DOI:10.1145/2133806.2133829

Technical Perspective
Who Knows? Searching for Expertise on the Social Web

By Ed H. Chi

IT IS DIFFICULT to remember what people had to do to find the answer to a question before the Web. Imagine it is 1990, before the age of search engines, and of course, Wikipedia. You have no access to a library, and available reference books are not enough. The only option you might have is to call a friend who might know the answer. In fact, this option is so important, it is baked into the game "Who Wants to be a Millionaire?" as one of the three lifeline options to take when you are stumped for an answer. This natural instinct to call someone is also baked into the DNA of Aardvark, the social question and answering (QA) engine described in the following paper by D. Horowitz and S. Kamvar.

When you turn to the phone, one of the first steps you have to figure out is who to call. This is the expertise location or question-routing problem in social QA research. At a high level, this seems like a great computer science problem. You have people as nodes, and their relationships and interactions as edges, and you want to model people's interests and expertise, as well as the frequency and recency of their interactions with each other. You will use these models to route questions. Your mind races with possible algorithms and user modeling approaches to apply. Yes, conceptually you would probably be correct with many of these ideas, but in practice, building Aardvark is much more complex and difficult.

First, you must figure out how to build accurate user models for each user on the whole Web, even with sparse data for many users, including brand new users of your system.

Second, you have to scale this system to millions of users, and be able to do the question routing in milliseconds. You also want to try and get answers that are good and return them as quickly as possible.

Third, potential answerers (read: humans) are finicky: they do not want you to spam them; they do not like being interrupted if they are in the middle of another conversation; and they don't like it when you call them up at weird hours of the day (no matter if you know their time zone or not). In order words, you must deal with the real human context and its associated social interaction.

Finally, you need to socially engineer the growth of this system, so that early users get experiences good enough to rave about your service and recommend it to other users. You want to build trust, and you want a network effect, such that, as each user joins the system, the whole system becomes even more useful to those already there.

Before Aardvark, social QA systems used a wide variety of techniques to route questions, most often using experience/reputation points or monetary rewards as incentives. Many services, such as Yahoo! Answers focused on building communities, and turning the act of answering into a point-based game. Instead, Horowitz, Kamvar and their team at Aardvark pushed the envelope and built a different service that focused on getting answers as quickly as possible; from someone you are socially connected to; and who is likely to be an expert on the topic. How did they do it?

On speed: By connecting Aardvark to a chat service (Google Chat), it exploited and delivered on that expectation of immediacy. In our age of instant knowledge via search engines and Wikipedia, this impedance match is a particularly nice touch that, perhaps ironically, humanizes the experience so it feels like a phone call. Interestingly enough, in a pioneering QA system called AnswerGarden, Mark Ackerman observed that users were often more satisfied when an answer came back quickly, even if the answer was somewhat less than perfect.[1]

On social interactions: Aardvark places emphasis on the social interaction just as much as getting the information—just as if there was a smart assistant who knew your Rolodex and made that phone call for you. Indeed, in Evans and Chi,[2] we showed how social interactions were present and pervasive throughout the information seeking episode—before, during, and after the core search task.

On expertise: Aardvark skillfully exploited the design knowledge gained from years of search engine research to scale the algorithms so that it can route those questions to others who are most likely able to answer it at that very moment. In fact, they followed the original meme of describing the "anatomy" of a search engine in describing their own system—a meme worth repeating in all areas of computer science involved in engineering real-world systems.

Users want one thing—getting their questions answered immediately. Search engines have played that role for many years now. It can be argued that the greatest impact computers have had on the human endeavor is the Web search engine, whose development and refinement seems to be the epitome of computer science. That was before the Web truly became social. In the brave new social Web, search will be different, and reading this paper will give you a sense of the direction social search engines are headed. C

References

1. Ackerman, M.S. Augmenting organizational memory: A field study of answer garden. *ACM Trans. Inf. Syst. 16*, 3 (July 1998), 203-224; http://doi.acm.org/10.1145/290159.290160
2. Evans, B.M. and Chi, E.H. An elaborated model of social search. *Information Processing & Management*; http://dx.doi.org/10.1016/j.ipm.2009.10.012

Ed H. Chi (chi@acm.org) is staff research scientist at Google Research, Mountain View, CA.