# SUMMARY

The exponential increase in new knowledge that characterizes our modern information age precludes depending solely on individual effort to keep up with new information. We must therefore develop new ways of "keeping up," and we must develop them quickly. The semantic web offers a promise that we can "keep up" by allowing software agents to roam in cyberspace on our behalf, where they can gather information of interest and synergistically assist us in decision making and in negotiating for our wants and desires. This dream, however, relies on agents being able to find and manipulate useful information, which, in turn, relies on having an abundance of ontologically described repositories. Hence, the fundamental enabling component for the semantic web is an ontological description of information, which provides for a shared understanding of a repository of information.

The semantic web represents a fundamental advance in web utility, but it is difficult to create semantic-web content because pages must be semantically annotated through processes that are mostly manual and require a high degree of engineering skill. Building a functional and effective semantic web truly is a grand challenge. Furthermore, users need an effective way to query the semantic web, but any burden we place on users to learn a query language is unlikely to garner sufficient user support and interest. If we want users to take advantage of the semantic web, we must devise a means for transforming existing, traditional web pages into semantic web pages, and we must provide a simple and unrestricted interface for processing user queries and helping users invoke service requests.

We propose using information extraction ontologies to handle these challenges. We show how a successful ontology-based data-extraction technique can (1) automatically generate semantic annotations for ordinary web pages, (2) support free-form, textual queries of semantically annotated sources, and (3) support free-form service requests. Our approach demonstrates how the semantic web can be created for and used by ordinary people. We intend to build a full-fledged, working prototype to demonstrate that our proposal works.

*Intellectual Merit.* Building the semantic web is a grand challenge, and many researchers are contributing to its development. In contrast to existing approaches, we propose to advance the state of the art by developing technology to view the existing web through a semantic lens. Our proposal is based on proven conceptual modeling and information extraction techniques and builds on our earlier NSF-funded work. Our contributions can help pave the way for transforming portions of the existing, traditional web into the semantic web. Specifically, we will show how to automatically annotate existing, data-rich web pages for the semantic web, query them using free-form questions for specific answers, and invoke services based on free-form user requests.

*Broader Impact.* This project has the potential to make a real impact on society. The problem is significant, and a solid resolution will be a major advance in web technology, enabling ordinary users to reap the benefits of the semantic web. We are also poised to make an impact on a number of students. Our research team is multi-disciplinary, coming from Computer Science, eBusiness, and Computational Linguistics. Thus, this project will directly impact students, faculty, and courses in three different colleges. Our student research team is diverse, including 4 PhD students (2 female), 4 masters students (1 female), and 1 undergraduate student, collectively from the US (4), China (3), and Syria (2). We intend to continue serving diverse and underrepresented segments of society. Our faculty and PhD students are actively involved with teaching at the graduate and undergraduate level. As we have done in the past, we will continue to publish our results and research artifacts on our web site and in peer-reviewed journals and conference proceedings.