

The GRID, the WEB and KNOWLEDGE

Hans-Georg Stork
Bollendorf

Abstract

The purpose of this note is to clarify these notions and how they may relate.

The GRID

First of all, there is no such thing as ‘the GRID’. GRID¹ stands for an allegedly new paradigm of large scale scientific computing (or “research networking”): the application of co-ordinated computing resources, interconnected via high-speed networks, to the solution of problems in fields such as High Energy Physics, Astrophysics, Nuclear Physics, Geophysics, Meteorology / Climatology, Neurobiology, Molecular Biology, Earth Observation, Operations Research, etc. GRIDs are *large scale distributed computing systems* providing mechanisms for the controlled sharing of computing resources. (The term has been borrowed from the electric ‘Power GRID’ that enables the sharing of energy resources.)

There are many GRID projects, the latest being *EuroGrid* and *DataGrid* (in Europe) and - for example - GriPhyN (Grid Physics Network) in the US. GRID projects can be classified roughly as *application* oriented (‘vertical’) or *infrastructure* oriented (‘horizontal’). Presumably, EuroGrid and DataGrid are examples of the latter.

Infrastructure oriented GRID projects aim to develop generic ‘middleware’ components, shielding specific applications from the details of accessing and using a configuration of heterogeneous resources, such as processors, storage and network connections. They guarantee resource interoperability through the use of standard protocols. The term *GRID technology*² usually refers to this kind of middleware.

An important function of GRID ‘middleware’ components is to ‘discover’ resources and information about these resources (including information coded as ‘metadata’) in order, for instance, to optimize their use. Middleware components also provide a range of directory and file services. They are needed to create a computing environment that appears to applications more or less as a single huge machine with vast storage resources and processing capabilities.

¹http://www.mkp.com/books_catalog/catalog.asp?ISBN=1-55860-475-8, The Grid: Blueprint for a New Computing Infrastructure; Edited by Ian Foster and Carl Kesselman; July 1998

²<http://www.globus.org/>: Globus project developing fundamental technologies needed to build computational grids.

Most GRID applications (e.g. within the above mentioned areas), apart from requiring enormous processing power, are dealing or expected to deal with huge to gigantic datasets (peta - exa orders of magnitude). Some application projects develop(ed) very specific solutions to resource sharing problems. However, the challenge is to create infrastructure components that support applications of as many kinds as possible. This is the reason why GRID infrastructure projects are currently given particular attention.

So far, industry involvement in GRID application developments has been fairly limited. Today, a single industrial application simply does not produce the large volumes of data that would warrant GRID type solutions (not even in “big engineering”). Of course, this may change. And it will change more rapidly if and when generic GRID middleware becomes more widely available. A likely commercial application domain is simulation of complex processes (e.g. in the aerospace or automotive industries).

For the time being GRID technology is largely (if not almost exclusively) driven by “big science”. ICT industry has a stake in this technology, that goes without saying, primarily as suppliers of software and hardware components. They must and will be involved in its development and in standardisation activities aiming at the specification of open interfaces.

It is by no means clear at this stage what impact GRID technology may have in the future on more general domains, closer to our everyday life. But it was not clear either, way back in the early seventies of the 20th century, when Arpanet³ became operational, what impact the Internet would make today.

Hence, GRID technology is definitely worth watching. It appears, for instance, that through this technology ‘Virtual Organisations’, not only within global research communities but also in more mundane quarters, become a real possibility. As GRID middleware becomes available less resource demanding applications (commercial or not) that nevertheless require large scale resource sharing, may benefit as well.

The Web

Most people have no doubts on what they are talking about when they talk about “the Web”. With probability very close to 1, they mean the World Wide Web, which does indeed have a direct impact on the everyday life of a steadily growing number of people.

Yet, the ‘WWW’ is but one instance, albeit the largest, perhaps most important and certainly the best known one, of a technology the principles of which have been ‘invented’ quite some time ago (e.g. Vannevar Bush’s MEMEX⁴ or Ted Nelson’s XANADU⁵). They have been experi-

³<http://webopedia.internet.com/TERM/A/ARPANET.html>, Wepopedia article on Arpanet, the precursor of the Internet

⁴<http://www.cs.brown.edu/memex/>, The Memex and Beyond web site is a major research, educational, and collaborative web site integrating the historical record of and current research in hypermedia. The name honors the 1945 publication of Vannevar Bush’s article *As We May Think* in which he proposed a hypertext engine called the Memex. (Maintained at Brown University)

⁵<http://xanadu.com.au/xanadu>, “the original hypertext and interactive multimedia system, under continuous development since 1960

mented with, based on comparatively ‘primitive’ precursor technologies⁶ (Interactive Videotex was one of them), a long time before ‘the Web’ became the most popular application of the Internet.

This technology can best be described as *distributed hypermedia systems* whose actual distribution may vary greatly in scale (hypermedia = hypertext + multimedia). The WWW is the outstanding example of a *very large scale* distributed hypermedia system.

While the World Wide Web as well as many local or company Webs (over intra- or extranets) are based on the *HTTP*-protocol for data transfer and *HTML* presentational markup for content display, there are a number of other distributed hypermedia systems built on top of the Internet infrastructure, using different protocols and content description schemes. Early examples developed before or around the time the WWW started to gain ground and momentum are MICROCOSM⁷ (University of Southampton) and HYPER-G⁸ (Technical University of Graz).

Largely due to the ‘simplicity’ of the WWW (e.g. in terms of ease of ‘putting content on the Web’) and its ensuing dominance, alternative designs never really took off. Yet some of them offered a sophistication well ahead of the original WWW structure and functions. (HYPER-G, for instance, offered bidirectional linking, links and content description separated from content, content management through client interfaces, etc., features that are only now entering the WWW world at large, based on a continuously growing set of W3C recommendations, including the *XML* and *RDF* families.)

Of course, saying that Webs are ‘distributed hypermedia systems’ only defers an explanation. The most concise way, perhaps, of characterizing such systems would be as interlinked *digital content* that resides on servers and that can be accessed, represented and interacted with through specific interface clients, known as ‘browsers’.

Digital content can be almost anything. With their capability of interpreting various forms of ‘telesoftware’ (e.g. Java applets, Java code or Javascript) and of hosting so called plug-ins, browsers can indeed deal with a large variety of transaction requirements and content types. Servers, on the other hand, are capable of assembling content on the fly, from all kinds of sources, including data base systems, document management systems and computing facilities in general, thus reacting to whatever request a user may issue through her browser. Processes running ‘behind the scenes’ can be of any degree of complexity.

Due to this generality Webs (and ‘the Web’) lend themselves to all sorts of applications. From a user’s point of view a Web is *simply an interface to the application* she is currently interacting with. It provides her, in particular, with the *address* of that application.

⁶http://www.w3.org/People/howcome/p/elektronikk-4-93/Dybvik_P_E.html, Paper by Per E Dybvik on the differences between Internet services and those of telecom administrations, and the culture clash between telecommunications and computer markets.

⁷<http://www.mmrg.ecs.soton.ac.uk/projects/microcosm.html>, The History of the Microcosm Project; by Wendy Hall

⁸<http://www.byte.com/art/9511/sec5/art4.htm>, Hyper-G Organizes the Web, by Udo Flohr, Byte Magazine, Nov 1995

Applications make use of *resources* (i.e. documents, data residing in data bases, computing facilities, data capturing devices, sensors, etc.). In fact, they may themselves be regarded as resources. This is why '*resource description*'⁹ is all important on Webs. It is a prerequisite for effective and efficient resource discovery and use, just as comprehensive catalogues are needed in order to make full use of brick-and-mortar libraries.

It is worth noting that current Web technology grew out of a research environment. But it is equally worth noting that it has been turned into big business very rapidly. This is not surprising: Web precursors (such as the above mentioned interactive videotex systems) had already been driven largely by business interests. They were targeted at a *mass market*. Yet they failed (except perhaps in France where Minitel received an additional push from the government), mainly because the technology, the underlying infrastructure and the organisational embedding were not sufficiently mature and open. Therefore, the new Internet-based Web technology came in handy, allowing to kill several birds with one stone: it accelerated PC penetration of homes while creating a vast new information and transaction space with many opportunities for incumbent and new entrepreneurs. Business models, however, still seem to be somewhat shaky.

Although Web technology greatly surpasses its - in retrospect - rather clumsy predecessors it has by no means reached its full potential. Whatever this may be: at least two issues have to be resolved, one being the 'semantic' access and use problem¹⁰ (i.e. access to and use of content and services, based on semantically sound resource description); the other being the universality of physical access via high-bandwidth local loops and broadband wireless channels. These are certainly moving targets.

The GRID and the Web

What do GRIDs and Webs have to do with each other? Well, everything and nothing. 'Everything' because in the digital world¹¹ everything can somehow be related to everything else. 'Nothing' because they address entirely different problems, tasks and functionalities.

While both are being operated on the Internet they are currently driven by different needs and interests: GRIDs by 'eScience' (including - more and more - 'industrial science'), Webs by 'eCommerce' and 'eContent' (of which more and more will be multimedia, e.g. for enter-, edu- and infotainment). GRIDs (and GRID technology) have fairly limited user communities; by contrast Webs (and 'the Web' in particular) address and potentially serve millions of people, i.e. the general public. GRIDs are about doing specialist computations on huge to gigantic datasets (cf. above) whereas data volumes flowing across Webs are far more modest, ranging from very small (e.g. a

⁹<http://www.dlib.org/dlib/may98/miller/05miller.html>, An Introduction to the Resource Description Framework; by Eric Miller; D-Lib Magazine; May 1998

¹⁰<http://www.w3.org/People/Berners-Lee/Weaving/>, Weaving the Web; by Tim Berners-Lee with Mark Fischetti; San Francisco 1999

¹¹<http://archives.obs-us.com/obs/english/books/nn/bdcont.htm>, Being Digital (Selected Bits); by Nicholas Negroponte; New York 1996

	GRIDS	Webs	Comments
main drivers	(big) eScience, eEngineering	scientific communication (initially, now:) eCommerce, eContent (multimedia)	<i>there is some and there may be more overlap in the future</i>
main functions	high performance computing, sharing of computing resources	information, communication, transactions	
applications	computationally hard and data intensive problems in science and engineering (e.g. realistic simulations)	I&C services, education & training, eBusiness, eCommerce (B2B, B2C, B2A, etc.), etc.	<i>Webs are mainly interfaces to 'behind the scene' applications</i>
data volumes	XXL (and bigger)	S - XL	<i>future GRIDs may also work on smaller volumes</i>
resources	storage (incl. caches), bandwidth, processor time, data files, ...	digital content and related services	<i>containers, conveyors & processors vs. content & applications</i>
users	special user groups (scientists, engineers)	general public, businesses, public administrations, etc.	<i>these are only the main target groups</i>
standards	middleware standards need to be agreed	many standards and recommendations exist	<i>GRID and Web communities are still fairly separate</i>

Table 1: A “comparison” of GRIDs and WEBS

transaction request) to very large (e.g. high-definition streamed video), also depending on the capacity of physical access paths.

Of course, there are some common basic problems which may have common solutions; GRID and Web developers may actually benefit from each other, for instance in the area of metadata codification. One has to bear in mind, however, that the characteristics of resources are quite distinct between GRIDs and Webs (cf. the above explanations). Table 1 summarizes the ‘differences’ highlighted so far.

To take ‘nothing’ for an answer to the question introducing this section may indeed be a bit too little. And it is certainly not necessary. The clue to a possibly correct understanding of the relationship between Webs and GRIDs lies in the statement “Webs are mainly interfaces to ‘behind the scene’ applications”. We noted that these applications can be arbitrarily

complex. And we do not usually care who or what is working ‘behind the scenes’. So it may be GRIDs (or isolated high performance computers or just an ordinary PC, or whatever). Indeed, GRID applications could render invaluable services even to the general public, via specialised professionals such as medical doctors or surgeons. These applications would be accessed via a Web and their output (e.g. visual representations of complex objects or simulations) translated into standard Web formats.

And GRIDs could provide the Web(s) with ‘knowledge’ as will become apparent from the remaining two sections of this note.

Knowledge

Again, we all believe we know what knowledge is. Yet, if asked what it means in contexts such as ‘knowledge *management*’, ‘knowledge *technologies*’, ‘knowledge *engineering*’, ‘knowledge *representation*’, ‘knowledge *acquisition*’, ‘knowledge *discovery*’, ‘knowledge-based *system*’, ‘knowledge-based *economy*’ and - last but not least - ‘knowledge *society*’, we probably find ourselves in a quandary (perhaps not only then ...). Now we even hear about an emerging ‘*Semantic Web*’¹² (of knowledge, presumably) and - more recently - there has been talk of a ‘*Knowledge Grid*’¹³.

Given this quasi inflationary use of the term our belief in knowing what knowledge is may indeed become less and less firm. Popular attempts to posit some kind of hierarchical or layered structure consisting of ‘data’, ‘information’ and ‘knowledge’ (bottom up in that order) do not seem to contribute much to clarifying the concept. It remains fuzzy in spite of thousands of years of philosophical rumination. And there is little reason to believe that too many of the contemporary proponents of the new knowledge terminology (as in the preceding paragraph) use this word for anything better than a convenient label.

What is behind it? And does it relate in any way to possible common sense ideas of what knowledge might be? Maybe in some of its appositions? The answer - we anticipate - is “yes” in most cases, providing we are a trifle more specific about the meaning of ‘knowledge’.

One common sense idea would be that knowledge is something very personal, that it is about something in the real world, that it can be gained, either through interpersonal communication or through personal experience, that it can be made operational through the decisions we take, the things we make, or - more generally - the way we behave. It can also be shared - through interpersonal communication - although we can never be sure the persons we share it with will gain the same knowledge we have. It should certainly be testable (we remember those days in school, don’t we ...): when queried the person tested should be able to phrase her knowledge,

¹²<http://www.scientificamerican.com/2001/0501issue/0501berners-lee.html>, The Semantic Web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities; by Tim Berners-Lee, James Hendler and Ora Lassila; Scientific American; May 2001

¹³http://www.researchcouncils.ac.uk/ltr/finalversion/LTTR_InformationToKnowledge.htm, Joint UK Research Councils: Long Term Technology Review of the Science & Engineering Base - Chapter 7: Information to Knowledge

preferably in terms of the tester's language, or behave in some other way, thereby demonstrating her knowledge.

Unfortunately, none of these characteristics distinguishes knowledge from other forms of mental content such as beliefs and opinions. So there must be more to it. Maybe 'verifiability'? For a statement (e.g. 'the moon consists of green cheese' which could also be the expression of a belief or an opinion) to qualify as a piece of knowledge we should require some kind of proof (e.g. by logical deduction from first principles or axioms, very much the way mathematicians would do it) or some factual evidence (e.g. a specimen of lunar matter that may be green cheese or not), possibly derived from experiment, observation or analysis.

That looks much better indeed. It tells us that the concept of knowledge implies a procedural element whereby statements (or things) claimed to be based on knowledge should ultimately be 'true' reflections of things existing or events happening in the so called '*real world*' (which - as we all know - is extremely complex and can only be mastered through suitable compartmentalisation, subdivision into domains, modularisation, filtering, etc.).

So, knowledge should have something to do with how we perceive the world around us, how much we perceive of it and how well our perceptions reflect what really exists or happens. This gives rise to a feedback loop: the way we perceive now and what, depend on previous perceptions and how well they matched reality. This feedback loop is usually called 'learning'. It leaves knowledge its subjective, personal touch, given that different people learn different things differently. Hence there is knowledge of different kinds and varying degree or 'depth' (which should be measurable somehow).

Yet, our discussion so far also insinuates - and strongly so - that the concept at issue goes beyond subjectivity, providing the '*real world*' does have an objective existence (and hence sets the standards of 'truth'). It is ultimately the '*real world*' that determines what is knowledge and what is not.

Knowledge itself, however, is in the mind, an intangible entity ('mindware' so to speak), a ('true') representation of facets of the real world. Hence, knowledge representation in humans boils down to implementing in biological 'wetware' a more or less elaborate *model (or 'abstraction') of the real world*. (This is another way of saying that humans 'learn'.) We are only now beginning to find out how this works. But we can definitely see the 'output of human knowledge': the changes we make to the '*real world*' guided by what we know about it. (We note, however, that too often these changes testify to the inadequacy of our world models!) And humans communicate their knowledge mainly in terms of formalisms based on discrete symbols. (Here we should note that the representational power of any such formalism is necessarily incomplete¹⁴.)

To summarize: humans have 'knowledge'; it is somehow represented in human brains; humans can acquire and perhaps even discover it ... Talking about representation, acquisition and discovery of knowledge in and by hu-

¹⁴<http://www.ddc.net/ygg/etext/godel/> (Gödel's theorem)

mans definitely makes sense. We may even describe human beings as rather sophisticated ‘knowledge-based systems’.

Which is not to say they are only that. In fact we may ask whether being ‘knowledge based’ is a distinguishing feature of humans? Or, more precisely: is awareness as experienced by humans (who presumably have some idea of what they know or do not know) a necessary prerequisite for gaining and using knowledge? Most likely not. Children learn most (i.e. develop mental models of the real world) when they seem to be the least aware of it. Animals learn. And we have long since succeeded in building machines that (learn to) recognize and classify correctly patterns, shapes, colours and all kinds of objects, and perform certain actions based on these classifications. But we are quite reluctant to qualify such machines as ‘conscious’ or ‘self-aware’.

Yet there must be some knowledge representation underlying the workings of these machines, that captures the relevant aspects of the world in which they have to carry out their tasks. Machines can be endowed with these representations in very much the same way as bio-wetware: they are either built in or gradually acquired or both. Well-known examples of the latter are machines whose design follows the ‘artificial neural network’ paradigm and hence mimic the mechanisms believed to make our neural tissue tick. In principle every computer programme that is not entirely nonsensical embodies assumptions on some mini-world. (It may of course still be rather nonsensical, depending on the validity of the assumptions.)

We conclude: If ‘knowledge technologies’ are about ‘artificial’ (i.e. man-made) methods and tools for creating, manipulating and sharing ‘knowledge representations’ (i.e. abstractions/models of the real world) then talking about ‘knowledge technologies’ definitely makes sense too. However, if it makes sense now it must have made sense already a fairly long time ago because humans started a long time ago to create increasingly complex external representations (i.e. outside their brains) of some of their knowledge at least. Needless to enumerate the stages of this evolution and the many artefacts invented.

What then are ‘knowledge technologies’ in our digital era? What is new in the digital era if anything? There are at least three *fundamental novelties*:

- We have already indicated the *first* one: our ability to construct machines that learn and develop knowledge representations as they go along. We may call what is being represented in such machines, ‘machine knowledge’ (although, of course, the structure of machine knowledge representations is already inherent in the design, but so is human knowledge representation). These machines are still very primitive, by comparison, but they prove the viability of the approach.
- The *second* fundamental novelty is, for the time being at least, perhaps even more dramatic: the ability to create, maintain and use external (symbolic) representations of human knowledge in hitherto unreachable dimensions, thanks to tools that are many orders of magnitude more powerful than pen, paper, the printing press or library catalogues.

- *Thirdly*, the digital technologies have enhanced drastically our ability to analyse what is going on in the world, to peruse vast amounts of data, searching for structure, thus refining our models of the world, and adding to human knowledge. These data are, by the way, mainly being collected through devices which themselves owe their existence to digital technologies. (To some extent, the ‘first novelty’ may in fact be considered a special case of the third one.)

These developments are having a steadily increasing impact: They are transforming industrial production processes, the way we create and distribute ‘content’ for human consumption, the way we do science, the way businesses are managed, the way public administrations work, etc. But none of these developments has come out of the blue. They have been going on for decades; they are the very gist of the evolution of digital technologies.

And given that human knowledge has always been expressed mainly through symbolic representations there has always been a discipline of ‘knowledge management’, that is: the systematic creation, structuring, updating, sharing and exploitation of sets (or repositories) of such representations (documents).

So, talking about ‘(human) knowledge management’ in the context for instance of the second fundamental novelty is certainly justified, providing we read ‘knowledge’ as: ‘external symbolic representations of human knowledge’. We should, therefore, be aware that systems grandly advertised as ‘knowledge management systems’ are in fact more or less sophisticated ‘document management systems’ (the notion of document being understood in its most general sense), with varying granularity and structural refinement. (This applies in particular to ‘Software Engineering Environments’ the core of which are systems managing documents called ‘software’, undoubtedly one of the most valuable kinds of knowledge representation in the digital era.)

Most ‘knowledge management systems’ are advertised as dealing with ‘corporate knowledge’, certainly a very challenging and profitable market. We note, however, that precisely this market has been the focus of past activities resulting in systems with varying labels: Management Information Systems (MIS), Office Automation Systems, Decision Support Systems, Expert Systems, Computer Supported Cooperative Work (CSCW) systems, Corporate Information Systems, etc. (not to mention the multitude of isolated or linked business application systems and tools for building such applications). It makes us wonder if ‘knowledge management systems’ now denotes the grand finale.

Of course, we have to admit quite emphatically that the new dimensions opened up through advances in networking, microelectronics, software and modelling techniques also present entirely new (research) challenges and (business) opportunities to develop further and apply the capabilities of systems of this kind. Maybe a new label is indeed necessary to draw sufficient attention to this fact.

New modelling techniques¹⁵ in particular, that have been developed and used in the AI (Artificial Intelligence) community for some time make it possible to formulate explicit representations of ‘corporate knowledge’, formal descriptions of a company’s ‘real world’ that can be interpreted according to some agreed semantics. A business process, for instance, would no longer be coded directly in software to be automated, but specified and documented as a piece of ‘corporate knowledge’.

[We refrain from commenting on non-technical terms such as ‘knowledge-based economy’ and ‘knowledge society’. These are largely political euphemisms better to be discussed under the heading ‘newspeak’. (We are waiting impatiently for ‘knowledge-based politics’ or at least a ‘knowledge-based government’ ...)]

The “Knowledge Grid” versus the Semantic Web

Back to GRIDs and Webs: The ‘Knowledge Grid’ appears as layer 3 of a generic Grid architecture described in the Grid section of the UK e-Science programme (started in 2000). The bottom layer is the Computation/Data Grid, and layer 2 of the e-Science Grid model is called ‘Information Grid’. These first two layers make up the technology explained in the GRID section of the present note.

A note (of September 1999) to the British Research Councils characterizes layer 3, the ‘knowledge layer’, as follows:¹⁶

“a knowledge grid superimposed on (b) [=layer 2] utilising KDD (knowledge discovery in database) technology of which a well-known component is ‘data mining’. The knowledge grid will also support intelligent assists to decision makers (from control room to strategic thinkers) and provide interpretational semantics on the information.”

In this wording the ‘Knowledge Grid’ addresses almost precisely what we highlighted in the previous section as the ‘third novelty’. However, from the same note we then learn about special requirements on this layer:

“The provision of a knowledge grid requires two major elements: firstly an agreed knowledge representation and then provision by elicitation from humans and/or discovery (inference), from databases, of knowledge in this representation and secondly the provision of homogeneous access over heterogeneous sources of scholarly publications and grey literature. This is likely to include facilities such as thesauri and/or domain ontologies to assist in understanding and multilingual facilities. Once again, coordination is the key to effective provision. ...”

¹⁵<http://www.ontoknowledge.org/>, On-To-Knowledge: Content-driven Knowledge-Management through Evolving Ontologies; a project in the Information Society Technologies (IST) Program for Research, Technology Development & Demonstration under the 5th Framework Program.

¹⁶<http://www.escience.clrc.ac.uk/KnowledgeInformationData.doc>, to the best of the author’s knowledge it appeared nowhere else denoting a similar concept

This, in turn, seems to address also the ‘second novelty’ mentioned above. And, most surprisingly, the objectives seem to coincide more or less with those of the W3C ‘Semantic Web’ activity. It even goes so far as to recommend a particular application that was one of the principal motivations of early Web development¹⁷. The W3C ‘Semantic Web’ activity has been described as follows:¹⁸

“The Semantic Web is a vision: the idea of having data on the Web defined and linked in a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications. In order to make this vision a reality for the Web, supporting standards, technologies and policies must be designed to enable machines to make more sense of the Web, with the result of making the Web more useful for humans. Facilities and technologies to put machine-understandable data on the Web are rapidly becoming a high priority for many communities. For the Web to scale, programs must be able to share and process data even when these programs have been designed totally independently. The Web can reach its full potential only if it becomes a place where data can be shared and processed by automated tools as well as by people.”

And:

“The Semantic Web approach proposes languages for expressing information and the relationships between information. Initially these languages provide the means for humans to encode meaning in relatively abstract ways that facilitate other machine processing with human intervention. Over time, these languages will accommodate additional formal systems techniques for verification of logical consistency and for reasoning.”

The alleged scope of the ‘Knowledge Grid’ appears indeed to be wider, encompassing both, knowledge representation formalisms and the act itself of representing knowledge (through ‘discovery’, elicitation, inference, analysis, etc.). (The latter is not explicit in W3C documents, and perhaps for good reasons. The ‘Semantic Web’ vision highlights features also seen as characteristics of the ‘Knowledge Grid’.) It is very much akin to the concept of ‘Semantic Web Technologies’ that underlies the IST 2001 action line of the same designation.

However, in the light of our ‘Grid versus Web’ discussion it is not quite clear why there should be a ‘Knowledge Grid’ (or ‘Knowledge Grids’) with ‘Semantic Web’ functionality separate from ‘the Web’. It (or they) should,

¹⁷<http://www.w3.org/History/1989/proposal.html>, Information Management: A Proposal; by Tim Berners-Lee, CERN; March 1989, May 1990; (This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system.)

¹⁸<http://www.w3.org/2001/sw/Activity>

on the contrary, be an integral part of it, providing the multitude of services made possible through its (or their) computational underpinning, supplying a good deal of the semantics of a 'Semantic Web', thus responding to the challenge of marrying the second and third novelty. Whether the offspring will still be branded 'Web' is an entirely open question.