto, 2005 winner. These applications were selected among 35 submissions over the past three years[13] and represent the best applications of Semantic Web applications in terms of SWC participant applications.

2003 winner: CS AKTive Space (CAS) In 2003, the first SWC took place, and a new style of exploring domain information by means of collecting and processing Semantic Web data won the first prize: CS AKTive Space (CAS)[14]. CAS is a smart browser interface for a Semantic Web application that provides ontologically motivated information about the UK computer science research community. The scenario used is based on a real-world community request: the desire for a funding council to be able to get a fast overview of the council's domain from multiple perspectives. This requires bringing together data from heterogeneous sources and constructing methods to present the possible relations in the data to a user quickly and effectively. CAS developers report that this scenario applies equally to any stakeholder of the domain. While a funding body may wish to know, for instance, how funding in an area is distributed geographically, a researcher may also wish to know who is working on which topic, where are they located, and who is in the community of practice for the top researchers in their area [22].

CAS provides multiple ways to look at and discover simple information or rich relations within the Computer Science domain of the UK. It facilitates querying, exploring and organizing information in ways that are meaningful to the users: where one user or group may be interested in seeing the relationship between funding, research area and geographical region, another may be interested in who the top researchers are in AI and their addres data. With CAS, people can formulate and see at a glance rich results like these, without having to string together large, complex queries.

CAS explores a wide range of semantically heterogeneous and distributed content relating to Computer Science research in the UK. There are almost 2000 research active Computer Science faculty, there are 24,000 research projects represented, many thousands of papers, and hundreds of distinct research groups. These entities are described by a number of existing sources, such as institutional information systems (university web sites, research council databases), bibliographic services and other third party data sets (geographical gazetteers, UK Research Assessment Exercise submissions) [37]. This content is gathered on a continuous basis using a variety of methods including harvesting and scraping of publically available data from institutional web sites, bulk translation from existing databases, and direct submissions by partner organisations, as well as other models for content acquisition. CAS supports both regularly scheduled harvesting to identify and deal with changes to existing data sources, and on-demand harvesting in response to changing user requirements [12] or update notifications from component sources.

---

[13] A full list of submissions is available at http://www.cs.vu.nl/ pmika/swc/submissions.html

[14] Demonstration available online at: http://triplestore.aktors.org/SemanticWebChallenge/
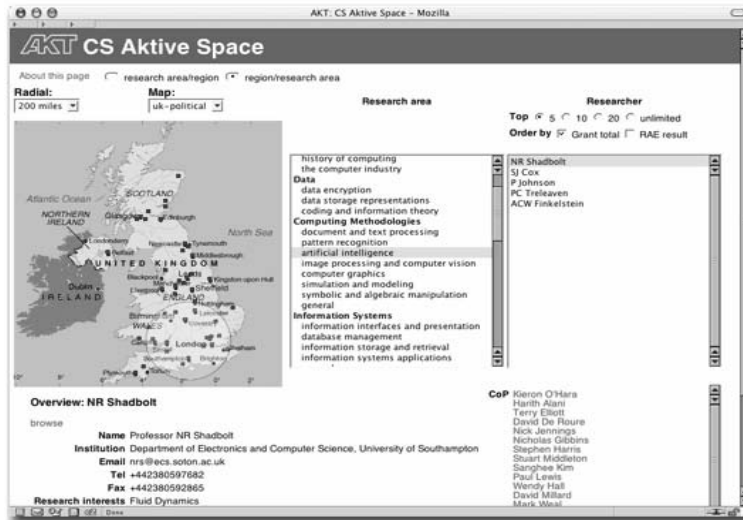
Fig. 1. Computer Science AKTive Space - 2003 SWC winner.

The content is mediated through an OWL ontology[15] which was constructed for the application domain and which incorporates components from other published ontologies. The content currently comprises around twenty five million RDF triples, and CAS developers provide scalable storage and retrieval technologies and maintenance methods to support its management [24]. As a user interacts with the system, the interface generates queries (expressed in RDQL) which are evaluated by the triplestore server. The strength of CAS, "lies in the separation of users from the activity 'under the hood', which enables users to answer questions that they might not be able to phrase." [22]. CAS also uses the OntoCoPI service [1] to get information about an individual's community of practice. OntoCoPI performs a spreading activation search, with weights assigned to the relationships in the ontology, in order to determine the ranked list of people with whom an individual associates. For example, the notion of a community of practice is calculated from a given researcher's coauthors, the projects that they are involved with, the institutions with which they are affiliated and the topics in which they conduct research.

2004 winner: Flink In 2004, the Flink system[16] won the SWC. Flink extracts, aggregates and visualizes online social networks [32]. It employs semantic technology for reasoning with personal information extracted from a number of electronic information sources including web pages, emails, publication archives and FOAF profiles. The acquired knowledge is used for the purposes of social network analysis and for generating a web-based presentation of the community. Flink is thus intended as a portal

---

[15] Available online at: http://www.aktors.org/ontology/

[16] Demonstration available online at: http://prauw.cs.vu.nl:8080/flink/

for anyone who is interested to learn about the work of the Semantic Web community, as represented by the profiles, emails, publications and statistics. The data collected by Flink can also be used for the purpose of social network analysis, in particular learning about the nature of power and innovativeness in scientific communities.

Flink is presented as a "who is who" of the Semantic Web by means of presenting professional work and social connectivity of Semantic Web researchers. Flink defines the community of Semantic Web researchers as those researchers who have submitted publications or held an organizing role at four international conferences related to the Semantic Web. That data set represents a community of 608 researchers from both academia and industry, covering much of the United States, Europe and to lesser degree Japan and Australia.
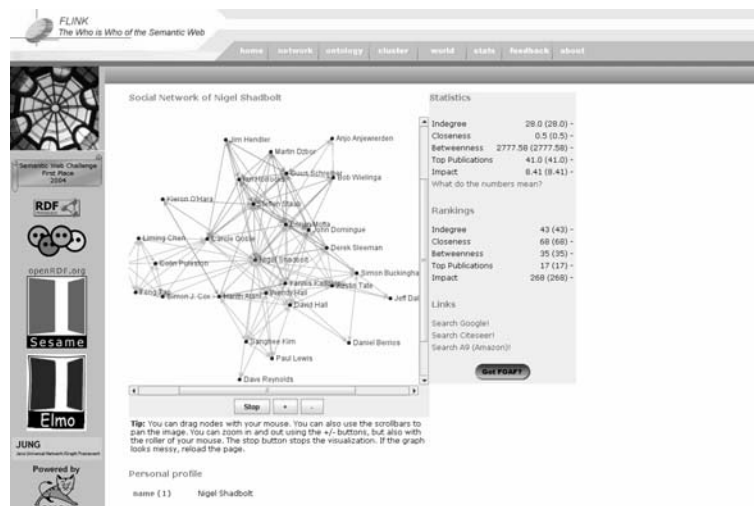

Fig. 2. Flink - 2004 SWC winner.

Flink takes a network perspective on the Semantic Web community, which means that the navigation of the website is organized around the social network of researchers. Once the user has selected a starting point for the navigation, the system returns a summary page of the selected researcher, which includes profile information as well as links to other researchers that the given person might know. The immediate neighbourhood of the social network (the ego-network of the researcher) is also visualized in a graphical form. The profile information and the social network is based on the analysis of webpages, emails, and publications. The displayed profile information includes the name, email, homepage, image, affiliation and geographic location of the researcher, as well as interests, participation at Semantic Web related conferences, emails sent to public mailing lists and publications written on the topic of the Semantic Web. The full text of emails and publications can be accessed by following external links. At the time of the SWC contest, Flink contained information for about 5147 publications authored

by members of the community and 8185 messages sent via five Semantic Web-related mailing lists.

The navigation from a profile can also proceed by clicking on the names of co-authors, addressees or others listed as known by this researcher. In this case, a separate page shows a summary of the relationship between the two researchers, in particular the evidence that the system has collected about the existence of this relationship. This includes the weight of the link, the physical distance, friends, interests and depictions in common as well as emails sent between the researchers and publications written together.

The information about the interests of researchers is also used to generate an ontology of the Semantic Web community. The concepts of this ontology are research topics, while the associations between the topics are based on the number of researchers who have an interest in the given pair of topics. An interesting feature of this ontology is that the associations created are specific to the community of researchers whose names are used in the experiment. This means that unlike similar lightweight ontologies created from a statistical analysis of generic web content, this ontology reflects the specific conceptualisations of the community that was used in the extraction process. Also, the ontology naturally evolves as the relationships between research topics changes

2005 winner: CONFOTO CONFOTO[17] won the 2005 SWC. CONFOTO is a browsing and annotation service for conference photos. It combines recent Web trends (tag-based categorization, interactive user interfaces, syndication) with the advantages of Semantic Web platforms (machine-understandable information, an extensible data model, the possibility to mix arbitrary RDF vocabularies). CONFOTO offers a variety of tools to annotate and browse pictures. Simple forms can be used to create multilingual titles, tags, or descriptions, while more advanced forms allow the relation of pictures to events, people, ratings, or copyright information. CONFOTO provides a tailored photo browser and gallery generator for pictures, and a generic RDF browser for other resource types. Although a central repository is used to store resource descriptions, it is not necessary to copy photo files to the server: The application supports uploaded pictures as well as pictures linked via a URL or described in external RDF/XML documents. RSS export functions facilitate photo sharing, and a SPARQL interface enables and encourages extended data re-use.

3.2   SWC and Killer Apps

Of particular interest to our work is the relation of the SWC criteria and the literature on killer apps. What the SWC is intended to uncover are new ways of exploiting information, particularly distributed information, and demonstrating the power of interrogation. In this respect, the challenge can only raise the profile of the SW, and help extend its community to more people and organisations. The SWC is an excellent vehicle for demonstrating where added value may come from. And as we have seen, increasing the size of the network will bring with it exponentially-increasing benefits.

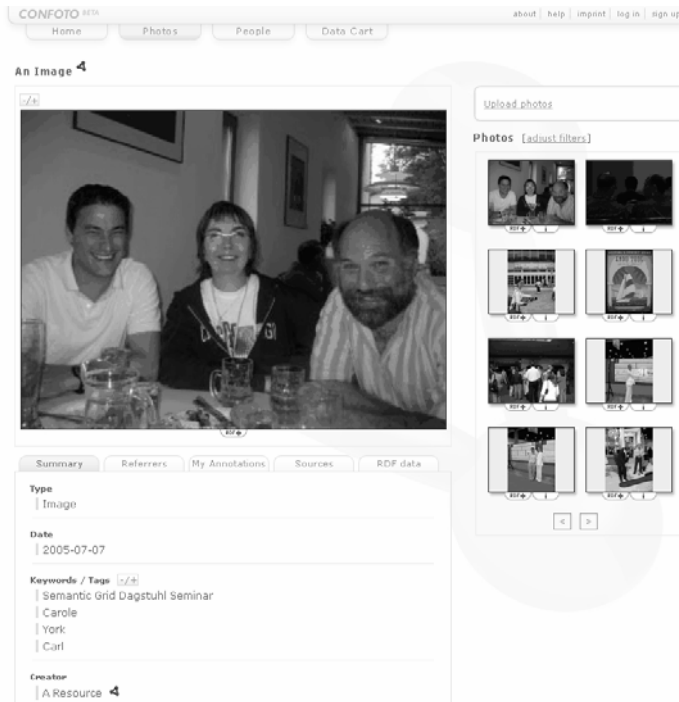[17] Demonstration available from http://www.confoto.org/home

Fig. 3. Confoto - 2005 SWC winner.

But the SWC looks unlikely to furnish us with a (prototype of a) killer app, because the criteria focus on interesting results, rather than on usability, superiority or the alteration of old habits. Some of the criteria are slightly double-edged. For example, it is essential that an application uses a formal description of the meaning of the data. This, of course, is a deliberate attempt to ensure that one of the most contentious aspects of the SW (one of the most commonly-cited causes of scepticism about the SWs prospects) is incorporated. That is, the use of ontologies to provide understanding of terms in knowledge from heterogeneous sources. However, the way the challenge is constructed means that what is bound to happen in many if not most applications is that the developers will create their system with a possibly very painstakingly constructed ontology in mind, rather than taking the more difficult option of employing a very lightweight system that could work with arbitrary ontologies. This is particularly true for 2003 and 2004 SWC winner applications whereas the 2005 winner relies less on a rich ontological substratum but the application does not make use of rich ontological terms. The situation is somewhat similar to the knowledge engineering Sisyphus challenges, where knowledge engineering methodologies were tested and compared by being applied to the same problem. However, as the methodologies were applied by their developers, the results were less than enlightening; a later attempt to try to measure how difficult

methodologies were to use by non-specialists suffered from an unwillingness of most developers to discover this key fact about their methods [38].

There is little here to create the genuine community (as opposed to a large network); to promote the idea that users have something of a responsibility not to free ride, and to publish RDF data. Neither is there much to promote personalisation within that community. There is little to protect privacy, little to reduce the pain of annotating legacy data or building ontologies, and, although the focus of the SWC is the results of information-processing, little to ensure that such processing can integrate into the organisational workflow. Surprisingly few of the traditional requirements, from a business perspective, appear in the SWC criteria.

Figure 4 summarises how the SWC winners fit into some of the general features of killer apps that we discussed earlier. It is interesting to observe the similarity between the first two applications which both targeted the same domain and community. CSAktiveSpace however focused on gathering information about where researchers are located and what areas they research, with some social network analysis based on addresses, research areas, projects, as well as co-authorship relations. Flink on the other hand is mainly for social network analysis based almost entirely of co-authorship information. The third winner, Confoto, allows for user networks to be formed and hence, in theory, it has the potential of generating a network value. Having said that, Confoto currently stand hopeless against giant rivals that proved extremely popular in the general WWW community, such as Flickr, although the later is probably based on less advanced technology. Generally speaking, all three winners are specifically designed and built for a relatively small community (i.e. computer scientists or academics in general) which inherently restricts their number of potential beneficiaries and limits their potential spread.

None of this, let us hastily add, is intended as a criticism of the SWC, which has publicised the SW and drawn a lot of attention to the extra power that it can provide. Our point is merely that there is a lot more to finding a killer app than producing an application that does brilliant things.

4 Discussion

Killer apps must provide a higher service quality, and evolve (pretty quickly) into something perceived as indispensable, conferring benefits on their users without extra costs or steep learning curves. Individual users should coalesce into a community of practice, and their old habits should change in accordance with the new possibilities provided by the app. This is particularly important as the SW is likely to impose new costs on users in the short term, for example through having to annotate legacy content, develop ontologies, etc. As the SW is in a relatively early stage of development, it is not currently clear exactly what threats and opportunities it provides (and, of course, the future form of the SW will conversely depend on what applications for it are successful). There has been some speculation about how the SW will develop, and what extensions of the WWW will be appropriate or desirable. For instance, consider a recent attempt by Marshall and Shipman to understand potential development routes for the SW [30], which sets out three distinct but related visions of the SW.

Application	Cost	Added Value	Community	Open Standard	Personalisation	Bootstrapping
CSAktiveSpace	Main cost is in data gathering, cleaning, and merging Simple browsing skills required for using existing service Use is free	Provides knowledge about researchers that is not easily obtainable from elsewhere But current system is only limited to Computer Science domain	Service targets a specific community No support for user interactions Metcalfe's network value can not be generated from the system	Based on RDF and knowledge base technology No API for external access, import/export, or for adding extensions Not open source	Users can not change or add to the data held in the system No user specific features are provided	System is setup with data from various sources Various techniques are used to collect the data No dependent on number of users
Flink	Main cost is in gathering and preparing the data A learning curve to understand the system output Use is free	Produces graph network profiles for researchers based on co-authorship information Limited to specific bibliographic sources for Semantic Web researchers	Service targets a specific community No support for user interactions Metcalfe's network value can not be generated from the system	Based on RDF and knowledge base technology No API for external access, import/export, or for adding extensions Some is open source	Users can not change or add to the data held in the system No user specific features are provided	System is setup with data from various sources Various techniques are used to collect the data No dependent on number of users
Confoto	Main cost is in developing the system Data gathering cost is minimal Use is free	Provides conference attendees with a semantically enriched photo annotation and sharing tool General field is currently dominated by Flickr	Users can interact by annotating each other's photos, profiles, and events Can generate a network value	Based on RDF and XML Users can import/export data No API for external access Not open source	Some non-persistent personalisation is provided Users can added to the data held in the system	System needs to be bootstrapped with user data (photos) Entirely dependent on user input

Fig. 4. Some general characteristics of Semantic Web Challenge winners

I *SW technology could bring an order and consistent structure to the chaotic Web*. So information access would be assisted by semantic metadata. This vision envisages that humans will continue to be the chief agents on the SW, but that information could now be represented and stored in ways to allow its use in situations far beyond those foreseen by its original authors. In other words, the SW will extend the existing Web, but exactly how is hard to predict. In order for potential SW Killer Apps to respond to this vision should, ideally, have the following properties in particular:

- They should help foster communities of users (that is, at some level, users should want to interact, and share experiences, with others). SW technology is expected to facilitate knowledge sharing and bringing more people together.
- Users should not feel submerged in a mass, but should retain their individuality with personalised products. With more machine readable information becoming available (eg FOAF), better personalisation should be feasible.
- Users should be able to bring as much of their legacy content up to date with relatively painless maintenance techniques. So, for example, applications should be able to leverage comparatively simple ontologies; ontology construction, merging and selection should be made easier, and we should be able to move away from handcrafting; annotation methods and interfaces have to be easy. In all these cases, the existence of a community of interested users will provide the initial impetus for the user to ascend the learning curve.

II *The Web will be turned, in effect, into a globally distributed knowledge base, allowing software agents to collect and reason with information and assist people with common tasks, such as planning holidays or organising diaries*. This vision seems close to Berners-Lee et als SW grand vision [4]. In many ways it is the composition of the other visions, assuming machine processing and global representation of knowledge. It is also a vision that will require more from a potential Killer App:

- They should exploit integrated information systems to make inferences that could not be made before, or to bring sources of information (for example, data held in relational databases) together in the context of more powerful information processing tools. Showing added value is key to encourage businesses and content providers to participate in the SW.
- They should help remove the rather painful need to annotate, build ontologies, etc.
- The new application should fit relatively smoothly into current work or leisure experiences. Little change in habits is acceptable, assuming some returned benefit, but too much change is a problem!

III *The SW will be an infrastructure, made up of representation languages, communication protocols, access controls and authentication services, for coordinated sharing of knowledge across particular domain-oriented applications*. Information is used largely for the original purposes of its author, but that much more machine processing will take place. If this vision prevails, then a prospective SW Killer app should pay special attention to the following:

- It should not compromise other important aspects of users lives, for instance by threatening privacy to a dangerous degree, either by making inappropriate surveillance possible, or by facilitating torts such as identity theft.

- Furthermore, if such a vision comes to pass, then the opportunities for killers apps are all the greater, in that any such standards-driven platform approach should make it possible for as many applications to flourish on top of it as possible. Whether such applications will ever be acknowledged as SW apps rather than general WWW applications is open to debate!

These conditions for each vision of the SW are, of course, necessary yet not sufficient! Furthermore, all of the conditions apply, to some degree, to each of the three visions.

We have seen that killer apps appear when there are opportunities to make progress on costs, communities, creativity and personalisation. All new technologies begin with a handicap in these areas. They impose costs, of retooling, of learning curves and of business process rescheduling. There is always a chicken and egg problem with the development of a community of users the technology of necessity precedes the community. The risks of creative thought become clearer at the outset; the benefits only appear later. And the dialectic between personalisation and creating economies of scale often means that the latter are pursued long before the former. As an added handicap, it is often the case that the costs are borne disproportionately by early adopters. The opportunities of the SW are also therefore counterbalanced by the risks. We note that we cannot predict where new killers will come from. The transformations that such applications wreak make the future very different from the present. Hence we cant be concretely prescriptive. But the general requirements for killer apps that emerge from our review of the business/management literature suggest certain routes for development, in addition to sensible lists of characteristics such as the criteria for the SW challenges, or the conditions listed at the beginning of this section.

So it is probably uncontroversial to assume that any SW killers will have to provide (1) a service that is not possible or practical under more traditional technologies, (2) some clear benefit to developers, data providers, and end users with minimum extra costs, and (3) an application that becomes indispensable to a user-base much wider than the SW researchers community. But additionally, research should be focusing on four important areas. First of all, perhaps most important, the cost issue should be addressed. Either the potentially large costs of annotating, ontology development, etc, should be mitigated, or side-stepped by thinking of types of application that can work with minimal non-automatic annotation, low cognitive overhead, or ontologies that sacrifice expressivity for simplicity. Secondly, another way of improving the cost/benefit ratio is to increase benefits, in which case the fostering of user communities looks like a sensible way forward. This means that applications in real-world domains (preferably in areas where the Internet was already important, such as media/leisure or e-science) look more beneficial than generic approaches. Thirdly, creativity is important, so radical business models are more interesting than simply redoing what the WWW already does. And fourthly, personalisation needs to be addressed, which means that extended user models are required.

When we look at the three visions outlined by Marshall and Shipman, vision III appears to be the one most amenable to the development of killer apps, in that it envisages a platform upon which applications sit the form of such applications is left relatively open. In contrast, vision I, for example, does not see too much of a change for

the WWW and the way it is used, and so there are fewer opportunities opening up as a result. Indeed, when we look at vision I, assuming that the SW does improve the navigation of the chaotic web, it might even be appropriate to say, not that there is a killer app for the SW, but rather that the SW is the killer app for the WWW. Whereas, with Marshall and Shipmans vision III, the vision is of a garden in which a thousand flowers bloom. Applications can be tailored to context; ontologies can remain as lightweight as possible, and data can be exploited as it is required. Publication of data can be tempered, thanks to privacy-enhancing tools, by supporting restrictions on its use. On this vision, it is the painstaking, pioneering and often tedious negotiations of standards that will be key; such standards need to support the right kind of research.

5   Conclusions

Killer apps are very difficult things to monitor. They are hard to describe, yet you know one when you see one. If you are finding difficulty persuading someone that something is a killer app, it probably is not! A lot depends on the bootstrapping problem for the SW – if the SW community is small then the chances of someone coming up with a use of SW technology that creates a genuinely new use for or way of producing information are correspondingly small. For it is finding the novelty that is half the battle. There is unlikely to be much mileage in simply reproducing the ability to do something that is already possible without the SW. Furthermore, it is likely that a killer app for the SW will exploit SW technology integrally; merely using RDF will not quite do the trick [39]. The willingness of the producers of already-existing killers to use SW technology, like Adobe, is encouraging, but again will not necessarily provide the killer app for the SW. The checklist of criteria for the SWC gives a good list of the essentials for a genuinely SW application.

This is not simply a matter of terminology. The SW is more than likely to thrive in certain restricted domains where information processing is important and expensive. But the ambitions of its pioneers, rightly, go beyond that. For that to happen, killer apps need to happen. We hope we have given some indication, via our examination of the business literature, of where we should be looking.

Acknowledgments

References

1. Alani, H., Dasmahapatra, S., O'Hara, K. and Shadbolt, N. (2003) 'Identifying Communities of Practice through Ontology Network Analysis', *IEEE Intelligent Systems*, Vol. 18, No. 2, pp. 18–25.
2. Berners-Lee, T. (1998) 'The semantic web road map', http://www.w3.org/DesignIssues/Semantic.html
3. Berners-Lee, T., Hendler, J., and Lassila, O. (2001) 'The semantic web', *Scientific American*, May.
4. Berners-Lee, T. (2003) 'Semantic Web: Where to direct our energy?', *Keynote speech at the 2nd International Semantic Web Conference (ISWC2003)*, Florida, USA, http://www.w3.org/2003/Talks/1023-iswc-tbl/
5. Berners-Lee, T. (2004) 'WWW 2004 keynote'. *13th Int. World Wide Web Conference*, New York, http://www.w3.org/2004/Talks/0519-tbl-keynote/
6. Berners-Lee, T. (1995-2005) 'The Fractal Nature of the Web', http://www.w3.org/DesignIssues/Fractal.html
7. Berners-Lee, T., Hall, W., Hendler, J.A., O'Hara, K., Shadbolt, N., and Weitzner, D.J. (2006) 'A Framework for Web Science', *Foundations and Trends in Web Science*, Vol. 1, No. 1, pp. 1–129.
8. Bilenko, M., Mooney, R., Cohen, W., Ravikumar, P., and Fienberg, S. (2003) 'Adaptive name matching in information integration', *IEEE Intelligent Systems*, pages 2–9, September/October.
9. Buckingham-Shum, S. (2004) 'Contentious, dynamic, information-sparse domains...and ontologies?', *IEEE Intelligent Systems*, pages 80–81, January/February.
10. Carroll, J.J, Dickinson, I., Dollin, C., Reynolds, D., Seaborne, A., and Wilkinson, K. (2003) 'Jena: Implementing the semantic web recommendations', Technical report, HP Laboratories Bristol, HPL–2003–146.
11. Christensen, C.M. (1997) *The Innovator's Dilemma*, Harvard Business School Press.
12. Ciravegna, F., Dingli, A., Guthrie, D., and Wilks, Y. (2003) 'Integrating Information to Bootstrap Information Extraction from Web Sites', *Proceedings of Workshop on Information Integration, IJCAI'03*, Acapulco, Mexico.
13. Coase, R. (1991) 'The Nature of the Firm', in E.O. Williamson amd S.G. Winter (eds.), *The Nature of the Firm: Origins, Evolution and Development*, Oxford University Press, Oxford. (Original work published in Economica 1937.)
14. Cox, S., Alani, H., Glaser, H., and Harris,S. (2004) 'The semantic web as a semantic soup', *Proceedings of the First Workshop on Friend of a Friend, Social Networking and the Semantic Web*, Galway, Ireland.
15. Cranor, L., Langheinrich, M., Marchiori, M., Presler-Marshall, M., and Reagle, J. (2002) 'The Platform for Privacy Preferences 1.0 (P3P1.0) Specification', World Wide Web Consortium, http://www.w3.org/TR/P3P/
16. Downes, L., and Mui, C. (2000) *Unleashing the Killer App*, Harvard Business School Press.
17. Eischen, K. (2002) 'The social impact of informational production: Software development as an informational practice', CGIRS working paper 2002-1, Center for Global International and Regional Studies, University of California, Santa Cruz.
18. Ellman, J. (2004) 'Corporate ontologies as information interfaces', *IEEE Intelligent Systems*, pages 79–80, January/February.
19. Evans, P. and Wurster, T. (2000) *Blown to Bits*. Harvard Business School Press.
20. Fensel, D., Bussler, C., Ding, Y., Kartseva, V., Klein, M., Korotkiy, M., Omelayenko, B., and Siebes, R. (2002) 'Semantic web application areas', *Proceedings of the Seventh International Workshop on Applications of Natural Language to Information Systems (NLDB 2002)*, Stockholm, Sweden.

21. Gilder, G. (1993) 'Metcalfe's law and legacy' *Forbes ASAP*, 13 September.

22. Glaser, H., Alani, H., Carr, L., Chapman, S., Ciravegna, F., Dingli, A., Gibbins, N., Harris, S., schraefel, m., and Shadbolt, N. (2004) 'CS AKTive Space: Building a Semantic Web Application', *Proceedings of First European Semantic Web Symposium (ESWS 2004)*, Crete, Greece, pp. 417–432.

23. Gruhl, D., Chavet, L., Gibson, D., Meyer, J., Pattanayak, P., Tomkins, A., and Zien, J. (2004) 'How to build a webfountain: An architecture for very large-scale text analytics', *IBM Systems Journal*, Vol. 43, No. 1, pp. 64–76.

24. Harris, S., and Gibbins, N. (2003) '3store: Efficient Bulk RDF Storage', *Proceedings of the ISWC'03 Practicle and Scalable Semantic Systems (PSSS-1)*, Sanibel Island, FL, USA.

25. Hendler, J. (2001) 'Agents and the semantic web', *IEEE Intelligent Systems*, pp. 30–37, march/april.

26. Kalfoglou, Y., Alani, H., Schorlemmeret, M., and Walton, C. (2004) 'On the emergent semantic web and overlooked issues', *Proceedings of the Third Internaitonal Semantic Web Conference (ISWC 2004)*, Hiroshima, Japan.

27. Kelly, K. (1998) *New Rules for the New Economy: 10 Radical Strategies for a Connected World*, Penguin Books.

28. Klein, M., and Fensel, D. (2001) 'Ontology versioning on the semantic web', *Proceedings of the First International Semantic Web Working Symposium*, pp. 75–91, Stanford University, CA, USA.

29. Markillie, P. (2005) 'Crowned at last', *The Economist*, March 31st.

30. Marshall, C., and Shipman, F.M. (2003) 'Which semantic web?', *Proceedings of the 14th HyperText Conference (HT 2003)*, pp. 57–66, Nottingham, UK.

31. McGuinness, D.L. (2002) 'Ontologies Come of Age', in D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster (eds.), *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*, MIT Press.

32. Mika, P. (2005) 'Flink: Semantic Web technology for the extraction and analysis of social networks', *Journal of Web Semantics*, Vol. 3, No. 2–3, pp. 211–223.

33. Nixon, L. (2004) 'Prototypical business use cases', FP6 IST NoE Deliverable D1.1.2, KnowledgeWeb EU NoE - FP6 507482, December. www.starlab.vub.ac.be/research/projects/knowledgeweb/KWebDel112v1.pdf

34. Noy, N.F., Sintek, M., Decker, S., Crubezy, M., Fergerson, R.W., and Musen, M.A. (2001) 'Creating semantic web contents with Protégé-2000', *IEEE Intelligent Systems*, pp. 60–71, March/April.

35. schraefel, m.c., Preece, A., Gibbins, N., Harris, S., and Millard, I. (2004) 'Ghosts in the semantic web machine?', *Proceedings of the First Workshop on Friend of a Friend, Social Networking and the Semantic Web*, Galway, Ireland.

36. Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N, de Velde, W.V., and Wielinga, B. (1999) *Knowledge Engineering and Management: The CommonKADS Approach*, MIT Press.

37. Shadbolt, N., schraefel, m., Gibbins, N., and Harris, S. (2003) 'CS AKTive Space: or how we stopped worrying and learned to love the semantic web', *Procedings of the Second International Semantic Web Conference*, Florida.

38. Shadbolt, N., O'Hara, K., and Crow, L. (1999) 'The experimental evaluation of knowledge acquisition techniques and methods: history, problems, and new directions', *International Journal of Human-Computer Studies (IJHCS)*, Vol. 51, pp. 729–755.

39. Uschold, M. (2003) 'Where are the semantics in the semantic web?', *AI Magazine*, Vol. 24, No. 3, pp. 25–36.

40. Uschold, M., and Gruninger, M. (2004) 'Ontologies and semantics for seamless connectivity', *SIGMOD Record*, Vol. 33, No. 4.

41. Varian, H.R. (2004) 'Competition and market power', in H.R. Varian, J. Farrell, and C. Shapiro (eds.), *The Economics of Information Technology: An Introduction*, Cambridge University Press, Cambridge, pp. 1–47.
42. Weitzner, D.J., Hendler, J., Berners-Lee, T., and Connolly, D. (2005) 'Creating a Policy-Aware Web: discretionary, rule-based access for the World Wide Web', in E. Ferrari and B. Thuraisingham (eds.), *Web and Information Security*, Idea Group Inc.
43. Wenger, E. (1998) *Communities of Practice: Learning, Meaning and Identity*, Cambridge University Press.
44. Wilks, Y. (2004) 'Are ontologies distinctive enough for computations over knowledge?', *IEEE Intelligent Systems*, Vol. 19, No. 1, pp. 74–75.
45. Williamson, O.E. (1975) *Markets and Hierarchies*, Free Press, New York.
46. Williamson, O.E. (1991) *The Nature of the Firm: Origins, Evolution and Development*, Oxford University Press, Oxford.