"Question answering" is an ill-defined notion and can practically encompass any system. For example, the call for papers for the symposium draws an analogy between question answering and the Turing test on one extreme, and finding answers on the internet on the other. As organizers of the symposium, our interest in question answering was as a technique to measure the competence of a knowledge base. While the symposium program evolves towards the goal of measuring the competence of a knowledge base, the purpose of this panel is to set the stage - to define the problem, assess the mood of the industry, and speculate about what we could expect knowledge bases to do. The participants in this panel are a good mix of down-to-earth industrialists to cutting edge academicians, who will answer questions such as:

What does a knowledge base mean to them? What are the two major accomplishments of knowledge bases in the last 10 years?

Have knowledge bases found any serious uses in industry? Does industry view knowledge bases as a critical part of their solution? What are the sources of skepticism?

Are ontologies the solution? For what? Do we need to do more? Less?

Have knowledge bases affected the state-of-the-art of database industry? Or the state-of-the-art of natural language processing?

What could knowledge bases do in 3, 5 and 10 years that would be impressive and realistic?

What are the critical factors for building a successful knowledge base? How do we test the flexibility of a knowledge base?
The Question Answering framework has been used in several recent government-sponsored evaluations. Such evaluations have helped the sponsors of research in relating problems that are meaningful to them to what can be accomplished using the technology, and to get a sense of differences between alternative approaches to solve the same problem. The evaluations have also helped the research community in focusing on a common set of "real" problems on which to try out new techniques and compare the results. However, a cynical view on such evaluations is that they have forced the participants to take a fairly "short sighted" approach to conduct research and urged them to simply try to beat a metric. They have also forced the researchers into focusing on low hanging fruit and converge on similar approaches giving "local maxima", and prevented solutions to real bottlenecks that can be addressed only by a much longer term view of research than is permitted by evaluations. The purpose of this panel is to take an introspective look at some of the recent evaluations, reconcile the two conflicting views suggested here, and perhaps, come up with methodological recommendations for evaluating question answering systems.

Each participant of this panel has been involved in one or more of government-sponsored evaluations. They will talk about their experience by addressing several of the following questions.

What was measured during the evaluation?

Was the problem domain sensitive or domain independent?

Were the measurements objective or subjective?

Was the evaluation competitive?

Were the experiments run by researchers or by a third party?

Where the test problems known before hand?

How much training time and data was provided?

What were the socio-dynamic responses to the evaluation?

What was the aftermath? Did evaluations improve the technology?
There's a proliferation of QA systems and NL chat systems on the Web and in Intranets (e.g. www.neuromedia.com or www.ask.com). The question arises: how far are these systems from "real" NLU? After all, they seem to follow Eliza. Do they? How do we settle such questions?

QA systems typically comprise of an NL understanding systems and a knowledge base/database. Given a user's query, what is the complexity of retrieving the information from the database? How do we approach this problem? What does it mean that one query is more complex than the other?

NLP is in demand. For this demand to be sustained we need a "Moore's Law for natural language engineering" so that the business community could understand what is possible when and at what price.

This panel will address the above three issues from at least the following three perspectives:

- lexicon, in particular complexity of lexical semantics.
- text and dialog, and their semantics complexity.
- abstract model, which supports the analyses presented.