This is a pre-publication version of a paper that will appear as *Shadbolt, N. R., & Smart, P. R. (2015) Knowledge Elicitation. In J. R. Wilson & S. Sharples (Eds.), Evaluation of Human Work (4th ed.). CRC Press, Boca Raton, Florida, USA.* (http://www.amazon.co.uk/Evaluation-Human-Work-Fourth-Wilson/dp/1466559616/).

Knowledge Elicitation: Methods, Tools and Techniques

Nigel R. Shadbolt[1] and Paul R. Smart[1]

[1]*Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK.*

Introduction

Knowledge elicitation consists of a set of techniques and methods that attempt to elicit the knowledge of a domain expert[1], typically through some form of direct interaction with the expert. Knowledge elicitation is a sub-process of knowledge acquisition (which deals with the acquisition or capture of knowledge from any source), and knowledge acquisition is, in turn, a sub-process of knowledge engineering (which is a discipline that has evolved to support the whole process of specifying, developing and deploying knowledge-based systems).

Although the elicitation, representation and transmission of knowledge can be considered a fundamental human activity – one that has arguably shaped the entire course of human cognitive and social evolution (Gaines, 2013) – knowledge elicitation had its formal beginnings in the early to mid 1980s in the context of knowledge engineering for expert systems[2]. These systems aimed to emulate the performance of experts within narrowly specified domains of interest[3], and it initially seemed that the design of such systems would draw its inspiration from the broader programme of research into artificial intelligence. In the early days of artificial intelligence, much of the research effort was based around the discovery of general principles of intelligent behaviour. Newell and Simon's (1963) General Problem Solver exemplified this approach. They were interested in uncovering a general problem solving strategy that could be used for any human task. In the early 1970s, however, a new slogan came to prominence: 'in the knowledge lies the power'. A leading exponent of this view was Edward Feigenbaum from the Stanford Research Institute. He observed that experts are experts by virtue of domain specific problem solving strategies together with a great deal of domain specific knowledge. This view received support from research into the psychology of problem solving that suggested that expert problem solving performance was attributable to the possession of domain specific facts and rules (Chi et al., 1988).

The realization that knowledge lay at the heart of expertise triggered a flurry of interest in knowledge elicitation and representation. Knowledge engineers soon discovered, however, that acquiring sufficient high-quality knowledge from individuals to build a robust and useful system was a very time-consuming and expensive activity. It seemed to take longer to elicit knowledge from

[1] It should be pointed out that although early conceptualizations of knowledge elicitation cast the process as one of extracting or mining knowledge from the heads of experts, more recent conceptualizations view the process as a modelling exercise. The idea is that the knowledge elicitor and domain expert work together in order to create a model of an expert's knowledge. This model may reflect reality to a greater or lesser extent.

[2] Experts systems are computer programs that embody domain-specific knowledge and that perform at the same level as human experts within some domain (although they do not necessarily solve problems in the same way as human experts).

[3] Some early examples of such systems are MYCIN (Shortliffe, 1976) for diagnosing bacterial infections and PROSPECTOR (Duda et al., 1979) for supporting decisions relating to geological exploration.

experts than to write the expert system software. This problem became widely recognized as the *knowledge acquisition bottleneck* (Hayes-Roth et al., 1983), and it spawned an interest in the development, evaluation and practical application of a broad range of knowledge elicitation techniques that continued throughout the 1980s and 1990s.

Today, the scope of knowledge engineering efforts are much broader than simply the development of expert systems. With the advent of the Web and Semantic Web[4], the focus of many knowledge engineering efforts has changed (Gil, 2011; Schreiber, 2013), and the development of formal computational ontologies[5] is now a major focus of attention for those concerned with the elicitation, representation and exploitation of human knowledge. There is also a broader recognition of the role that knowledge elicitation can play in corporate knowledge management. There are many different characterizations of knowledge management, but the central assumption is that knowledge is a valuable asset that must be managed (Nonaka & Takeuchi, 1995; Stewart, 1997). What we are looking for in knowledge management is a means to get the right knowledge to the right people at the right time and in the right form. These are difficult challenges, and many of them are identical to those encountered with the attempt to develop early knowledge-based systems (Hayes-Roth et al., 1983). There is thus a growing appreciation of the value of incorporating knowledge elicitation techniques into knowledge management initiatives, and it has been suggested that the tools, techniques, methods and approaches of knowledge engineering are well suited to the knowledge management enterprise (Gavrilova & Andreeva, 2012; Milton et al., 1999). One topic of particular interest concerns the use of knowledge elicitation techniques to support the transformation of tacit knowledge into explicit knowledge as part of the cycle of organizational knowledge creation (Nonaka & Takeuchi, 1995). Many of the knowledge elicitation techniques presented below can assist with this process, and they may thus play important roles in enabling organizations to realize their innovative potential.

This chapter will discuss the problem of knowledge elicitation for knowledge intensive systems in general. These systems now come in a bewildering range of forms, from conventional expert systems through to intelligent tutoring systems, adaptive interfaces and workflow support tools. In many cases, the goal of knowledge elicitation is simply to generate representations of knowledge that may or may not be exploited in the context of computerized systems. One of the aims of knowledge elicitation, for example, may be to document the work-related knowledge and expertise that has developed within an organization over a period of time. In addition, there may be a requirement to capture the knowledge of individuals who are about to leave an organization or who have recently retired. These kinds of knowledge elicitation efforts often form part of an effort to

[4] The Semantic Web is a set of technologies that provide a common framework for the representation and exchange of knowledge and data in the context of the World Wide Web (Berners-Lee et al., 2001; Shadbolt et al., 2006).

[5] A 'computational ontology', in this case, is a formal, machine-readable representation of knowledge in some domain of interest. In the context of the Semantic Web, ontologies are typically created using the representational formalisms provided by the family of languages that goes under the heading of the Web Ontology Language or OWL. Such languages have both a formal semantics and an RDF/XML-based serialization. The formal semantics provide the basis for forms of machine-based reasoning in which a system is able to infer additional information based on the data that is explicitly represented, while the RDF/XML-based serialization enables knowledge to be published and exploited within the distributed infrastructure of the World Wide Web.

preserve organizational knowledge and expertise by making the knowledge available to new recruits.

Another goal of knowledge elicitation and modelling, especially in more recent times, is to create computational ontologies that can be used in the context of the Semantic Web. The Semantic Web is a vision of how information can be represented and exchanged in the distributed computing environment of the World Wide Web. The essential idea is that information should be represented in a common form and with common semantics. This enables data to be shared, reused and processed across application, enterprise and community boundaries. Unlike the case with the conventional Web, which is designed largely for human consumption, the aim of the Semantic Web is to support greater levels of machine intelligence and more advanced forms of human-machine interaction. In this respect, it is important to bear in mind that the Semantic Web is not a replacement for the conventional Web; rather, it is something that sits alongside the conventional Web and extends the range of capabilities and forms of interaction that can be delivered:

> *"The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation" (Berners-Lee et al., 2001)*

Ontologies play an important role in the context of the Semantic Web. They provide machine-readable representations of human knowledge that specify the knowledge structures of interest in some domain. Such forms of knowledge representation may serve a variety of purposes. As is the case with any form of Web-accessible content, it is not always easy to anticipate the kind of ways in which these epistemic resources will be exploited. They may be used to support the implementation of intelligent systems, they may be used to support data interoperability and exchange solutions, or they may simply be used to enable semantic search through domain-specific resource repositories.

Many problems arise before the elicitation of detailed domain knowledge can begin. Firstly, we need to fully understand the goal of a knowledge engineering project. Sometimes a key failure is in formulating the role of a knowledge-based system; on other occasions it is a failure to appreciate what it is realistic to build. Systems can fail because no one has thought of the social and organisational problems that must be resolved in deploying a system. Very often the effort and resources required to build systems are underestimated – this occurs in both the development and maintenance of systems. A particularly difficult situation arises when one is expected to conjure up knowledge for areas in which no evidence of systematic practice exists at all. Here, one is expected to provide theories for domains where there is no theory.

In term of the actual process of knowledge elicitation, one may be able to gather information from a variety of non-human resources: textbooks, technical manuals, case studies and so on. However, in most cases one needs to consult a practising expert. This may be because there isn't the documentation available, or because real expertise derives from practical experience in the domain rather than from a reading of standard texts. Few knowledge-intensive systems are ever built without recourse to experts at some stage. Those systems not informed by actual expert understanding and practice are often the poorer for it. One of the recent slogans to emerge from the knowledge and cognitive engineering community is that the 'gold is not in the documents':

> *"The gold is not in the documents. Document analysis is useful in bootstrapping researchers into the domain of study...but experts possess knowledge and strategies that do not appear in documents and task descriptions. Cognitive engineers invariably rely on interactions with experts to garner implicit, obscure, and otherwise undocumented expert knowledge." (Hoffman & Lintern, 2006, p. 215)*

Given the need for expert involvement, it is typically the case that a knowledge engineer will be responsible for eliciting the expertise of experts. The main challenge here is to find a means by which the expert is enabled to communicate their knowledge to the person responsible for developing a knowledge solution. How can we establish the conditions that enable the expert to communicate the knowledge that underlies their expertise? This is a hard enough problem in itself, but there are a variety of circumstances that contrive to make the problem even harder. Much of the power of human expertise lies in laid-down experience, gathered over a number of years, and represented as heuristics[6]. Often the expertise has become so routinized that experts no longer know how they accomplish particular tasks. In other cases, the knowledge required to build a system is distributed across an organisation and resides in the minds of a number of experts.

Of course, it is not just the capacity to elicit knowledge from an expert that is important. We would also like the knowledge elicitation process to be highly efficient and address the aforementioned knowledge acquisition bottleneck. Ideally, we would like to be able to use techniques that minimise the effort spent in gathering, transcribing and analysing an expert's knowledge. We would also like to minimise the time spent with expensive and scarce experts. And, of course, we would like to maximise the yield of usable knowledge.

These sorts of issues lie behind the development of the many knowledge elicitation techniques that have become available over the past 20-30 years. A number of surveys of these techniques are now available (Cooke, 1994, 1999; Hoffman, 1987, 1989; Hoffman et al., 1995; Milton, 2012; Shadbolt, 2005; Shadbolt & Burton, 1995), and the current chapter builds on these existing surveys. We begin by describing, in sufficient detail for the reader to apply them, examples of major knowledge elicitation techniques. We then consider the features of domain experts and their associated expertise that are likely to directly affect the knowledge elicitation process. We also describe some of the issues that surround the appropriate selection of knowledge elicitation techniques as part of a programme of knowledge elicitation. Our attention then turns to some of the available software tools that support the knowledge elicitation process, typically by providing computerized versions of one or more knowledge elicitation techniques. Finally, we discuss some of the implications of the Web and Semantic Web for knowledge elicitation efforts.

Knowledge Elicitation Techniques

There are a range of techniques that can be used to elicit knowledge from domain experts. The techniques we will describe are methods that we have found in our previous work to be both useful and complementary to one another. We can subdivide them into *natural* and *contrived* methods. The distinction is a simple one. A method is described as natural if it is one an expert might informally adopt when expressing or displaying expertise. Such techniques include interviews or the observation of actual problem solving. There are other methods we will describe in which the expert

[6] An heuristic is defined as a rule of thumb or generally proven method to obtain a result given particular information.

undertakes a contrived task. Examples here include concept sorting and the repertory grid technique. In the case of contrived tasks, the task elicits expertise in ways that are not usually familiar to an expert, and experts may feel uncomfortable when asked to perform them. Indeed, experts may feel they are performing badly with such methods, and they may question the value of such methods in tapping into their expertise. In this respect, it is worth noting that we have found that an expert's own opinion of the worth of a technique is no guide as to its actual value (Schweikert et al., 1987). In addition, contrived techniques can sometimes prove more efficient than their non-contrived counterparts when it comes to knowledge elicitation (Burton et al., 1990). For these reasons, it is often useful to incorporate the use of contrived techniques into a program of knowledge elicitation, although time will often be required to explain the use of these techniques to domain experts.

Interviews

Almost everyone starts in knowledge elicitation by determining to use an interview. The interview is the most commonly used knowledge elicitation technique, and it takes many forms. Three kinds of interview are generally recognized within the knowledge engineering community. These are the *unstructured, semi-structured* and the *structured* interview. In all cases, the main aim of the interview is to elicit information regarding how a particular task is performed or how a particular decision is made.

The starting point for most new knowledge engineering efforts will be an unstructured interview since this is the best means of establishing rapport between the knowledge elicitor and the expert. In addition, unstructured interviews provide a useful means of 'bootstrapping' the elicitor's understanding of the target domain – they provide an opportunity for the elicitor and the expert to discuss the domain in an informal setting with no constraints as to what can be discussed. Unfortunately, this is also one of the main drawbacks of the unstructured interview. By virtue of being unstructured, the interview can easily allow the elicitor and expert to dwell on irrelevant topic areas or cover important areas in insufficient depth. For these reasons, there is often a requirement to resort to more structured interviewing methods.

The structured interview is a formal version of the interview in which the person eliciting the knowledge plans and directs the session[7]. A significant benefit of the structured interview is that it provides structured transcripts that are easier to analyse than unstructured conversations. This serves to improve the efficiency of the structured interview, and it also enables the elicitor and expert to focus their attention on a limited subset of important topics.

Although it is common to see the structured interview as a single technique, it is probably best to think of it as a class of techniques (Hoffman et al., 1995). There are, in fact, many varieties of structured interviews. In *forward scenario simulation interviews*, for example, the expert is walked through the problem verbally by the elicitor who presents decision- or task-relevant information to the expert and the expert is asked to respond accordingly (Cordingley, 1989; Grover, 1983). Another

[7] In practice, we have found that it is often useful to involve the expert in the planning of a structured interview session. Expert input at the planning stage can be useful in terms of identifying important areas, and it also enables the expert to have an understanding of what topics will be discussed in advance of the knowledge elicitation session.

kind of structured interview is the *fixed probe interview* in which specific probe questions are used to elicit domain knowledge. A template for such an interview is as follows:

1. Ask the expert to give a brief (10 minute) outline of the target task, including the following information:
   a description of the possible solutions or outcomes of the task;
   a description of the variables that affect the choice of solutions or outcomes; and
   a list of the major rules or procedures that connect the variables elicited to the solutions or outcomes.
2. Take each rule or procedure elicited in Stage 1, ask when it is appropriate and when it is not, and if it is a procedure ask how it is performed. The aim is to reveal the scope (generality and specificity) of each existing rule and hopefully generate some new rules.
3. Repeat Stage 2 until it is clear that the expert will not produce any additional information.

A useful way of obtaining a domain overview (Stage 1 of the structured interview) is to ask probe questions that relate to an individual's specific experience. It is also important in this technique to be specific about how to perform Stage 2. We have found that it is helpful to constrain the elicitor's interventions to a specific set of *probes*, each with a specific function. Here is a list of probes (P) and functions (F) that can help in the first two stages of the interview.

P1.1 Could you tell me about a typical case?
F1.1 Provides an overview of the domain tasks and concepts.

P1.2 Can you tell me about the last case you encountered?
F1.2 Provides an instance-based overview of the domain tasks and concepts.

P2.1 Why would you do that?
F2.1 Converts an assertion into a rule.

P2.2 How would you do that?
F2.2 Generates *lower order* rules.

P2.3 When would you do that?
Is <the rule> always the case?
F2.3 Reveals the generality of the rule and may generate other rules.

P2.4 What alternatives to <the prescribed action/decision> are there?
F2.4 Generates more rules.

P2.5 What if it were not the case that <currently true condition>?
F2.5 Generates rules for when current condition does not apply.

P2.6 Can you tell me more about <any subject already mentioned>?
F2.6 Used to generate further dialogue if the expert dries up.

P2.7 Can you tell me about an unusual case you encountered/heard about from some other expert?
F2.7 Refines the knowledge to include rare cases and special procedures.

The idea here is that the elicitor engages in a type of slot/filler dialogue. The provision of template questions about concepts, relations, attributes and values makes the elicitor's job much easier. It also provides sharply focused transcripts that facilitate the process of extracting usable knowledge. Of course, there will be instances when none of the above probes are appropriate (such as the case when the elicitor wants the expert to clarify something). However, you should try to keep these interjections to a minimum. The point of specifying such a fixed set of linguistic probes is to constrain the expert to giving you all, and only, the information you want.

The sample of dialogue below is taken from a real interview of this kind. It is the transcript of an interview by a knowledge engineer (KE) with an expert (EX) in the domain of geological analysis[8].

KE: **What would you do at this stage?**
EX: **I would look at the grain size of the hand specimen and see how fine it was.**
KE: **Why would you look at the grain size?**
EX: **That will tell me if the rock has been formed near to the surface or deep inside the earth. The finer the grain size the faster it cooled. Coarse crystals indicate that the rock was cooling slowly + forming deeper down + we say its emplacement is plutonic + if it cooled near the surface its emplacement is volcanic.**
KE: **Are there any alternatives to coarse and fine grain size?**
EX: **There are glasses + you can't see any structure here because the rock cooled so fast.**
KE: **What would you look at next?**
EX: **Colour is important + the lighter the rock the more acidic it is.**
KE: **Why is a lighter rock more acidic?**
EX: **Acidic rocks are higher in quartz and colour is a good indicator of quartz content – leucocratic or light things have a lot of quartz – melanocratic that is darker rocks have olivines and pyroxines.**

This is quite a rich piece of dialogue. From this section of the interview alone we can extract numerous rules such as:

```
IF      grain size is large
THEN    rock is plutonic

IF      rock is leucocratic
THEN    rock has high quartz content
```

Of course, these rules may need refining in later elicitation sessions, but the text of the dialogue shows how the use of the specific probes has revealed a well-structured response from the expert[9].

Techniques exist to impose a lesser amount of structure on an interview. These kind of techniques can be referred to as types of semi-structured interview. One example of a semi-structured interview is the *knowledge acquisition grid* (LaFrance, 1987). This is a matrix of knowledge types and forms – examples of knowledge forms are *layouts* and *stories*, while some examples of question types are *grand tour* and *cross-checking*. A grand tour involves such things as distinguishing domain boundaries and the overall organization of goals; cross-checking involves the engineer attempting to validate the acquired knowledge by, for example, playing devil's advocate.

Another form of semi-structured interview technique is the *teachback technique* of Johnson and Johnson (1987). In this technique, the expert explains something to the elicitor who then attempts to explain it to the expert – the knowledge is effectively 'taught back' to the expert. The expert then has an opportunity to check and, if necessary, amend the information.

---

[8] In the transcripts we use the symbol + to represent a pause in the dialogue.
[9] In fact, a possible second-phase elicitation technique would be to present these rules back to the expert and ask about their validity, scope and so forth.

Unstructured interviews have no agenda (or, at least, no *detailed* agenda) set either by the knowledge elicitor or by the expert. Of course, this does not mean that the elicitor has no goals for the interview, but it does mean that she has considerable scope for proceeding. As mentioned earlier, the unstructured interview is useful for a variety of reasons. Firstly, the approach can be used whenever one of the goals of the interview is to establish a rapport between the expert and the knowledge elicitor. There are no formal barriers to the discussion covering whatever material either participant sees fit. Secondly, one can get a broad view of the topic easily; the knowledge elicitor can 'fill in the gaps' in her own perceived knowledge of the domain. Thirdly, the expert can describe the domain in a way with which he is familiar, discussing topics that he considers important and ignoring those he considers uninteresting.

The disadvantages are clear enough: the lack of structure can lead to inefficiency; the expert may be unnecessarily verbose; the expert may concentrate on topics whose importance he exaggerates; the coverage of the domain may be patchy; and the data acquired may be difficult to integrate, either because it does not form a coherent body of content, or because there are inconsistencies (this will be even more likely if the information provided by *several* experts is to be collated).

In all of the interview techniques mentioned so far (and in some of the other techniques as well) there exist a number of dangers that have become familiar to practitioners of knowledge elicitation. One problem is that in an interview experts will only produce what they can verbalise. If there are non-verbalisable aspects to the domain, the interview will not recover them. It may be that the knowledge was never explicitly represented or articulated in terms of language (consider, for example, pattern recognition expertise). Then there is the situation where the knowledge was originally learnt explicitly in a propositional or language-like form. However, in the course of experience such knowledge has become routinised or automatised[10]. This can happen to such an extent that experts may regard the complex decisions they make as based only on hunches or intuitions. In actual fact, these decisions are based upon large amounts of remembered data and experience and the continual application of that knowledge. In this situation they tend to give *black box* replies such as 'I don't know how I do that...' or 'It is obviously the right thing to do...'.

Another problem arises from the observation that people (and experts in particular) often seek to justify their decisions in any way they can. It is a common experience of the knowledge elicitor to get a perfectly valid decision from an expert, and then to be given a spurious justification as to why it was made and how it originated.

For these and other reasons one should always supplement interviews with additional elicitation methods. In general, knowledge elicitation should always consist of a programme of techniques and methods (see section on 'Methodologies and Programmes').

Protocol Analysis

Protocol Analysis (PA) is a generic term for a number of different ways of performing some form of analysis of the expert(s) actually solving problems in the domain. In all cases, the elicitor takes a record of what the expert does using written notes or (preferably) an audio or video recording.

[10] We often use a computing analogy to refer to this situation and speak of the expert as having *compiled* the knowledge.

Transcripts or protocols are then made from these records and the elicitor tries to extract meaningful structure, rules and processes from the protocols.

We can distinguish two general types of PA: *online* and *offline*. In online PA the expert is recorded solving a problem and *concurrently* a commentary is made. The nature of this commentary specifies two sub-types of the online PA method. The expert performing the task may be describing what they are doing as problem solving proceeds. This is called *self-report*. A variant on this is to have another expert provide a running commentary on what the expert performing the task is doing. This is called *shadowing*.

Offline PA allows the expert to comment retrospectively on the problem solving session, usually by being shown an audio-visual record of it. This may take the form of a *retrospective self-report* by the expert who actually solved the problem. Alternatively, it may take the form of a retrospective report by another expert – this has recently been referred to as *collegial verbalization* (Erlandsson & Jansson, 2007) – or there could be group discussion of the protocol by a number of experts including its originator. In situations where only a behavioural protocol (such as a video recording) is obtained then some form of retrospective verbalisation of the problem-solving episode will obviously be required.

In many cases, the focus of protocol analysis is on verbal data. In this case, the technique is typically referred to as verbal protocol analysis (see Bainbridge & Sanderson, 2005). Other types of events, such as eye movements, gestures and other non-verbal behaviours may also be the focus of protocol analysis, although this is rarely seen in practice. Combining the analysis of (e.g.) eye movements with verbal reports may be useful in some cases, particularly in situations where the aim is to better understand the allocation of attention to particular environmental cues and sources of task-relevant information. In one study, for example, Van Gog et al (2005) used a combination of eye movement data and concurrent verbal protocol analysis in order to explore expertise-related differences in electrical circuit troubleshooting performance.

In deciding between the various kinds of PA technique on offer, it is worth bearing in mind a number of issues. Firstly, in their classic treatment of protocol analysis, Ericsson and Simon (1996) recommend the use of concurrent verbal reports (i.e., online self-reports) over retrospective ones. One of the possible problems with retrospective reports is that the conditions associated with verbalization in the two cases may differ, and this may affect information processing accordingly. In general, it is assumed that the longer the delay between performance and report, the greater this problem becomes. As a result, it is predicted that more immediate retrospective reports are the most similar to concurrent ones. On the other hand, concurrent verbalization techniques can present a number of problems for experts, such as interference with the execution of skilled actions. Ericsson and Simon (1996) suggest a number of conditions under which verbal report procedures should succeed or fail. For instance, verbal reports are not as effective for eliciting knowledge when the problem is novel or the reporter has low verbal ability or is inhibited in some way. When these sorts of conditions are encountered in the context of a programme of knowledge elicitation, it may be beneficial to incorporate more retrospective PA techniques.

In addition to decisions about the choice between online and offline PA, decisions also have to be made about the extent to which other experts (other than those actually performing the task) are involved in the verbal commentary. Typically, the individual performing the task provides the verbal report, either concurrently or retrospectively. However, other techniques, such as that of collegial verbalization[11] have also been the focus of recent attention (Erlandsson & Jansson, 2007, 2013). One issue of interest here concerns the extent to which the reports provided by other experts matches those provided by the performing expert. In one study comparing collegial verbalization with retrospective self-report, Erlandsson and Jansson (2013) found a number of similarities between the protocol data delivered by the two techniques, suggesting that collegial verbalization may be as effective as retrospective self-report. Clearly, in a situation where a video record of expert performance is available, a number of protocols can be obtained using multiple experts. This may serve to improve the reliability and completeness of the resulting knowledge base.

In trying to decide when it is appropriate to use PA bear in mind that it is alleged that different knowledge elicitation techniques differentially support the elicitation of particular kinds of information. This is commonly known as the *differential access hypothesis* (Hoffman et al., 1995). With PA, it is claimed that the sorts of knowledge elicited include the "when" and "how" of using specific knowledge. It can reveal the problem solving and reasoning strategies, evaluation procedures and evaluation criteria used by the expert, and procedural knowledge about how tasks and sub-tasks are decomposed. A PA gives you a complete episode of problem solving. It can be useful as a verification method to check that what people say is actually what they do. It can also take you deeper into a particular problem. It is, however, intrinsically a narrow method since it can only be used to analyze a relatively small number of problems within the domain.

Before PA sessions can be held, a number of pre-conditions should be satisfied. The first of these is that the elicitor is sufficiently acquainted with the domain to understand the expert's tasks. Without this, the elicitor may completely fail to record or take note of important parts of the expert's behaviour.

A second requirement is the careful selection of problems for PA. This sampling of problems is crucial. PA sessions may take a relatively long time, and usually only a few problems can be addressed in any programme of acquisition (Shadbolt & Burton, 1989). Therefore, the selection of problems should be guided by how representative they are. Asking experts to sort problems into some form of order (Chi et al., 1981; Chi et al., 1982) may give an insight into the classification of types of problems and help in the selection of suitable problems for PA (see also the following sections on concept sorting, repertory grids and laddered grids for methods that can be used to help classify and structure problems).

A further condition for effective PA is that the expert(s) should not feel embarrassed about describing their expertise in detail. It is preferable for them to have experience in thinking aloud. Uninhibited thinking aloud has to be learned in the same way as talking to an audience. One or two short training sessions may be useful. In these training sessions a simple task, such as long multiplication, can be used as an example. This puts the expert at ease and familiarises them with the task of talking about their problem solving.

[11] Collegial verbalization is based on the procedure of videotaping practitioners while they perform their normal work tasks in their normal work setting. This is followed up by having a close colleague of the practitioner watch the video recordings and verbalise.

In order to collect protocols, the expert is asked to 'think aloud' while performing some task, and the resulting commentary is typically recorded and transcribed. In terms of recording techniques, it is preferable to use video recordings rather than audio recordings. This is because video recordings capture more information about the context in which problem-solving occurs, which can help to support the resulting analysis. In particular, the following two advantages of video recording techniques have been noted by Bainbridge and Sanderson (2005):

1. Firstly, video recordings often help to disambiguate what is being referred to in the case of situated forms of problem-solving activity. Subjects often make use of pronouns, such as 'when *it's* at 55', and the presence of a visual record can help to disambiguate what is being referred to. Also, as noted by Bainbridge and Sanderson (2005), video recordings can help when people use general anaphoric references supplemented by pointing; for example, '*that* is too high so I'll lower *this* until it is between *these*'.
2. A second advantage of video recording techniques relates to the fact that is often useful to have information about the total task environment in which problem solving occurs. This can be used at a later time to assess to what extent people's behaviour is influenced by features of the environment that are not explicitly mentioned in the verbal report.

One of the main drawbacks of video recording techniques is, of course, the amount of data they make available for analysis. It can be difficult to avoid the temptation to scale up the analytic effort when confronted with such detailed records, and discipline is often required to limit attention to information of relevance to the knowledge elicitation effort.

When actually conducting a PA the following are a useful set of tips to help enhance its effectiveness.

1. Present the problems and data in a realistic way. The way problems and data are presented should be as close as possible to a real situation.
2. Transcribe the protocols as soon as possible. The meaning of many expressions is soon lost, particularly if the protocols are not recorded.
3. Avoid long self-report sessions. Because of the need to perform a double task – combining expert performance with verbal commentary – the process of thinking aloud is significantly more tiring for the expert than being interviewed. This is one reason why shadowing is sometimes preferred.
4. In general, the presence of the elicitor is required in a PA session. Although the elicitor adopts a background role, her very presence suggests a listener to the interviewee, and lends meaning to the talking aloud process. Therefore, comments on audibility, or even silence by the elicitor, are quite acceptable.

When a verbal or behavioural transcript has been obtained we next have to undertake its analysis. A number of approaches to the analysis of verbal protocols have been described in previous work, such as that by Bainbridge and Sanderson (2005). In general, however, it is acknowledge that there are no objective independent techniques for doing these analyses, and this means that analysts "have to use both their own natural language understanding processes, and their knowledge of the task, in order to make sense of what is going on, to infer missing passages, and to interpret the results of summary analyses (Bainbridge & Sanderson, 2005, p .166). For the purposes of most knowledge elicitation exercises, the analysis will typically involve the 'encoding' of the protocol transcript into 'chunks' of knowledge (actions, assertions, propositions, keywords, etc.), and it should result in a rich domain representation with many elicited domain features together with a number of specified links between those features. The example below is from a self-report of an expert geologist. It is immediately apparent that protocols can be extremely dense sources of information. A very significant amount of work is required to analyse and structure the content in this very small fragment of a self report concerning one rock specimen.

> **To start off with it's obviously a fairly coarse-grained rock ... and you've got some nice big orthoclase crystals in here – this is actually SHAP GRANITE – I know it just because everybody's seen SHAP GRANITE – or it's a very strong possibility that it's SHAP GRANITE ... it's a typical teaching specimen – as I say the obvious things are these very big orthoclase crystals pink colouration and you can certainly see some cleavage in some of them – you can certainly make out there are feldspar cleavages in there – it's a coarse-grained rock anyway, you can see the crystals nice and coarsely – these large porphyritic crystals – you can see, in the ground mass, you can see quartz – get some light on it (HOLDS SPECIMEN UP TO WINDOW) quartz, which is this fairly clear mineral you can actually look into it and see through it as opposed to calcite or feldspars where it's more cloudy – you can't actually see any good crystal faces on these cut sections – small flakes of biotite, black micacious looking – small plates, you can certainly see some on this specimen even without a hand lens.**

There are a number of principles that can guide the protocol analysis. For example, analysis of the verbalization resulting in the protocol can distinguish between information that is attended to during problem-solving, and that which is used implicitly. A distinction can be made between information brought out of memory (such as a recollection of a similar problem solved in the past), and information that is produced 'on the spot' by inference. The knowledge chunks referred to above can be analysed by examining the expert's syntax, or the pauses he takes, or other linguistic cues. Syntactical categories (e.g., use of nouns, verbs, etc.) can help distinguish between domain features and problem-solving actions, etc. In general, for multiple analysts to perform the encoding independently. This provides insight into the reliability of certain forms of encoding, and it also serves to highlight areas of contention that may need to be the focus of future knowledge elicitation sessions.

The focus and depth of the analytic efforts is typically is dictated by the goals of the knowledge elicitation exercise. If the aim is to understand the sequential ordering of tasks in the context of some larger business process, this will require a less detailed form of protocol analysis compared to situations where the aim is to develop a computational model of the mental processes associated with problem-solving behaviour.

When appropriately elicited, verbal and non-verbal protocols can help to illuminate the normal sequential flow of working and thinking, and they are thus valuable components of the analyst's knowledge elicitation toolkit. In spite of this, protocol analysis does have its have limitations. Firstly, protocol analysis techniques share with the unstructured interview the problem that they may deliver unstructured transcripts that are hard to analyse. Moreover, they focus on particular problem cases and so the scope of the knowledge produced may be very restricted. It is difficult to derive general domain principles from a limited number of protocols. These are some of the practical disadvantages of protocol analysis. However, there are more subtle problems. For example, two actions, which look exactly the same to the knowledge elicitor, may be very different in their extent and intent. For example, our geologist who applies a particular test to a specimen may apply that same test to another but with a quite different purpose. The knowledge elicitor simply does not know enough to discriminate the actions.

Another source of concern stems from the possibility of distorted information – the risk that protocol analysis may yield information that is not an accurate reflection of what takes place in task settings where the technique is not being employed. The causes of these distortions are outlined by Bainbridge and Sanderson (2005). They include:

1. The fact that being asked to give a verbal protocol changes the nature of the task that is being performed. A task that typically involves a number of concurrent actions may instead be performed in a sequential fashion as a result of the constraints imposed by the need to verbalize what one is doing. In cases where the a multiple ways a accomplishing a task, an expert may resort to a method that is easier to verbalize. Self-report techniques may also interfere with expert performance. There is some empirical evidence that attending to the components of a well-learned skill can impair performance (Beilock et al., 2002; Gray, 2004), and it thus seems likely that by asking an expert to think aloud we are changing the nature of the task being performed. Some cases of skilled performance are probably best demonstrated when the expert is left to perform the task automatically without the kind of attentional reorganization that is required by protocol analysis. This may also be the case with certain types of decision making expertise. By asking the expert to verbalise, one is in some sense destroying the point of doing protocol analysis – to access procedural, real-world knowledge.
2. The temporal constraints involved in giving a verbal protocol. In situations where people are working under time constraints, there may be limits to what people can verbalize. In particular, there may be insufficient time to report task-relevant information that is brought to mind and then quickly forgotten as a result of the tempo of task performance.
3. The fact that giving a self-report is a socially-situated activity involving self-presentation issues. People may, for example, want to appear to rational and knowledgeable to a professional observer, and this may influence the content of the self-report accordingly.
4. The fact that some aspects of the task may be performed automatically, and the expert may not have conscious access to the knowledge that is being used. This particularly the case with tasks involving advanced perceptual-motors skill.
5. The limited scope of the technique. By focusing on a limited number of tasks, protocol analysis may inadequately sample the total knowledge possessed by a expert. As noted by Bainbridge and Sanderson (2005) "knowledge about the components, mechanisms, functions and causal relations in a machine, memories of specific events, and helpful categories will be mentioned explicitly only if the task involves some problem solving that requires the person to review this sort of evidence" (p. 162).

Having pointed to these drawbacks, it is also worth remembering that context is often important for memory – and hence for problem solving. For most non-verbalisable knowledge, and even for some verbalisable knowledge, it may be essential to observe the expert performing the task in a naturalistic setting. It may be that this is the only situation in which the expert is actually able to demonstrate their expertise.

Critical Decision Method

The Critical Decision Method (CDM) is "a retrospective interview strategy that applies a set of cognitive probes to actual nonroutine incidents that required expert judgement or decision making" (Klein et al., 1989, p. 464). As a knowledge elicitation technique, the CDM contains elements of both interviewing and protocol analysis but in a context that stresses the examination of problem solving in naturalistic decision making contexts (Zsambok & Klein, 1997). The technique involves the expert being guided through the recall and elaboration of previously encountered cases, especially ones that were, in some sense, unusual, difficult or otherwise involved critical decisions. Such cases are often particularly memorable for the domain expert, and this serves as an aid to the elicitation of important information, such as the information the expert needs to make decisions in particular contexts. At the same time, incidents that are difficult or nonroutine are typically ones that provide the richest source of information about the knowledge and capabilities of domain experts. Detailed presentations of this method, along with summaries of studies illustrating its use, can be found in Klein et al (1989), Crandall et al (2006), O'Hare et al (1998) and Hoffman et al (1998).

As originally presented by Klein et al (1989), a CDM session is organised into five steps.

1. **Select incident.** In the first step, the expert is guided in the recall and recounting of a specific incident and its associated context. As mentioned above the aim is to select an incident that is unusual or nonroutine. The expert may be asked to "select an incident that was challenging and that, in his or her own decisionmaking, might have differed from someone with less experience" (Klein et al., 1989, p. 466). As a second example, experts may be asked to focus on incidents that are "in some manner unusual and difficult (i.e., where the best choice was not clear cut) in which the [expert] felt that their expertise and experience made a critical difference to the outcome" (O'Hare et al., 1998, p. 1700).
2. **Obtain unstructured incident account.** In the second step, the expert is asked to describe the incident from their own perspective. This step accomplishes a number of goals. Firstly, it provides the basis for an analysts initial understanding of the incident in question. Secondly, it serves to activate the expert's memory of an incident as the basis for subsequent questioning.
3. **Construct incident timeline.** After the incident has been described by the expert, a timeline of the account is constructed. This serves to establish the sequence and duration of each event reported by the expert.
4. **Decision point identification.** Once a timeline has been constructed, decision points in the timeline are identified, and specific decisions are marked for further probing. In general, decisions are subjected to further probing if the expert feels that additional courses of action are possible, or if another expert might have chosen a different course of action.
5. **Decision point probing.** Any decision points that were marked for further probing in step 4 are analyzed in more detail using a set of cognitive probes.

Table 1 contains a range of probe question types with exemplars that we have found to be particularly useful when applying the CDM. Although these are typically used in step 5 of the CDM method, there is no reason why these questions cannot be used in the context of other steps. In addition, the probes listed in Table 1 do not exhaust the range of probes that could be used in the context of the CDM. O'Hare et al (1998), for example, present an extended set of cognitive probes that are designed to "obtain additional information on the perceptual and cognitive structures and processes that appear to mediate expertise" (p. 1700).

Probe Type	Probe Examples
Cues	What were you seeing, hearing, smelling?
Knowledge	What information did you use in making this decision? How was it obtained?
Analogues	Were you reminded of any previous incidents?
Scenarios	Does this case fit a standard or typical scenario? Does it fit a scenario you were trained to deal with?
Goals	What were your specific goals and objectives at the time?
Options	What other courses of action were considered or available?
Choice	How was this option selected/other options rejected? What rule was being followed?
Anticipation	Did you imagine the possible consequences of this action? Did you imagine the events that would unfold?
Experience	What specific training or experience was necessary or helpful in this decision? What more would have helped?
Decision making	How much time pressure was involved in making the decision? How long did it take to make the decision?
Aiding	What training, knowledge or information could have helped?
Situation assessment	If you were asked to describe the situation to a colleague at this point, how would you summarise the situation?
Errors	What mistakes are likely at this point? How might a novice have behaved differently?
Hypotheticals	If a key feature of the situation had been different, what differences would it have made in your decision?

Table 1: Sample CDM probe questions.

The outcome of the CDM is a range of products, which can be used to support training and system development activities (Klein et al., 1989). One of the most important products is referred to as the Critical Cue Inventory (CCI) that is a collection of all the perceptual cues that are used to guide the consideration and selection of particular decisions. In the case of medical decision-making, for example, the CCI could include a list of cues for recognizing critical conditions, such as early signs of cardiopulmonary distress (see Klein et al., 1989). Another important product of the CDM is the Situation Assessment Record (SAR). The SAR records the changes in goals and cue usage associated with situation assessment processes. It typically combines information about the cues being sought or identified, the expectancies generated by these cues, the goals activated by the current situation, and the selected course of action resulting from knowledge about the assessed situation.

A typical CDM session can last around 2 hours. Depending on the domain, much of this time may be spent recollecting a rich complex incident. In other settings, the majority of the effort may be devoted to examining counterfactual situations. The CDM does have its limitations. In distributed problem solving situations no one individual may handle more than one element of a task. The individuals, in this case, would never know whether their judgements or assessments were correct within the context of the larger socially-distributed process. In addition, in high workload environments, we have sometimes observed that incidents and events can become merged. When responding to an opening query one sometimes sees an expert recount an incident but then become confused when asked for a timeline or other details. Despite these shortcomings, the style of interview and the attention paid to particular incidents often provides a rich output from which the elicitor can extract important task-relevant knowledge. An added bonus is that the case studies resulting from the application of the CDM can often serve as important training materials.

Concept Sorting

Unlike interview techniques and PA, concept sorting is a form of contrived knowledge elicitation technique that is likely to be unfamiliar to the domain expert. The technique is useful when we wish to elicit the different relationships that exist between a fixed set of concepts. In the version of concept sorting we describe here an expert is presented with a number of cards on each of which is printed a concept word. The cards are shuffled and the expert is asked to sort the cards into either a fixed number of piles or else to sort them into any number of piles the expert finds appropriate. This process is repeated many times.

Using this task one attempts to get multiple views of the structural organisation of knowledge by asking the expert to do the same task over and over again. Each time the expert sorts the cards, he should create at least one pile that differs in some way from previous sorts. The expert should also provide a name or category label for each pile on each different sort. This is often referred to as the *dimension* along which concepts are sorted (see Table 3), and it typically identifies a particular property or attribute associated with a class of objects (e.g., 'grain size' may be represented as an attribute of the 'rock' class).

Performing a card sort requires the elicitor to have some basic conception of the domain. Cards have to be made with the appropriate labels before the session. However, no great familiarity is required as the expert provides all the substantial knowledge in the process of the sort. We now provide an example from our geology domain to show the detailed mechanics of a sort.

The concepts printed on a set of cards are the names of igneous rocks drawn from a structured interview with the expert. He had previously described 18 rock types, which are presented in Table 2.

1	adamellite	10	granite
2	andesite	11	lherzolite
3	basalt	12	microgranite
4	dacite	13	peridotite
5	diorite	14	picrite basalt
6	dolerite	15	rhyodacite
7	dunite	16	rhyolite
8	gabbro	17	syenite
9	granodiorite	18	trachyte

Table 2: The names of 18 types of igneous rock elicited from a geologist as part of a structured interview.

The expert was shown possible ways of sorting cards in a *toy* domain as part of the briefing session. He was then asked to sort the real elements in the same way. The dimensions/piles which the expert used for the individual card sorts are presented in Table 3.

Sort #	Dimension	Piles
1	grain size	1=coarse, 2=medium, 3=fine
2	colour	1=melanocratic, 2=mesocratic, 3=leucocratic
3	emplacement	1=intrusive, 2=extrusive
4	presence of olivine	1=always, 2=possibly, 3=never
5	presence of quartz	1=always, 2=possibly, 3=never
6	percentage of silica	1= >68%, 2= <68%, 3= about 68%
7	density	1=very light, 2=light, 3=medium, 4=dense, 5=very dense

Table 3: The results of seven card sorts undertaken as part of a concept sorting knowledge elicitation session with a geologist.

Table 4 shows the piles into which each of the rock types in Table 2 was placed as part of the sequence of card sorts. As can be seen from Table 4, many of the elements are distinguishable from one another, even with this limited number of card sorts.

	Rock																	
Sort	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
1	1	3	3	3	1	2	1	1	1	1	1	2	1	3	3	1	1	3
2	3	2	2	2	2	1	1	2	3	3	1	3	1	1	3	3	3	3
3	1	2	2	2	1	1	1	1	1	1	1	1	1	2	2	2	1	2
4	1	3	2	3	3	2	1	2	3	3	1	3	1	1	3	3	3	3
5	1	2	2	2	2	2	3	2	1	1	3	1	3	3	1	1	1	2
6	2	2	2	3	2	2	2	2	3	1	2	1	2	2	1	1	2	2
7	1	3	4	2	3	4	5	4	1	1	5	1	5	4	2	1	3	2

Table 4: The positioning of cards representing different types of igneous rock (see Table 2) in the piles resulting from seven card sorts with a geologist (see Table 3).

Using the results of the card sorts, we can attempt to extract decision rules directly. An example of a rule extracted from the card sorting data is:

```
IF      the grain size is fine           (sort 1/pile 3)
AND     the color is mesocratic          (sort 2/pile 2)
AND     its emplacement is extrusive     (sort 3/pile 2)
AND     it does NOT contain olivine      (sort 4/pile 3)
AND     may possibly contain quartz      (sort 5/pile 2)
AND     it contains less than 68% silica (sort 6/pile 2)
AND     its density is medium            (sort 7/pile 3)
THEN    the rock is andesite             (outcome 2)
```

As can be seen from this example, card sorts often produce long and cumbersome rules. In fact many of the clauses may be redundant. For example, once you have established that the grain size is small, then it is going to be an extrusive rock. The utility of the technique, however, does not reside solely in the production of decision rules. We can use it, as we have said, to explore the general inter-relationships between concepts in the domain. We can also use the technique to elicit the features of concepts[12] that might not otherwise surface in the context of other techniques.

The advantages of concept sorting can be characterised as follows. It is fast to apply and easy to analyse. It also serves to make explicit the implicit structure that experts impose on their expertise. In fact, the process of performing concept sorting is often instructive to the expert – a sort can lead the expert to see structure that he himself has not consciously articulated before. Concept sorting can also be a highly efficient technique, especially when computerised support is available for the implementation and analysis of the sorting procedure. Unlike the case with interviews and protocol analysis, time can often be saved by not having to transcribe and analyse lengthy verbal reports[13].

[12] It is important to bear in mind that although the name of the technique suggests that its use is limited to concepts, the technique can, in fact, be applied to knowledge elements of any type. The cards used in a card sorting task, for example, might name tasks, goals, actions, resources, and so on; the only restriction is that in any sorting session the cards should be of the same knowledge type.

[13] Although it is not necessary to make an audio recording of concept sorting sessions, we recommend that such records are, in fact, made. An expert makes many asides, comments and qualifications in the course of sorting ranking and so on. In fact one may choose to use the contrived methods as a means to carry out auxiliary structured interviews. The structure this time is centred on the activity of the technique.

Finally, in domains where the concepts are perceptual in nature (i.e. X-rays, layouts and pictures of various kinds), then the cards can be used as a means of presenting these images and attempting to elicit names for the categories and relationships that might link them.

The techniques does, of course, have its disadvantages. Experts can often confound dimensions by not consistently applying the same semantic distinctions throughout an elicitation session. Alternatively, they may over simplify the categorisation of elements, missing out important caveats.

Repertory Grids

This technique has its roots in the psychology of personality (Fransella et al., 2003; Jankowicz, 2003; Kelly, 1955). It is designed to reveal a conceptual map of a domain in a fashion similar to the concept sorting technique discussed above. The work of Mildred Shaw and Brian Gaines was particularly important in promoting the use of the technique (Shaw & Gaines, 1987), and the development of computerized versions of the technique was an important step in making the repertory grid a standard element of the knowledge elicitation technique palette (the technique as developed in the 1950s was very time-consuming to administer and analyse by hand). One example of repertory grid software is WebGrid 5, which can be accessed from the WebGrid website[14]. WebGrid 5 is the latest version of the Web-based implementation of the repertory grid technique that was described by Gaines and Shaw (Gaines & Shaw, 1997; Shaw & Gaines, 2001), as part of their attempt to make knowledge acquisition technologies accessible via the World Wide Web. The software provides an excellent means of experimenting with the approach and indeed undertaking machine-supported elicitation sessions

As part of the repertory grid technique subjects are presented with a range of domain elements and asked to choose three, such that two are similar, and different from the third. This is known as the method of triadic elicitation (e.g., Caputi & Reddy, 1999)[15]. In order to demonstrate this technique, suppose we were trying to uncover an astronomer's understanding of the planets within our own solar system. We might present her with a set of planets, and she might choose Mercury and Venus as the two similar elements and Jupiter as different from the other two. The expert is then asked for her reason for differentiating these elements, and this dimension is known as a construct. In our example, 'size' might be a suitable construct that differentiates between the selected elements. The remaining elements are then rated with respect to this construct.

This process continues with different triads of elements until the expert can think of no further discriminating constructs. The result is a matrix of similarity ratings, relating elements and constructs. This is can be analyzed using a variety of statistical techniques, of which the most popular is probably called cluster analysis. Cluster analysis can reveal clusters of concepts, some of which may not have been articulated using other kinds of techniques (e.g., interviews).

Figure 1 shows the results of a repertory grid applied to the domain of planets (within our own solar system). We can see that the expert has so far generated seven constructs along which the planets vary. In this case, a nine point rating scale has been used, and, in the case of the 'size' (small/large) construct, the smallest planet, Mercury, has been given a rating 1 and the largest planet, Jupiter, a

---

[14] See http://gigi.cpsc.ucalgary.ca/.

[15] In fact, Kelly (1955) describes a number of variations on the general triadic elicitation procedure. More information about these variations can be found in Fransella and Bannister (Fransella & Bannister, 1977).

rating of 9. The other planets have been rated in a comparative manner along the size construct[16]. The analysis has already revealed clusters of both constructs and elements. Thus, Jupiter and Saturn are clustered together at around 84% similarity, Neptune and Uranus at around 88% similarity, and these two pairs of clusters are themselves clustered together at around 79% similarity[17]. An astronomer might well observe that this group of four planets constitutes the gas giants. A new concept – gas giant – has thus been uncovered, which might be distinguished from the other planets; i.e., the rocky or terrestrial planets. Note that Pluto bears very little similarity to other planets in the grid. In fact, it appears to occupy a category all by itself (although it does bear more similarity to the rocky planets than the gas giants). This is clearly interesting given the debate concerning the ontological status of Pluto as a proper planet[18].

Constructs can also be the focus of cluster analysis. With respect to Figure 1, we can see that the constructs relating to temperature and distance from the Sun are clustered, as are the presence of rings and multiple moons. Such associations can reveal causal or other law-like relations in the domain; for example, the relationship between rings and moons may indicate some sort of causal relationship between the two.

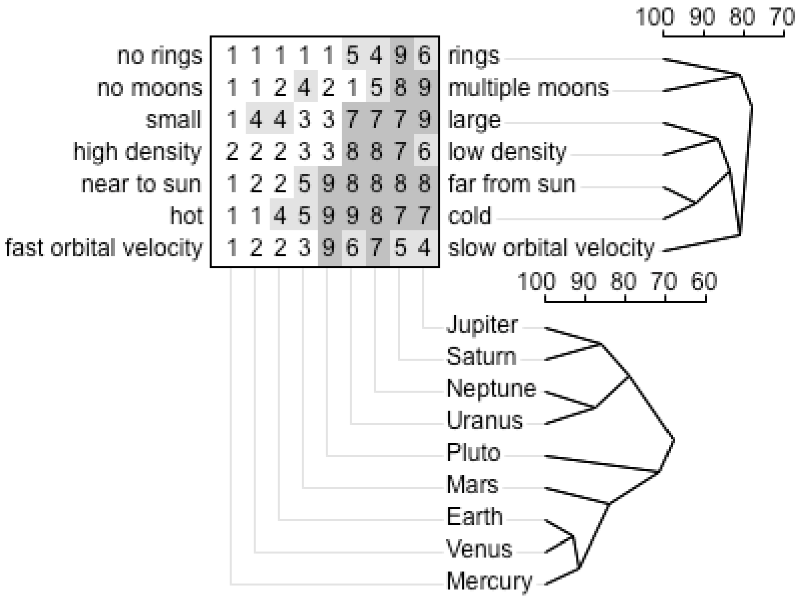

Figure 1: The results of the repertory grid technique applied to the domain of planets (implemented using WebGrid 5).

Variants on the repertory grid technique allow you to run sociogrids (e.g., Shaw, 1980). These allow you to compare one individual's view of the domain with another's, and this can be important in terms of highlighting areas of consensus and difference among experts.

[16] In Figure 1, shading in the matrix is also used to highlight ratings. Heavy shading designates a high value for an element on a construct.

[17] The similarity ratings between the individual elements and element clusters, in this case, are based on the FOCUS algorithm described by Jankowicz and Thomas (1982). The percentage similarity between adjacent elements in the grid is computed as $((-100 * d) / c(n-1)) + 100$, where d is the sum of the absolute differences between the ratings of adjacent elements, c is the number of constructs in the grid (i.e., 7), and n is the largest rating possible (i.e., 9).

[18] See http://news.bbc.co.uk/1/hi/5282440.stm.

Laddered Grids

Another somewhat contrived technique that you will need to explain carefully to the expert before starting is the laddered grid technique. As part of this technique the expert and elicitor construct a graphical representation of the domain in terms of the relations between domain or problem solving elements. The result is a two-dimensional, hierarchically-structured graph where nodes are connected by labelled arcs. No extra elicitation method is used here; expert and elicitor construct the graph together by negotiation.

In using the technique the elicitor enters the conceptual map of the domain (see 'Concept Mapping and Process Mapping') at some point and then attempts to move around it with the expert. A formal specification of how we use the technique is shown below together with an example of its use.

- Start the expert off with a seed item.
- Move around the domain map using the following prompts:
    - To move DOWN the expert's domain knowledge:
        - Can you give examples of <ITEM>?
    - To move ACROSS the expert's domain knowledge:
        - What alternative examples of <CLASS> are there to <ITEM>?
    - To move UP the expert's domain knowledge:
        - What have <SAME LEVEL ITEMS> got in common?
        - What are <SAME LEVEL ITEMS> examples of?
    - To elicit essential properties of an item:
        - How can you tell it is <ITEM> ?
    - To discriminate items:
        - What is the key difference between <ITEM 1> and <ITEM 2>?

The elicitor may move around the knowledge map in any order which seems appropriate or convenient. As the session progresses, the elicitor keeps track of the elicited knowledge by drawing up a network on a large piece of paper, or, if computer supported, via some other graphical characterisation. This representation allows the elicitor to make decisions (or ask questions) about what constitutes higher or lower order elements in the domain and what differences exist between elements in the network. In order to give the reader a flavour of the technique, there follows an extract from a laddered grid elicitation session. Once again, the knowledge domain is geology.

KE: **So how could you tell something was dacite?**
EX: **Well + examine the fresh surface and the weathered surfaces first + looking at grain size, the relationship between the grains**
KE: **Can I just stop you there. What type of grain size is it?**
EX: **Coarse, medium, fine grain, oh, you want me to actually say what dacite is?**
KE: **The grain, in dacite what would it be?**
EX: **Er + medium grained.**
KE: **Medium grained, right. So can you give me other examples of medium grained rocks?**
EX: **Medium grained rocks + dolerite... Granodiorite as well... And we'll stay with that.**
KE: **Right, erm, what alternative is there to a medium grained rock?**
EX: **Well, you can have a coarse grained one or a fine grained one, those are sort of the three major ones.**
KE: **Right, can you give me examples of coarse grained rocks?**
EX: **Er, gabbro, granite... hmm, yeah, those two.**

KE: **And any examples of fine-grained rocks?**
EX: **Er, basalt... er andesite, trachyte...microgranite as well.**
KE: **Right, erm so. What about others**
EX: **Some of these are sort of a metamorphic ones where you're going to get large grains in a fine-grained matrix. There are phenocrysts in them, that's what we call the large grains**
KE: **Is, is there a word for that kind of texture or?**
EX: **Porphyritic mixture**
KE: **Can you give me the examples of the porphyritics...**
EX: **Nepheline-syenite, oh and Kentallenite**
KE: **How would you go about telling the difference between dolerite and granodiorite? What is the key difference?**
EX: **Whether it's got quartz or hasn't got quartz or the percentage of quartz present will define whether it's an acidic rock or a basic rock, basic not having any quartz in it at all, and then er if there's a low amount, that's going to be an intermediate rock**
KE: **Which, which are the intermediate?**
EX: **Dacite + you've got high quartz are granite, microgranite, and andesite, and no quartz gabbro, basalt, dolerite and trachyte, intermediate dacite.**

In the course of this laddered grid interview the elicitor drew up a hierarchical representation of the domain as shown in Figure 2. This is only one of a number of representations that could have been made. In this case the concepts of fine, medium and coarse grained rocks have been understood to be classes of rock type. Similarly the concept of an acidic, intermediate or basic rock has been treated as a class of rock type. However, the grain size and acidity (amount of quartz) could have been represented as properties of the particular rock types.

This hierarchy gives rise to the following set of rules that could be included in the knowledge base of a knowledge intensive system for geological rock classification.

```
IF      the rock is of medium grain size
AND     the rock is intermediate
THEN    the rock may be dacite

IF      the rock is of coarse grain size
AND     the rock is acidic
THEN    the rock may be granite

IF      the rock is of coarse grain size
AND     the rock is basic
THEN    the rock may be gabbro
```

As is the case with many knowledge elicitation techniques, it helps to keep an audio record of the session for future review or transcription. Laddering is an excellent way of carrying out a structured interview. In addition, it is a technique that can be applied to a variety of knowledge types besides concepts; for example, actions, tasks, goals, resources, and so on can be the subject of a laddered grid knowledge elicitation session.

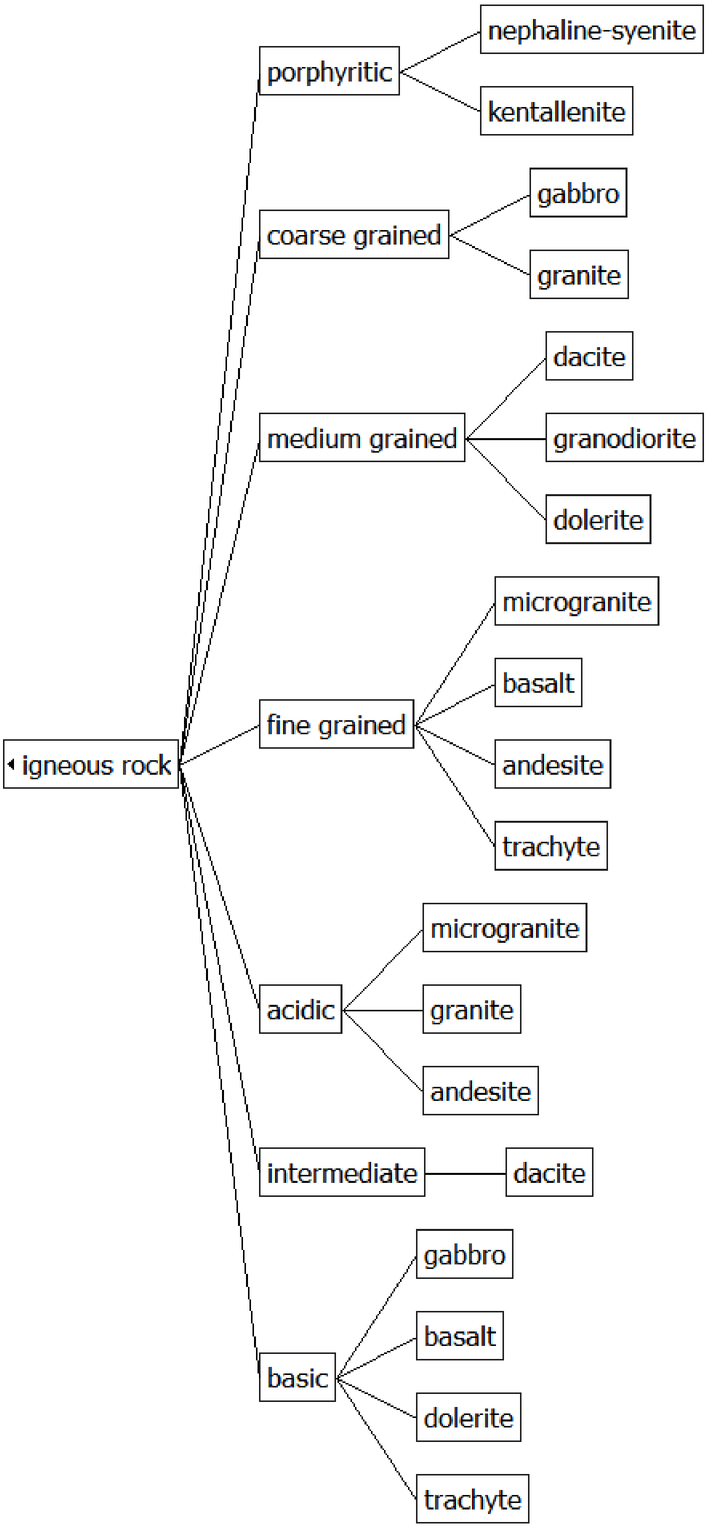

Figure 2: Example of a laddered grid in the geology domain (this grid was developed using the Ladder Tool that is available as part of the PCPACK knowledge editing toolkit[19]).

Limited Information Task

A technique which can prove an excellent complement to the methods already outlined is a technique called the *limited information task* (Hoffman, 1987) or *20 questions* (Grover, 1983). Using

[19] See http://www.tacitconnexions.com/.

this technique, the expert is provided with little or no information about a particular problem to be solved, and the expert must therefore ask the elicitor for specific information that will be required to solve the problem. The information that is requested, along with the order in which it is requested, provides the elicitor with an insight into the expert's problem solving strategy. One difficulty with this method is that the elicitor needs a good understanding of the domain in order to make sense of the expert's questions and to provide meaningful responses. The elicitor should have forearmed themselves with a problem from the domain together with a crib sheet of appropriate responses to the questions.

In one of the versions of the limited information task that we use, we tell the expert that the elicitor has a scenario in mind and the expert must determine what it is. The scenario might represent a problem, a solution or a problem context. The expert is told that they may ask the elicitor for more information, though what the elicitor gives back is terse (e.g., it may consist of simple 'yes' or 'no' responses) and does not go much beyond what was asked for in the question. The expert may be asked to explain why each of the questions was asked.

The limited information task is useful because it provides information about the relative importance of particular items of information as part of a problem-solving process. Often traditional knowledge-based systems gather the right data but the order in which the data is gathered and used can be very different from how an expert works. This can decrease the acceptability of any implemented system if other experts are to use it, and it also has consequences for the intelligibility of any explanations the system offers in terms of a retrace of its steps to a solution.

The drawbacks to this technique are that the elicitor needs to have constructed plausible scenarios, and the elicitor has to be able to cope with the questions that are asked. The experts themselves are sometimes uncomfortable with this technique; this may well have to do with the fact that, as with other contrived techniques, it is not a natural means of manifesting expertise. In addition, whilst a few scenarios may reveal some of the general rules in a domain, the elicitation is very case specific. In order to get a broad range of knowledge, many different scenarios need to be constructed and used.

An interesting variation on this method is a form of telephone consultancy. Here we take two domain experts and place them at opposite ends of a table and ask them to imagine that one is a 'client' who is ringing up the other, a 'consultant', to ask for advice concerning a particular problem. They then engage in a conversation in which the 'consultant' tries to elicit the nature and context of the problem, and finally attempts to offer appropriate advice. In this variation of the limited information task you can rely on one of the experts to generate interesting cases. In addition, the expert playing the role of the 'client' can provide appropriate responses to the 'consultant's' enquiries. The only drawback is that sometimes experts construct extremely difficult cases for each other in order to test each other's mettle!

Concept Mapping and Process Mapping

Concept mapping and process mapping are both examples of diagramming techniques (Milton, 2012) that focus on the structure of conceptual and procedural knowledge, respectively. Concept mapping is probably one of the most widely used knowledge elicitation techniques, in part due to the popularity of the CmapTools software that was developed by the Institute for Human and Machine Cognition (IHMC) (see below).

The artefacts that result from concept mapping (i.e., concept maps) are collections of propositions that are commonly displayed as a 2-dimensional network of labelled grids and nodes (see Figure 3). Concept mapping has been reported to be a very efficient knowledge elicitation technique with the technique yielding an average of two useful propositions per session minute (Hoffman et al., 2001). The technique has also demonstrated its utility in a variety of disparate domains, with meteorology (Hoffman & Lintern, 2006) and intelligence analysis (Derbentseva & Mandel, 2011) serving as just a couple of examples.

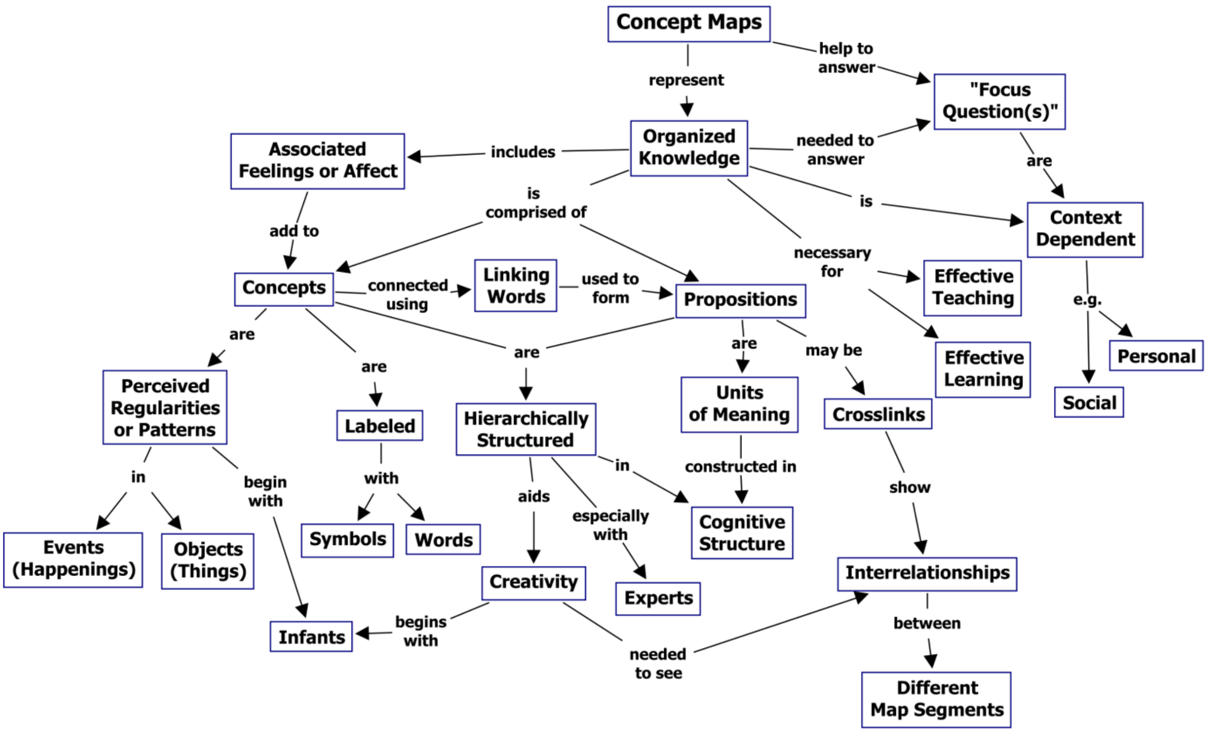

Figure 3: A concept map intended to explore the notion of a 'concept map' (source: http://cmap.ihmc.us).

Both concept and process mapping can be performed with popular knowledge acquisition toolkits, such as PCPACK and CmapTools (see below). In practice, however, CmapTools, tends to be used primarily for concept mapping, while the features of the Diagram Tool within PCPACK make it ideally suited for process mapping. One of the features of the PCPACK Diagram Tool is a capability to 'drill down' into a process, detailing the structure of its constituent subprocesses. It also provides a range of process-oriented graphical notations that are consistent with those seen in popular modelling paradigms (e.g., UML activity diagrams).

Classification of Knowledge Elicitation Techniques

We have now sampled some of the major approaches to knowledge elicitation and, where appropriate, given a detailed description of techniques that are likely to be of use. There are many variants on the methods we have described. Below we have provided a taxonomy of methods with which we are familiar together with a primary reference for each one.

- Non-contrived/Natural
    - Interviews
        - Structured
            - Fixed Probe (Shadbolt & Burton, 1990a; Wood & Ford, 1993)

            - Focused Interviews (Hart, 1986; Scott et al., 1991)
            - Forward Scenario Simulation (Grover, 1983)
            - Critical Decision Method (Hoffman et al., 1998)
        - Semi-Structured
            - Knowledge Acquisition Grid (LaFrance, 1987)
            - Teach Back (Johnson & Johnson, 1987)
        - Unstructured (Weiss & Kulikowski, 1984)
    - Protocol Analysis
        - Verbal
            - Online (Johnson et al., 1987)
            - Offline (Elstein et al., 1978)
            - Shadowing (Clarke, 1987)
            - Collegial Verbalization (Erlandsson & Jansson, 2007)
        - Behavioural (Ericsson & Simon, 1996)
- Contrived
    - Diagramming
        - Laddered Grid (Corbridge et al., 1994; Walker & Crittenden, 2012)
        - Concept Mapping (Novak & Cañas, 2006)
        - Process Mapping (Milton, 2012)
    - Sorting and Rating
        - Concept Sorting (Gammack, 1987)
        - Repertory Grid (Shaw & Gaines, 1987)
        - Pathfinder (Schvaneveldt et al., 1985)
    - Constrained Processing
        - Limited-Information Task (Hoffman, 1987)
        - 20 Questions (Grover, 1983)

This is, of course, only one possible structure for a taxonomy of knowledge elicitation techniques. A number of alternative classifications appear in the literature based on a variety of perspectives, such as the nature of the interaction between elicitor and expert, the type of knowledge (conceptual vs. procedural) elicited from the expert, and the kind of materials required by the task or delivered as outputs from the task. Gavrilova and Andreeva (2012) categorize knowledge methods based on the level of involvement of an expert and an elicitor and type of interaction/collaboration between them. They distinguish between 'active' (analyst-leading) and 'passive' (expert-leading) techniques, where an active technique requires "the active position of an analyst, who 'pulls' the knowledge from the expert with the help of specially prepared questions" and a passive technique is a technique in which "the analyst's interference into the process in which the expert is engaged is very limited" (Gavrilova & Andreeva, 2012, p. 529). As another example of the taxonomic organization of knowledge elicitation techniques, Milton (2012) organizes knowledge elicitation techniques into three categories, namely questioning techniques (e.g., laddering), task-based techniques (e.g., concept sorting) and diagramming techniques (e.g., concept mapping).

None of the existing taxonomies (including the one presented here) are necessarily complete with respect to the range of knowledge elicitation techniques that have been discussed in the literature. In part, this stems from the fact that the goals of knowledge elicitation and the kind of task contexts in which knowledge elicitation is deemed important have changed over time. As pointed out by Hoffman and Lintern (2006), the methodology of knowledge elicitation could be folded into the broader methodology of cognitive task analysis, which is a focal point for human factors and cognitive systems engineering. This serves to blur the distinction between knowledge engineering and cognitive engineering, and it tends to result in a greatly expanded palette of knowledge elicitation methods. A variety of ethnographic methods, for example, could be seen as forms of knowledge elicitation (see Hutchins, 1995).

Other techniques that are sometimes presented as knowledge elicitation techniques are the various methods associated with data mining (Witten & Eibe, 2005) machine learning (Mitchell, 1997) and rule induction (Hart, 1986). These techniques are not covered in detail here because they are not techniques that are typically used in conjunction with domain experts. There are, however, some exceptions. In particular, there have been a number of recent attempts to combine expert input with machine learning techniques in order to improve the quality of the knowledge that results from the machine learning process. Typically, the kind of outputs delivered by machine learning tend to prove difficult for experts to understand and extend, and this presents problems in terms of the maintenance of the knowledge base and the trust that experts place in automated decision-making processes. Argument-based machine learning (ABML) is a technique which was developed to address some of these issues (Mozina et al., 2008). The technique is intended to combine expert knowledge with machine learning processes, and it requires the expert to explain the reasons for decisions in particular cases. Groznik et al (2013) describe a recent application of the technique, wherein ABML is used to elicit knowledge from neurologists in order to develop a decision support system concerned with neurological diagnoses.

Experts and Expertise

As the source of much of the knowledge that is captured as part of a knowledge engineering initiative, domain experts are a critical focus of attention for those involved in knowledge engineering. Failing to pay adequate attention to the differences among experts, as well as the level of expertise they possess, is likely to have a profound effect on the efficiency of the knowledge elicitation process, as well as the quality of the knowledge that gets elicited.

One of the first challenges that must be addressed in any knowledge engineering project is the identification of individuals with the relevant expertise. In some cases, it may be obvious who the experts are within a given domain; in other cases, however, it may not at all be clear how experts should be identified. Factors such as the possession of professional qualifications, experience and occupational position, as well as the results of testing and screening processes, may all be used as the basis for expert identification; however, none of these methods is without its problems (Farrington-Darby & Wilson, 2006). For example, the position held by an individual is a commonly used criterion for expert selection; however, the reasons for individuals being awarded a position within a given occupational setting may have very little to do with their actual expertise (see Farrington-Darby & Wilson, 2006). In terms of experience, a general rule of thumb is that expertise develops after about 10,000 hours of practice. Recent research, however, has suggested that expertise in some domains, such as weather forecasting, may take considerably longer (see Hoffman & Lintern, 2006). In spite of the difficulties, it is worth spending some time considering who and who is not an expert. As Burton et al (1990) note:

> *"Inadequate expertise is likely to continue to be a problem for those working in applied settings. We suggest that considerable time be put into the original selection of an expert. External validation of an expert's suitability will save considerable time and wasted effort in future sessions." (p. 177)*

Once experts have been identified, it is important to consider the differences between experts, as well as the nature of the expertise they manifest. Experts can be differentiated in a number of ways; however, one scheme that we have found useful in practice is to distinguish between three kinds of experts: the *academic*, the *practitioner*, and the *samurai*. Each of these types of expert differs along a number of dimensions[20]. These include the outcome of their expert deliberations, the problem solving environment they work in, the state of the knowledge they possess (both its internal structure and its external manifestation), their status and responsibilities, their source of information, and the nature of their training.

How are we to tell these different types of expert apart when we encounter them? The academic type regards their domain as having a logically organised structure. Generalisations over the laws and behaviour of the domain are important to them; theoretical understanding is prized. Part of the function of such experts may be to explicate, clarify and teach others. They thus talk a lot about their domains. They may feel an obligation to present a consistent story both for pedagogic and professional reasons. Their knowledge is likely to be well structured and accessible. These experts may suppose that the outcome of their deliberations should be the correct solution of a problem. They believe that the problem can be solved by the appropriate application of theory. They may, however, be remote from everyday problem solving.

The practitioner class, on the other hand, are engaged in constant day-to-day problem solving in their domain. For them, specific problems and events are the reality. Their practice may often be implicit, and what they desire as an outcome is a decision that works within the constraints and resource limitations in which they are working. It may be that the generalised theory of the academic is poorly represented and articulated by the practitioner. For the practitioner, heuristics may dominate and theory is sometimes thin on the ground.

The samurai is a pure performance expert – their only reality is the performance of action to secure an optimal performance. Practice is often the only training, and responses are often automatic.

One can see this sort of distinction between experts in any complex domain. Consider, for example, medical domains where we have professors of the subject, busy doctors working the wards, and medical ancillary staff performing many important but repetitive clinical activities.

The knowledge elicitor must be alert to these differences because the various types of expert will perform very differently in knowledge elicitation situations. The academic will be concerned to demonstrate mastery of the theory. They will devote much effort to characterising the scope and limitations of the domain theory. Practitioners, on the other hand, are driven by the cases they are solving from day to day. They have often compiled or routinised any declarative descriptions of the theory that supposedly underlies their problem solving. The performance samurai will more often

[20] In practice, of course, experts do not tend to fall in one or other categories; rather, they embody elements of all three types of expert.

than not turn any knowledge elicitation interaction into a concrete performance of the task, simply exhibiting their skill.

Another important distinction between experts is with respect to their level of expertise. A number of models of expertise development have been proposed within the cognitive science and human factors communities, and these may serve as the basis for a second dimension along which experts can be classified – one that is largely orthogonal to the previously mentioned distinction between academics, practitioners and samurais. One model, proposed by Dreyfus and Dreyfus (1986), suggests that expertise develops via the progression through five sequential stages: novice, advanced beginner, competent, proficient and expert. The transition between these stages is assumed to depend on the accumulation of situated practical experience within the relevant domain. Another classification scheme derives from the Craft Guilds of the Middle Ages (Hoffman, 1998; Hoffman et al., 1995). In this case, the developmental scale ranges from a 'Naivette' (i.e., one who is totally ignorant of a domain) through to a 'Master' who is regarded as one of an elite group of experts – the expert of experts.

Recognizing the developmental stage of an expert can be important for the purposes of knowledge elicitation. Clearly, individuals with well-developed levels of expertise are important targets for knowledge elicitation, since they are the ones who are likely to possess the greatest amount of domain-relevant knowledge. Having said that, expertise development tends to be associated with a shift from explicit to tacit knowledge, and thus individuals at different points on the developmental trajectory from novice to master may be differentially responsive to particular kinds of knowledge elicitation technique. In certain kinds of domains, for example, a 'Journeyman' or 'Expert' may have greater conscious access to domain-relevant knowledge as compared to a 'Master'. For this reason, techniques such as interviews may yield more information from those at intermediate levels of expertise development as compared to those further along the developmental scale.

Clearly, the expertise embodied by experts is not of a homogenous type (Feltovich et al., 1997). In constructing any knowledge-intensive system, it is likely that very different types of knowledge will be uncovered, and these are likely to have very different roles in the system under development. In general, we can distinguish between four kinds of knowledge (three of these – the domain, inference and task knowledge categories – are explicitly represented within knowledge engineering methodologies, such as the CommonKADS methodology (see Schreiber et al., 2000)):

- **Domain Knowledge.** Firstly, we can distinguish what is called domain knowledge. This term is being used in the narrow sense of knowledge that describes the concepts and elements in the domain and relations between them. This sort of knowledge is sometimes referred to as declarative knowledge – it describes what is known about things in the domain.
- **Inference Knowledge.** There is also knowledge and expertise that has to do with what we might call the inference level. This is knowledge about how the components of expertise are to be organised and used in the overall system. It tells us the type of inferences that will be made and what role knowledge will play in those inferences. This is quite a high level description of expert behaviour and may often be implicit in expert practice.
- **Task Knowledge.** Another type of expert knowledge is the task level. This is sometimes called procedural knowledge. This is knowledge concerned with goals, sub-goals, tasks and

sub-tasks. Thus, in a classification task there may exist a number of tasks to perform in a particular order so as to utilise the domain level knowledge appropriately.

- **Strategic Knowledge.** Finally, there is a level of expert knowledge referred to as strategic knowledge. This is information that monitors and controls the overall problem solving process.

Within any of these categories of knowledge, the information may be either implicit or explicit. Thus, in some domains, the expert may have no real notion of the strategic knowledge they are following, whilst in others this knowledge is very much at the forefront of their deliberations.

Methodologies and Programmes

We turn next to the question of how knowledge elicitation techniques should be assembled to form a programme of knowledge acquisition. There are a number of articles and books on how to undertake knowledge elicitation as part of knowledge engineering project. Milton (2007), for example, describes the processes involved in knowledge elicitation and modelling in the form of a step-by-step guide. The choice as to which knowledge elicitation technique to use in any particular situation is guided by a variety of criteria, including the characteristics of the domain, of nature of the domain expert, and the requirements associated with the proposed knowledge system solution. Furthermore, it is clear that some techniques are going to be more costly in terms of time with the expert, or else the effort required for the analysis of elicited material. In order to select an appropriate knowledge elicitation technique, one needs to understand which method best fits the particular problem and situation. This calls for empirical evaluations of each of the techniques with respect to factors such as the nature of experts and their associated expertise. Although there are a variety of difficulties associated with the evaluation of knowledge elicitation techniques (Shadbolt et al., 1999), the available research has provided some general conclusions as to their relative efficacy (Burton et al., 1987; Burton et al., 1990; Hoffman et al., 1995; Shadbolt & Burton, 1990b). It has also provided some guidelines as to when to use particular kinds of knowledge elicitation technique. Gammack and Young (1985), for example, offer a mapping of knowledge techniques onto domain types. Their analysis requires that domain knowledge be separated into different categories, and they provide suggestions about which techniques are most likely to be effective within each category.

One of the main criteria for choosing between different techniques within a programme of knowledge elicitation is likely to be the type of knowledge that needs to be elicited. In this respect, the distinction between explicit and tacit knowledge has proven to be of significant interest. Different knowledge elicitation techniques are thus deemed to be differentially effective at eliciting explicit or tacit knowledge (see Figure 4). Another knowledge dimension that is often seen as important is the distinction between conceptual and procedural knowledge. Here, techniques such as process mapping are considered to be more effective for the elicitation of conceptual knowledge and techniques such as concept mapping and concept sorting are deemed to be more effective for the elicitation of conceptual knowledge. Figure 4 summarizes the differential suitability of a number of knowledge elicitation techniques with respect to these two knowledge dimensions (i.e., explicit/implicit and conceptual/procedural).

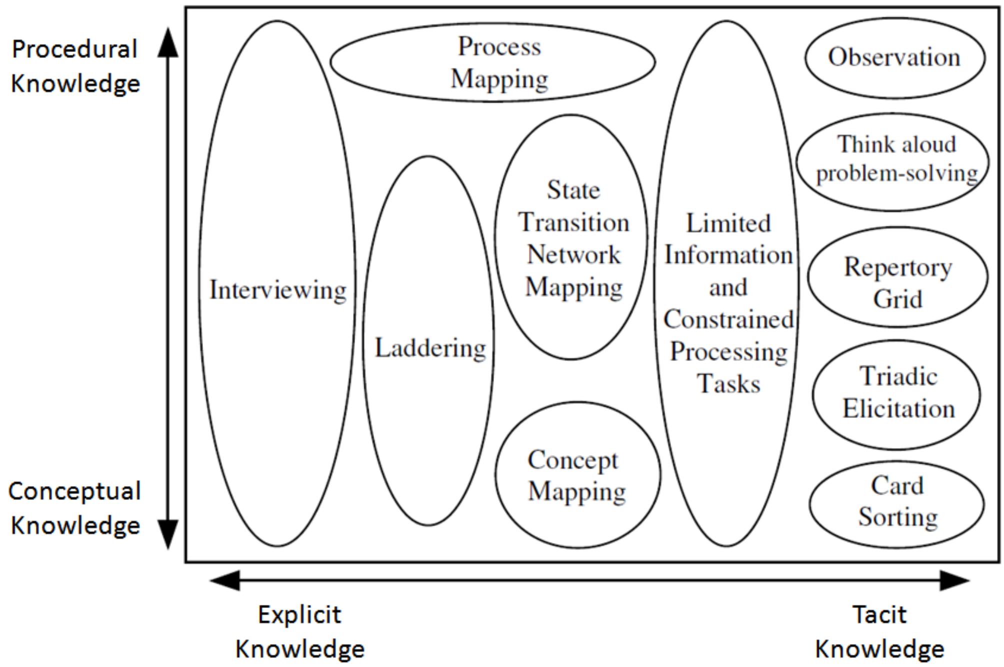

Figure 4: Differential utility of knowledge elicitation techniques with respect to the elicitation of different kinds of knowledge (source: Milton, 2003).

The notion that different knowledge elicitation methods are differentially effective at eliciting particular kinds of knowledge forms part of what has become known as the differential access hypothesis (Hoffman et al., 1995). Although some empirical support for the hypothesis has been found, a strong version of the differential access hypothesis (namely the idea that certain kinds of knowledge can *only* be elicited via the use of particular techniques) remains a point of contention within the knowledge engineering community (Hoffman & Lintern, 2006). When it comes to the notion of tacit knowledge, for example, Hoffman and Lintern (2006) suggest that the different knowledge elicitation techniques establish different conditions under which the verbalization of tacit knowledge is more or less likely. They suggest knowledge elicitation techniques should be seen as 'scaffolds' that support the expression or communication of knowledge. With this in mind, the key aim in knowledge elicitation becomes one of establishing the right kind of conditions under which experts can articulate, or otherwise communicate, their expertise. These kind of conditions are clearly influenced by the kind of technique that is used, since each technique is associated with different forms of social interaction, access to mnemonic cues, the use of different diagrammatic representations, and so on. With this in mind, it might be argued that something like tacit knowledge should not be seen as a form of knowledge that can never, in principle, be verbalized by experts; rather, it should be seen as a form of knowledge that is more easily articulated in certain situations as opposed to others. This, suggest Hoffman and Lintern (2006), has shifted the debate from a consideration of differential access to one of differential utility when it comes to the selection of knowledge elicitation techniques:

> *"The hypothetical problem of differential access has given way to a practical consideration of differential utility. Any given method might be more useful for certain purposes, might be more applicable to certain domains, or might be more useful with certain experts having certain cognitive styles. In other words, each knowledge elicitation method has its strengths and weaknesses. Some of these are purely*

methodological or procedural (e.g., transcription and protocol analysis takes a long time), but some relate to the content of what is elicited." (Hoffman & Lintern, 2006, pp. 216-217)

In spite of this change in perspective, however, it should be clear that there remains a compelling reason to exploit a variety of techniques within any programme of knowledge elicitation. Even when it appears that only one particular body of knowledge is being dealt with – one which shows no internal differentiation with respect to (e.g.) explicit/tacit or procedural/conceptual distinctions – it is still advisable to use a variety of techniques. One reason for this stems from the possibility that the knowledge elicited by different techniques may predict actual performance to a greater or lesser extent. Studies have thus found that the content of verbal reports and the details of actual performance are not always the same. Cooke and Breedin (1994), for example, discovered a dissociation between the written explanations that were offered for physics trajectory problems and the actual predictions that were made concerning those trajectories. These results suggest that the results of multiple techniques should be compared with each other in order to evaluate the connection between knowledge and performance.

One of the factors that may inform the design of knowledge elicitation programmes is the methodological framework in which knowledge elicitation and modelling is undertaken. Although a number of methodologies exist for the development of ontologies within the context of the Semantic Web (e.g., Sure et al., 2003), such methodologies typically ignore the early steps of the knowledge engineering process and place little emphasis on knowledge elicitation. CommonKADS (Schreiber et al., 2000) is one of the few methodologies that explicitly incorporates the use of knowledge elicitation techniques. One way in which CommonKADS helps to structure the knowledge elicitation activity is by the distinction it makes between domain, task and inference knowledge (see above). These different kinds of knowledge are represented as distinct 'layers' within a CommonKADS knowledge model specification, and mappings are established between the layers (e.g., between elements of inference and domain knowledge) in order to flexibly link different kinds of knowledge together in the context of a particular knowledge solution (see Figure 6). CommonKADS also offers a range of reusable components that can be used as points of departure for the selection and implementation of knowledge elicitation activities. The reusable components include a set of domain schemas, a catalogue of inference types and a library of task templates. These are useful not only in terms of improving the efficiency of the modelling process, they also serve to focus attention on the kinds of knowledge that needs to be acquired in the context of a particular kind of knowledge-based activity. Each of the CommonKADS task templates (see Figure 5) thus highlights the typical pattern of inferences that are associated with each kind of task, and it also links these inferences with particular bodies of domain knowledge (e.g., concepts) (see Figure 6). This kind of information can be extremely valuable in terms of highlighting the kind of knowledge to elicit and the kind of behavioural patterns to look for in expert performance.

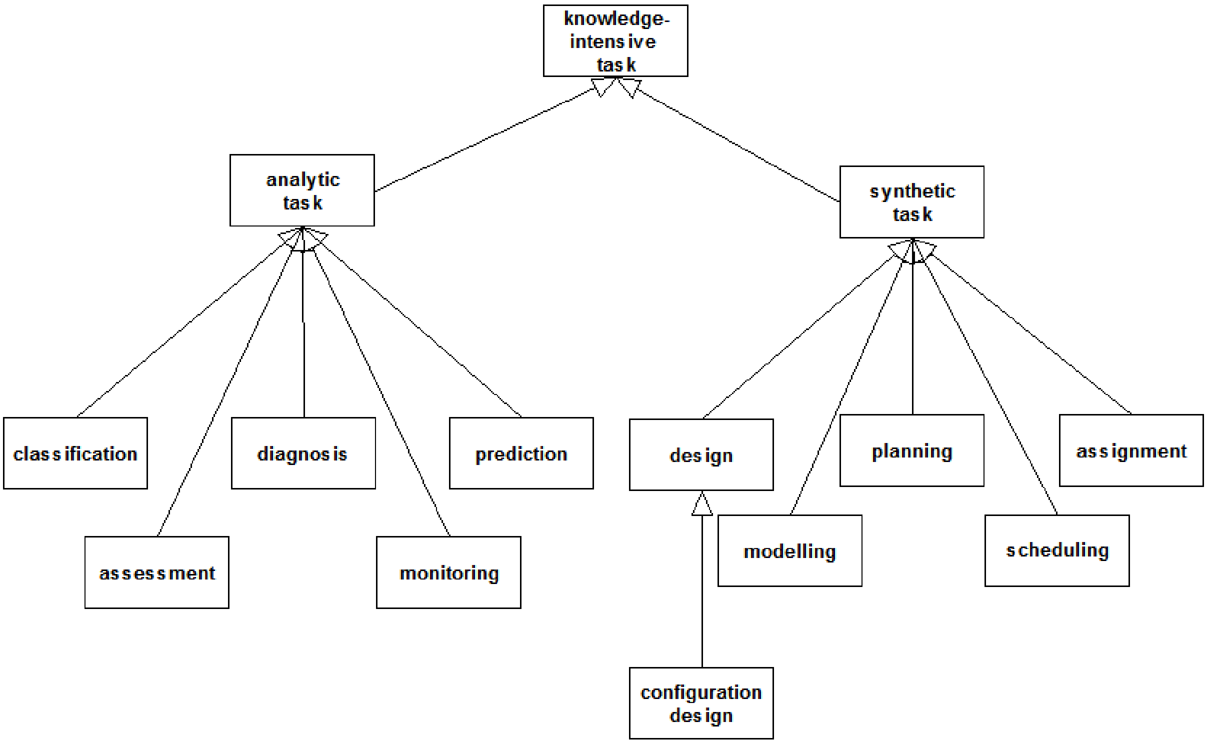

Figure 5: Knowledge-intensive tasks recognized by the CommonKADS methodology. Each of these tasks are associated with default inferences, control structures and template domain schemas.

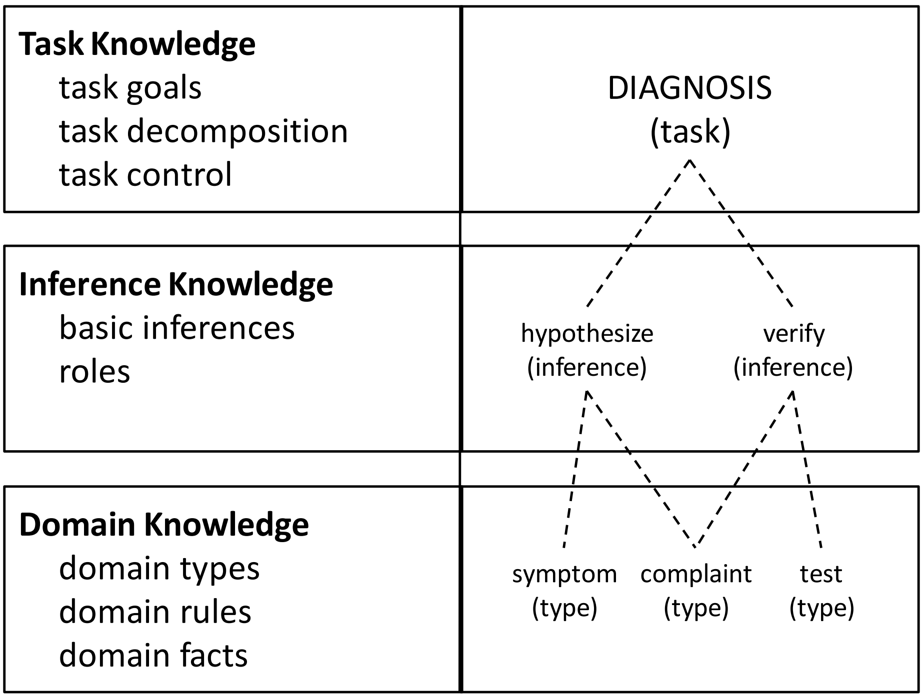

Figure 6: Linkages between the various layers of the CommonKADS knowledge model for a particular kind of knowledge-intensive task – in this case, diagnosis. Each task is associated with specific types of inferences that are themselves linked with particular elements at the level of domain knowledge.

Knowledge Elicitation Tools

As indicated in the previous section, the attempt to improve our understanding of the conditions under which knowledge elicitation techniques are most effective, as well as how to adapt those techniques within specific knowledge elicitation programs, is the focus of recent and ongoing research attention. Another focus of attention concerns the development of software tools to support the knowledge elicitation process.

The software tools that are presented in this section – PCPACK, Protege and CmapTools – have a long history of development and use within the knowledge acquisition community. The recent development of these tools has been strongly influenced by the Web[21] and, in particular, the Semantic Web. All the tools have thus been extended in particular ways to accommodate the representational frameworks associated with the Semantic Web. Recent versions of PCPACK thus provide support for RDF export, while knowledge elicitation plug-ins for Protégé interoperate with the Protégé-OWL plug-in in order to provide support for knowledge elicitation in the context of ontology development (Wang et al., 2006). There has also been a recent effort to extend CmapTools in order to provide support for the visualization and editing of OWL ontologies (Eskridge & Hoffman, 2012; Hayes et al., 2005).

PCPACK

PCPACK is an integrated suite of knowledge elicitation tools that has a long history of use within the knowledge engineering community (Schreiber et al., 2000, chapter 8). Early versions of PCPACK provided computerized support for many of the knowledge elicitation techniques described earlier in this chapter (O'Hara et al., 1998; Shadbolt & Milton, 1999); however, more recent versions of the software have settled on those tools that provide the greatest level of support to those engaged in corporate knowledge engineering and management initiatives. The current version of PCPACK is maintained and distributed by Tacit Connexions, and a fully operational demonstration version of the software can be downloaded from the Tacit Connexions website[22]. PCPACK includes a variety of tools to support knowledge elicitation and modelling, and all of these tools are integrated with a single knowledge repository such that any changes to the knowledge base made using one tool are immediately reflected in other components of the tool suite. Among the tools included with PCPACK is the Ladder Tool, which is used for creating laddered grids of various kinds (e.g., taxonomic and meronymic concept hierarchies); a Diagram Tool, which can be used for process and concept mapping; a Protocol Tool, which can be used for protocol analysis; an Annotation Tool, which is used to provide an HTML editing interface for knowledge objects; and a Publisher Tool, which enables knowledge models to be published as Web-accessible 'Knowledge Webs'. Other tools provide support for RDF import/export, annotation template management and the matrix-based editing of knowledge object properties and relationships. Figure 7 shows a screenshot of one of the PCPACK tools, namely the Ladder Tool.

[21] For example, all the tools reviewed in this section support the publication of HTML versions of knowledge models. This enables the models to be accessed in the context of the conventional, document-centered Web, as well as the more recent data-centric Web or Web of Linked Open Data (see Heath & Bizer, 2011).
[22] See http://www.tacitconnexions.com/PCPACK%20download%20promo%20page.htm.

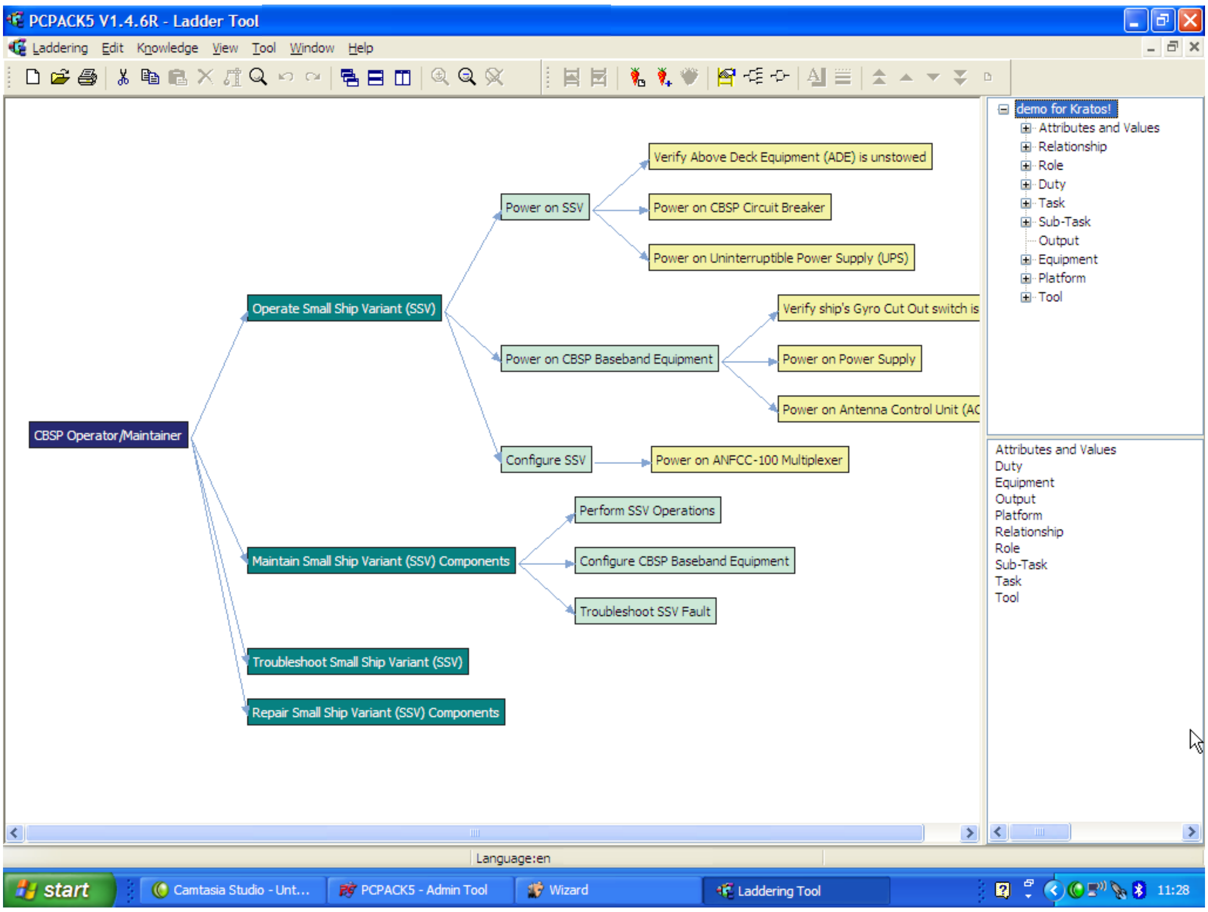

Figure 7: The PCPACK Ladder Tool.

Protégé

As with PCPACK, the Protégé knowledge editor[23] has a long history of use within the knowledge engineering community. As a flexible and customizable knowledge editing environment, Protégé is able to provide support for a variety of knowledge engineering methodologies and modelling frameworks. However, ever since the advent of the Semantic Web and the development of the Protégé-OWL plug-in (Knublauch et al., 2004; Knublauch et al., 2005) it is probably fair to say that the primary use of the tool is to develop (OWL-based) ontologies.

Unlike PCPACK, Protégé does not provide an integrated suite of knowledge elicitation tools as standard. The primary purpose of the tool is to support the editing of elicited knowledge rather than to support the process of knowledge elicitation itself. There have, however, been a number of attempts to provide computerized versions of the knowledge elicitation techniques as plug-ins to the Protégé environment. Wang et al (2006) thus describe the attempt to implement card sorting and laddering plug-ins in order to support the use of knowledge elicitation techniques as part of the ontology development process.

Protégé is available as a free, open-source download from the Protégé website. It has typically been implemented as a Java-based desktop application; however, recent development efforts have seen the release of WebProtégé (Tudorache et al., 2013), which is a lightweight, Web-based version of the original Protégé environment.

[23] See http://protege.stanford.edu/.

CmapTools

Another widely used knowledge elicitation and knowledge modelling tool is CmapTools, which is developed and maintained by the Institute for Human and Machine Cognition (IHMC)[24]. CmapTools provides support for the development of concept maps, which can be developed in conjunction with a domain expert and then published on the Web. The tool enables the user to establish links between concept maps, which are collectively referred to as a 'knowledge model'. In addition, links to other resources, such as images, videos, text documents, and so on, can be associated with any node in the concept map diagram.

As with other knowledge engineering technologies, the development of CmapTools is currently being influenced by the Semantic Web. Researchers at the IHMC are currently exploring the potential to combine concept mapping capabilities with the representational formalisms encountered in the context of the Semantic Web (Eskridge & Hoffman, 2012; Hayes et al., 2005). Ultimately, this effort will enable the CmapTools concept mapping system to be used for the construction, sharing and visualization of OWL ontologies.

Knowledge Elicitation, Knowledge Engineering and the World Wide Web

The Web and the Semantic Web have had a profound impact on the discipline of knowledge engineering (Gil, 2011; Schreiber, 2013). In many cases, the Web now serves as both the starting point (e.g., by providing access to a rich source of domain-relevant knowledge and information) as well as the end point (e.g., by serving as a platform for knowledge publication and distribution) for knowledge engineering efforts. The specifications and recommendations that have emerged in the context of the Semantic Web initiative (Berners-Lee et al., 2001; Shadbolt et al., 2006) (for example, RDF, RDF-S and OWL) have served as a Procrustean bed that has affected nearly all knowledge representation frameworks and knowledge engineering technologies. The Web is also an environment that can be used for the purposes of knowledge elicitation, especially when the elicitation effort requires collaboration from multiple stakeholders. Finally, of course, the Web serves as an environment for the implementation of a whole variety of intelligent systems and knowledge-based solutions.

Perhaps the most notable feature of the Web, when it comes to knowledge elicitation, is the role that the Web plays as a knowledge source. The Web provides access to a rich range of resources that are relevant to the construction of any prospective knowledge-intensive system. If one takes the domain used throughout much of this chapter – the classification of rocks and minerals – one is able to find a wealth of online resources. These range from dictionaries and definitions of terms, succinct summaries of the processes of rock formation, and extensive online databases. Such resources can serve as an important focus for the initial stages of knowledge elicitation, particularly for purposes of domain familiarization. They can also provide access to a range of materials that can be incorporated into knowledge elicitation exercises (e.g., images of different kinds of rocks can be used as the basis for card sorting exercises).

Web-based resources may also serve as the direct target of knowledge acquisition efforts. Although, such resources by themselves are unlikely to provide all the required information – recall the aphorism 'the gold is not in the documents' – they can yield knowledge structures (for example,

[24] See http://cmap.ihmc.us/.

concept lists) that are subsequently refined and extended in the course of face-to-face knowledge elicitation sessions.

Complementing the use of manual knowledge acquisition methods is the use of a range of advanced knowledge discovery techniques that can be used to extract knowledge from online sources. These kind of automated techniques are vitally important given the scale of the Web and the range of resources that are now available. Information extraction and natural language processing (NLP) technologies are one focus of ongoing research attention in this area (Sarawagi, 2008), as are opinion mining and sentiment analysis techniques (Feldman, 2013; Pang & Lee, 2008). There is also interest in the use of ontology learning techniques to create initial ontological structures from large-scale bodies of domain-relevant information (Maedche & Staab, 2003). These kind of analytic and learning techniques are likely to become all the more important as we move into an era where Linked Open Data assets (see Heath & Bizer, 2011) become increasingly prevalent on the Web.

Web resources may also be used as part of an integrated knowledge acquisition effort that combines Web access with the use of conventional knowledge elicitation techniques and other forms of advanced machine-based processing, such as NLP. Mendonça et al (2012) thus used NLP to isolate initial concepts and then refined these in conjunction with domain experts using a variety of knowledge elicitation techniques (namely interviews, sorting and matrix-based techniques). This was followed by a knowledge validation phase in which the Web was used to support the collaborative validation of elicited knowledge. This study highlights how the Web can be exploited at several stages of the knowledge elicitation process: it can be used as an initial resource to support domain familiarization and extract initial concepts (perhaps using machine-assisted techniques, such as NLP), and it can also be used to validate the elicited knowledge – the knowledge is published on the Web and made available to a global community of experts who can validate and refine the elicited knowledge as a precursor to (e.g.) ontology development. Further research in this area should consider the kind of opportunities the Web makes available for knowledge elicitation and adapt knowledge engineering methodologies to exploit these opportunities.

The main problem, of course, when it comes to use of Web-based resources concerns their varying quality and coverage. The information provided by the sources is often of unknown origin and there is often no prior history with many of the sources that may be used to assess their reputation. One focus of ongoing research within the Web Science community is how to determine whether to trust a particular piece of information provided by a source.

In addition to the use of the Web as a knowledge source, the Web also provides a platform for active knowledge elicitation from individual experts or expert communities. Unfortunately, there are very few examples, at the present time, of Web-based tools that could be used for collaborative knowledge elicitation. Perhaps one reason for this relates to a shift in our appreciation of how the Web can be used as a mechanism for knowledge acquisition. When one looks at examples of large knowledge repositories on the Web – for example, Wikipedia – what one tends to encounter is a system in which knowledge content has emerged as a result of the collaborative efforts of multiple individuals. This has led to our traditional notions of expert-centred knowledge engineering being supplemented with an approach that draws on the contributions of large numbers of users, very few of whom are perhaps regarded as experts within the target domain. The point is that sometimes the actions of a large number of users can yield useful knowledge outputs (although whether these outputs can ever serve as a substitute for the kind of outputs obtained in face-to-face knowledge elicitation sessions with domain experts is currently a moot point). Folksonomies (Wu et al., 2006) represent one example here, as do the structured resources that emerge from the cumulative editing actions of Wikipedia users; e.g., DBpedia (Bizer et al., 2009). In general, there is an increasing recognition of the way in which certain classes of Web-based systems – sometimes referred to as social machines – can be used to leverage the contributions of human user communities, often at large scale. Knowledge acquisition is often a key focus of such systems (Shadbolt, 2013); however, the systems can also (on occasion) yield collective problem-solving performances that parallel those of individual human experts. In such cases, it may be possible to see a social machine as a form of biotechnologically hybrid intelligent system that dynamically exploits the complementary contributions of both human individuals and conventional computing systems.

One final point that is worth reiterating here relates to the way in which the Semantic Web has impacted knowledge engineering efforts. As mentioned previously, many of the tools used for knowledge elicitation have been influenced by the advent of ontology languages that have been developed for the Semantic Web, and the output of many knowledge engineering efforts now consists in the generation of resources (e.g., OWL ontologies) that are compliant with the standards and recommendations of the Semantic Web community. It is tempting to think of the Semantic Web, in this case, as a large-scale knowledge repository that is the distributed counterpart of the more centralized knowledge bases encountered in the era of expert systems development. There are, however, a number of differences between the Semantic Web and conventional knowledge bases, of which the most obvious relate to the heterogeneity, scale, and diverse quality of Semantic Web knowledge content (d'Aquin et al., 2008). It is also fair to say that the content of the Semantic Web tends to be used in a manner that is unlike that seen in the case of conventional expert systems. As Brueker (2013) notes “Ontologies are rarely used as knowledge bases, but rather as (shallow) vocabularies for managing large information repositories” (p. 179). Indeed, as is evidenced by systems such as IBM’s Watson (Ferrucci et al., 2010), intelligence on the Semantic Web is likely to emerge as a result of the ability to exploit large amounts of available data rather than an ability to carry out sophisticated reasoning (d'Aquin et al., 2008). Although Watson does use ontologies for some inferences, its answers are, for the most part, based on sophisticated information retrieval capabilities and the ability to integrate probabilistic evidence from many diverse sources.

The ability to treat the Web as an epistemic resource and press maximal benefit from an ever-expanding quantity of linked data assets is likely to be a key focus area for research into the next generation of intelligent systems. To what extent computational ontologies will play a role in the realization of these capabilities is unclear; however, what is largely beyond dispute is that, in the near future, the Web is likely to serve as means by which human knowledge is made available for a variety of purposes, and, in view of this, the interest in knowledge elicitation and the need for robust knowledge elicitation techniques is likely to continue.

Conclusion

Despite a range of scientific and technical advances (including the continued development of the Web and Semantic Web), the problem of knowledge elicitation remains an important area of research attention and practical application. This chapter has described some of the methods and techniques that are used in this enterprise. We have also sought to provide an indication of the difficulties inherent in doing this kind of work. Knowledge elicitation is itself a form of complex expertise. Experienced knowledge engineers come to recognise the characteristics of expert thinking, and they develop skills that allow them to capture an expert's knowledge despite the many obstacles they face. Continued research into the differential effectiveness of knowledge elicitation techniques in different situations is likely to inform our understanding of how to structure and manage the knowledge acquisition process; however, there really is no substitute for real-world practical experience when it comes to knowledge elicitation. Just as expertise in other areas only comes at the expense of many hours of practical experience within the relevant domain, so a mastery of knowledge elicitation often requires many hours of active engagement in the knowledge elicitation process.

References

Bainbridge, L., & Sanderson, P. (2005) Verbal protocol analysis. In J. Wilson, T. Megaw & N. Corlett (Eds.), *Evaluation of Human Work*. Taylor & Francis, London, UK.

Beilock, S. L., Wierenga, S. A., & Carr, T. H. (2002) Expertise, attention, and memory in sensorimotor skill execution: Impact of novel task constraints on dual-task performance and episodic memory. *The Quarterly Journal of Experimental Psychology: Section A*, 55(4), 1211-1240.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001) The Semantic Web. *Scientific American*, 284(4), 34-43.

Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., & Hellmann, S. (2009) DBpedia-A crystallization point for the Web of Data. *Web Semantics: Science, Services and Agents on the World Wide Web*, 7(3), 154-165.

Brueker, J. (2013) A cognitive science perspective on knowledge acquisition. *International Journal of Human-Computer Studies*, 71(2), 177-183.

Burton, A. M., Shadbolt, N. R., Hedgecock, A. P., & Rugg, G. (1987) A Formal Evaluation of Knowledge Elicitation Techniques for Expert Systems: Domain 1. In D. S. Moralee (Ed.), *Research and Development in Expert Systems IV*. Cambridge University Press, New York, New York, USA.

Burton, A. M., Shadbolt, N. R., Rugg, G., & Hedgecock, A. P. (1990) The efficacy of knowledge elicitation techniques: A comparison across domains and levels of expertise. *Knowledge Acquisition*, 2(2), 167-178.

Caputi, P., & Reddy, P. (1999) A comparison of triadic and dyadic methods of personal construct elicitation. *Journal of Constructivist Psychology*, 12(3), 253-264.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981) Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121-152.

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.) (1988) *The Nature of Expertise*. Erlbaum, Hillsdale, New Jersey, USA.

Chi, M. T. H., Glaser, R., & Rees, E. (1982) Expertise in Problem Solving. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence: 1*. Lawrence Erlbaum, Hillsdale, New Jersey, USA.

Clarke, B. (1987) *Knowledge Acquisition for Real Time Knowledge Based Systems*. First European Workshop on Knowledge Acquisition for Knowledge Based Systems, Reading, England, UK.

Cooke, N. J. (1994) Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies*, 41(6), 801-849.

Cooke, N. J. (1999) Knowledge Elicitation. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay & M. T. H. Chi (Eds.), *Handbook of Applied Cognition* (pp. 479-510). John Wiley & Sons, Ltd, Chichester, England, UK.

Cooke, N. J., & Breedin, S. D. (1994) Constructing naive theories of motion on the fly. *Memory & Cognition*, 22(4), 474-493.

Corbridge, C., Rugg, G., Major, N. P., Shadbolt, N. R., & Burton, A. M. (1994) Laddering: Technique and tool use in knowledge acquisition. *Knowledge Acquisition*, 6(3), 315-341.

Cordingley, E. S. (1989) Knowledge elicitation techniques for knowledge-based systems. In D. Diaper (Ed.), *Knowledge Elicitation: Principles, Techniques and Applications*. John Wiley & Sons, New York, New York, USA.

Crandall, B., Klein, G., & Hoffman, R. R. (2006) *Working Minds: A Practitioner's Guide to Cognitive Task Analysis*. MIT Press, Cambridge, Massachusetts, USA.

d'Aquin, M., Motta, E., Sabou, M., Angeletou, S., Gridinoc, L., Lopez, V., & Guidi, D. (2008) Toward a new generation of Semantic Web applications. *Intelligent Systems*, 23(3), 20-28.

Derbentseva, N., & Mandel, D. R. (2011) *A Concept Map Knowledge Model of Intelligence Analysis.* Defence Research and Development Canada (DRDC), Toronto, Canada. (Ref: TR 2011-077)

Dreyfus, H. L., & Dreyfus, S. E. (1986) *Mind over Machine: The Power of Human Intuition and Expertise in an Era of the Computer*. Free Press, New York, New York, USA.

Duda, J., Gaschnig, J., & Hart, P. (1979) Model design in the PROSPECTOR consultant system for mineral exploration. In D. Michie (Ed.), *Expert Systems in the Micro-Electronic Age*. Edinburgh University Press, Edinburgh, Scotland, UK.

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978) *Medical Problem Solving: An Analysis of Clinical Reasoning* (Vol. 2). Harvard University Press, Cambridge, Massachusetts, USA.

Ericsson, K. A., & Simon, H. A. (1996) *Protocol Analysis: Verbal Reports as Data*. MIT Press, Cambridge, Massachusetts, USA.

Erlandsson, M., & Jansson, A. (2007) Collegial verbalisation - a case study on a new method on information acquisition. *Behaviour & Information Technology*, 26(6), 535-543.

Erlandsson, M., & Jansson, A. (2013) Verbal reports and domain-specific knowledge: A comparison between collegial and retrospective verbalisation. *Cognition, Technology & Work*, 15(3), 239-254.

Eskridge, T. C., & Hoffman, R. (2012) Ontology Creation as a Sensemaking Activity. *Intelligent Systems*, 27(5), 58-65.

Farrington-Darby, T., & Wilson, J. R. (2006) The nature of expertise: A review. *Applied Ergonomics*, 37(1), 17-32.

Feldman, R. (2013) Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4), 82-89.

Feltovich, P. J., Ford, K. M., & Hoffman, R. R. (Eds.) (1997) *Expertise in Context: Human and Machine*. MIT Press, Cambridge, Massachusetts, USA.

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., Lally, A., Murdock, J. W., Nyberg, E., & Prager, J. (2010) Building Watson: An overview of the DeepQA project. *AI Magazine*, 31(3), 59-79.

Fransella, F., & Bannister, D. (1977) *A Manual for Repertory Grid Technique*. Academic Press, London, England, UK.

Fransella, F., Bell, R., & Bannister, D. (2003) *A Manual for Repertory Grid Technique* (2nd ed.). John Wiley & Sons, Inc, Chichester, England, UK.

Gaines, B. R. (2013) Knowledge acquisition: Past, present and future. *International Journal of Human-Computer Studies*, 71(2), 135-156.

Gaines, B. R., & Shaw, M. L. G. (1997) Knowledge acquisition, modelling and inference through the World Wide Web. *International Journal of Human-Computer Studies*, 46(6), 729-759.

Gammack, J. G. (1987) Different Techniques and Different Aspects of Declarative Knowledge. In A. L. Kidd (Ed.), *Knowledge Acquisition for Expert Systems: A Practical Handbook*. Plenum Press, New York, New York, USA.

Gammack, J. G., & Young, R. M. (1985) Psychological Techniques for Eliciting Expert Knowledge. In M. Bramer (Ed.), *Research and Development in Expert Systems*. Cambridge University Press, Cambridge, England, UK.

Gavrilova, T., & Andreeva, T. (2012) Knowledge elicitation techniques in a knowledge management context. *Journal of Knowledge Management*, 16(4), 523-537.

Gil, Y. (2011) Interactive knowledge capture in the new millennium: How the Semantic Web changed everything. *The Knowledge Engineering Review*, 26(1), 45-51.

Gray, R. (2004) Attending to the execution of a complex sensorimotor skill: Expertise differences, choking, and slumps. *Journal of Experimental Psychology: Applied*, 10(1), 42-54.

Grover, M. D. (1983) *A pragmatic knowledge acquisition methodology*. 8th International Joint Conference on Artificial Intelligence, Karlsruhe, Germany.

Groznik, V., Guid, M., Sadikov, A., Možinaa, M., Georgiev, D., Kragelj, V., Ribaričc, S., Pirtošekb, Z., & Bratko, I. (2013) Elicitation of neurological knowledge with argument-based machine learning. *Artificial Intelligence in Medicine*, 57(2), 133-144.

Hart, A. (1986) *Knowledge Acquisition for Expert Systems*. Kogan Page, London, UK.

Hayes-Roth, F., Waterman, D. A., & Lenat, D. B. (1983) *Building Expert Systems*. Addison-Wesley, Reading, Massachusetts, USA.

Hayes, P., Eskridge, T. C., Saavedra, R., Reichherzer, T., Mehrotra, M., & Bobrovnikoff, D. (2005) *Collaborative knowledge capture in ontologies*. 3rd International Conference on Knowledge Capture, Banff, Alberta, Canada.

Heath, T., & Bizer, C. (2011) Linked Data: Evolving the Web into a Global Data Space. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 1(1), 1-136.

Hoffman, R. R. (1987) The problem of extracting the knowledge of experts from the perspective of experimental psychology. *AI Magazine*, 8(2), 53-66.

Hoffman, R. R. (1989) A survey of methods for eliciting the knowledge of experts. *ACM SIGART Bulletin*, 108, 19-27.

Hoffman, R. R. (1998) How can expertise be defined? Implications of research from cognitive psychology. In R. Williams, W. Faulkner & J. Fleck (Eds.), *Exploring Expertise*. Macmillan, New York, New York, USA.

Hoffman, R. R., Coffey, J. W., Ford, K. M., & Carnot, M. J. (2001) *Storm-LK: A human-centered knowledge model for weather forecasting*. Human Factors and Ergonomics Society Annual Meeting, Minneapolis, Minnesota, USA.

Hoffman, R. R., Crandall, B., & Shadbolt, N. R. (1998) Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis. *Human Factors*, 40(2), 254-276.

Hoffman, R. R., & Lintern, G. (2006) Eliciting and representing the knowledge of experts. In K. A. Ericsson, N. Charness, P. Feltovich & R. R. Hoffman (Eds.), *Cambridge Handbook of Expertise and Expert Performance*. Cambridge University Press, New York, New York, USA.

Hoffman, R. R., Shadbolt, N. R., Burton, A. M., & Klein, G. (1995) Eliciting knowledge from experts: A methodological analysis. *Organizational Behavior and Human Decision Processes*, 62(2), 129-158.

Hutchins, E. (1995) *Cognition in the Wild*. MIT Press, Cambridge, Massachusetts, USA.

Jankowicz, D. (2003) *The Easy Guide to Repertory Grids*. John Wiley & Sons Ltd, Chichester, England, UK.

Jankowicz, D., & Thomas, L. (1982) An algorithm for the cluster analysis of repertory grids in human resource development. *Personnel Review*, 11(4), 15-22.

Johnson, L., & Johnson, N. (1987) Knowledge Elicitation Involving Teachback Interviewing. In A. Kidd (Ed.), *Knowledge Elicitation for Expert Systems: A Practical Handbook*. Plenum Press, New York, New York, USA.

Johnson, P. E., Zualkernan, I., & Garber, S. (1987) Specification of expertise. *International Journal of Man-Machine Studies*, 26(2), 161-181.

Kelly, G. A. (1955) *The Psychology of Personal Constructs*. W. W. Norton and Company, New York, New York, USA.

Klein, G. A., Calderwood, R., & Macgregor, D. (1989) Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man and Cybernetics*, 19(3), 462-472.

Knublauch, H., Ferguson, R. W., Noy, N. F., & Musen, M. A. (2004) *The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications*. 3rd International Semantic Web Conference (ISWC'04), Hiroshima, Japan.

Knublauch, H., Horridge, M., Musen, M., Rector, A., Stevens, R., Drummond, N., Lord, P., Noy, N. F., Seidenberg, J., & Wang, H. (2005) *The Protégé OWL Experience*. OWL Experiences and Directions Workshop at ISWC2005, Galway, Ireland.

LaFrance, M. (1987) The knowledge acquisition grid: A method for training knowledge engineers. *International Journal of Man-Machine Studies*, 26(2), 245-255.

Maedche, A., & Staab, S. (2003) Ontology Learning. In S. Staab & R. Studer (Eds.), *Handbook on Ontologies*. Springer-Verlag, Berlin, Germany.

Mendonça, F. M., Coelho, K. C., de Andrade, A. Q., & Almeida, M. B. (2012) *Knowledge acquisition in the construction of ontologies: A case study in the domain of hematology*. 3rd International Conference on Biomedical Ontology, Graz, Austria.

Milton, N. (2003) *Personal Knowledge Techniques.* University of Nottingham, Nottingham, England, UK.

Milton, N. (2012) Acquiring Knowledge from Subject Matter Experts. In J. Kantola & W. Karwowski (Eds.), *Knowledge Service Engineering Handbook*. CRC Press, Boca Raton, Florida, USA.

Milton, N., Shadbolt, N., Cottam, H., & Hammersley, M. (1999) Towards a knowledge technology for knowledge management. *International Journal of Human-Computer Studies*, 51(3), 615-641.

Milton, N. R. (2007) *Knowledge Acquisition in Practice: A Step-By-Step Guide*. Springer, London, England, UK.

Mitchell, T. M. (1997) *Machine Learning*. McGraw-Hill, New York, New York, USA.

Mozina, M., Guid, M., Krivec, J., Sadikov, A., & Bratko, I. (2008) *Fighting Knowledge Acquisition Bottleneck with Argument Based Machine Learning*. 18th European Conference on Artificial Intelligence, Patras, Greece.

Newell, A., & Simon, H. A. (1963) GPS, A Program that Simulates Human thought. In E. Feigenbaum & J. Feldman (Eds.), *Computers and Thought*. McGraw-Hill, New York, New York, USA.

Nonaka, I., & Takeuchi, H. (1995) *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, New York, New York, USA.

Novak, J., & Cañas, A. J. (2006) *The Theory Underlying Concept Maps and How to Construct Them.* Florida Institute for Human and Machine Cognition, Pensacola, Florida, USA.

O'Hara, K., Shadbolt, N. R., & Van Heijst, G. (1998) Generalized directive models: Integrating model development and knowledge acquisition. *International Journal of Human-Computer Studies*, 49(4), 497-522.

O'Hare, D., Wiggins, M., Williams, A., & Wong, W. (1998) Cognitive task analyses for decision centred design and training. *Ergonomics*, 41(11), 1698-1718.

Pang, B., & Lee, L. (2008) Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.

Sarawagi, S. (2008) Information extraction. *Foundations and Trends in Databases*, 1(3), 261-377.

Schreiber, G. (2013) Knowledge acquisition and the Web. *International Journal of Human-Computer Studies*, 71(2), 206-210.

Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N. R., Van de Velde, W., & Weilinga, B. (2000) *Knowledge Engineering and Management: The CommonKADS Methodology*. MIT Press, Cambridge, Massachusetts, USA.

Schvaneveldt, R., Durso, F., Goldsmith, T., Breen, T., Cooke, N., Tucker, R., & Maio, J. (1985) Measuring the Structure of Expertise. *International Journal of Man-Machine Studies*, 23(6), 699-728.

Schweikert, R., Burton, A. M., Taylor, N. K., Corlett, E. N., Shadbolt, N. R., & Hedgecock, A. P. (1987) Comparing knowledge elicitation techniques: A case study. *Artificial Intelligence Review*, 1, 245-253.

Scott, C. A., Clayton, J. E., & Gibson, E. L. (1991) *A Practical Guide to Knowledge Acquisition*. Addison Wesley, Boston, Massachusetts, USA.

Shadbolt, N., & Burton, A. M. (1989) The empirical study of knowledge elicitation techniques. *ACM SIGART Bulletin - Special Issue on Knowledge Acquisition*, 108, 15-18.

Shadbolt, N., Hall, W., & Berners-Lee, T. (2006) The semantic web revisited. *IEEE Intelligent Systems*, 21(3), 96-101.

Shadbolt, N., & Milton, N. (1999) From knowledge engineering to knowledge management. *British Journal of Management*, 10(4), 309-322.

Shadbolt, N. R. (2005) Eliciting Expertise. In J. R. Wilson & N. Corlett (Eds.), *Evaluation of Human Work* (3rd ed.). CRC Press, Boca Raton, Florida, USA.

Shadbolt, N. R. (2013) Knowledge acquisition and the rise of social machines. *International Journal of Human-Computer Studies*, 71(2), 200-205.

Shadbolt, N. R., & Burton, A. M. (1990a) Knowledge Elicitation. In J. R. Wilson & E. N. Corlett (Eds.), *Evaluation of Human Work: A Practical Ergonomics Methodology* (2nd ed., pp. 406-440). Taylor and Francis, London, England.

Shadbolt, N. R., & Burton, A. M. (1990b) Knowledge elicitation techniques: Some experimental results. In K. L. McGraw & C. R. Westphal (Eds.), *Readings in Knowledge Acquisition*. Ellis Horwood, New York, New York, USA.

Shadbolt, N. R., & Burton, M. (1995) Knowledge Elicitation. In J. R. Wilson & N. Corlett (Eds.), *Evaluation of Human Work: A Practical Ergonomics Methodology*. Taylor & Francis, London, England.

Shadbolt, N. R., O'Hara, K., & Crow, L. (1999) The experimental evaluation of knowledge acquisition techniques and methods: History, problems and new directions. *International Journal of Human-Computer Studies*, 51(4), 729-755.

Shaw, M. L. G. (1980) *On Becoming A Personal Scientist: Interactive Computer Elicitation Of Personal Models Of The World*. Academic Press, London, UK.

Shaw, M. L. G., & Gaines, B. R. (1987) An Interactive Knowledge Elicitation Technique using Personal Construct Technology. In A. L. Kidd (Ed.), *Knowledge Acquisition for Expert Systems: A Practical Handbook*. Plenum Press, New York, New York, USA.

Shaw, M. L. G., & Gaines, B. R. (2001) WebGrid: Knowledge Elicitation and Modeling on the World Wide Web. In R. Rajkumar (Ed.), *Industrial Knowledge Management: A Micro-Level Approach* (pp. 335-348). Springer-Verlag, London, UK.

Shortliffe, E. H. (1976) *Computer-Based Medical Consultations: MYCIN*. Elsevier, New York, New York, USA.

Stewart, T. A. (1997) *Intellectual Capital: The New Wealth of Organizations*. Nicholas Brealey, London, UK.

Sure, Y., Staab, S., & Studer, R. (2003) On-To-Knowledge Methodology (OTKM). In S. Staab & R. Studer (Eds.), *Handbook on Ontologies*. Springer-Verlag, Berlin, Germany.

Tudorache, T., Nyulas, C., Noy, N. F., & Musen, M. A. (2013) WebProtégé: A collaborative ontology editor and knowledge acquisition tool for the Web. *Semantic Web*, 4(1), 89-99.

Van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2005) Uncovering expertise-related differences in troubleshooting performance: Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, 19(2), 205-221.

Walker, B. M., & Crittenden, N. (2012) The Use of Laddering: Techniques, Applications and Problems. In P. Caputi, L. L. Viney, B. M. Walker & N. Crittenden (Eds.), *Personal Construct Methodology*. Wiley-Blackwell, Chichester, England, UK.

Wang, Y., Sure, Y., Stevens, R., & Rector, A. (2006) *Knowledge elicitation plug-in for Protégé: Card sorting and laddering*. 1st Asian Semantic Web Conference, Beijing, China.

Weiss, S., & Kulikowski, C. (1984) *A Practical Guide to Designing Expert Systems*. Rowman and Allanheld, Totowa, New Jersey, USA.

Witten, I. H., & Eibe, F. (2005) *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier, San Francisco, California, USA.

Wood, L. E., & Ford, J. M. (1993) Structuring Interviews with Experts During Knowledge Elicitation. In K. M. Ford & J. M. Bradshaw (Eds.), *Knowledge Acquisition as Modeling*. Wiley, New York, New York, USA.

Wu, H., Zubair, M., & Maly, K. (2006) *Harvesting social knowledge from folksonomies*. 17th Conference on Hypertext and Hypermedia, Odense, Denmark.

Zsambok, C. E., & Klein, G. A. (Eds.) (1997) *Naturalistic Decision Making*. Lawrence Erlbaum Associates, Mahwah, New Jersey, USA.