

The Madsci Network: Direct Communication of Science from Scientist to Layperson

Ricky SETHI^{a,c,*} & Lynn BRY^{a,c}

^a*USC Information Sciences Institute, USA*

^b*Harvard University, USA*

^c*The Madsci Network, USA*

*rickys@sethi.org

Abstract: Internet-based volunteer communities collaboratively contribute to the expansion of human knowledge and cognition. Their popularity is evidenced not only by reference sites like Wikipedia but also by niche communities designed to help people answer complex queries, especially in relation to highly technical or scientific subjects that are beyond the ability of their peers to answer. Ask-A-Scientist sites like The Madsci Network are a subset of Ask-An-Expert sites that rely upon expert volunteers to disseminate information directly from the expert or scientist to the layperson.

This transmission of knowledge, directly from scientist to layperson, promises to transform how societies interpret scientific knowledge and how that knowledge is used in a social, cultural, and political sense. This is especially significant for children, who are fast becoming online consumers of science and are faced with the increasingly difficult problem of discerning legitimate science from pseudo-science. In this paper, we present details of the organization, implementation, and research efforts of The Madsci Network virtual community.

Keywords: Ask-A-Scientist, Virtual Community, Ask-An-Expert, Q&A, Question & Answer Site

1. Introduction

It is often very difficult to find answers to specific, specialized queries on the web, especially when they are scientific in nature. Virtual organizations like The Madsci Network¹, which can be defined as Internet-based volunteer communities, are indispensable in such cases since finding reliable scientific information is not easy on the web; harder still is to discern legitimate science from pseudo-science. This is especially significant for children, who are fast becoming online consumers of science (Dawson, 2000; Forte & Bruckman, 2006; Osborne & Collins, 2001; Sethi, 2005).

As the Internet evolves, it will transform human society in a multitude of ways. In fact, the possible effects of the Internet on human cognition and science education become apparent when examining how students use the Internet to learn about science. Our goal is to have a positive social impact on science education and add societal value to how students use the Internet to learn about science.

Given the prevalence of pseudo-science, filtered scientific reporting, and conflicting ideas, an Ask-A-Scientist virtual community such as The Madsci Network can give both students and the general public direct access to legitimate scientific knowledge in an easily comprehensible manner. This dissemination of knowledge directly from scientist to the layperson promises to transform how societies interpret scientific knowledge and how that knowledge is used in a social, cultural, and political sense.

1.1 Background on The Madsci Network

¹ <http://www.madsci.org>

The Madsoci Network (Baram-Tsabari, Sethi, Bry, & Yarden, 2006, 2008) is a virtual organization, an Internet-based volunteer community, that functions primarily as a human-mediated Question & Answer (Q&A) site. It is an online, non-profit, Ask-A-Scientist website, a subset of Q&A websites, with over 700 volunteer scientists and 28 volunteer moderators distributed globally fielding 90-150 questions a day from visitors from all over the world. It handles questions in 26 different scientific disciplines and consistently gets over 650,000 unique visitors and 4,000,000 page views per year.

Established in 1996, The Madsoci Network facilitates the direct flow of legitimate scientific information from scientists to laypeople, particularly to students (Baram-Tsabari et al., 2006), by processing questions from students, laymen, and scientists covering topics such as Physics, Astronomy, Biology, Computer Science, Earth Sciences, etc. It maintains a searchable archive of over 40,000 answered questions, experiments, and other areas of interest to learning science. It has received recognition from many organizations (including the U.S. Department of Education, Science magazine, New Scientist, and the BBC) and has received numerous awards, as well.

2. Social Question & Answer (Q&A) Communities

There is a fast-growing interest in studying Question & Answer (Q&A) communities, as evidenced by the public fascination with question-answering systems like IBM's Watson (Ferrucci et al., 2010; Kalyanpur, Murdock, & Fan, 2011) and the increasing business demand for Question & Answering social, scientific, and medical applications, such as on Yahoo! Answers, Naver, WebMD, Quora, WikiHow, Scientific American's Ask-A-Scientist, Baidu Knows, etc. (Etzioni, 2011; F. Maxwell Harper, Moy, & Konstan, 2009; F. Maxwell Harper, Raban, Rafaeli, & Konstan, 2008; Nam, Ackerman, & Adamic, 2009; Wren, 2011). These Q&A sites are quickly becoming repositories of collective knowledge; Yahoo! Answers alone has over 120 million users and 400 million answers while Naver has accumulated over 70 million questions and answers with an average of 110,000 answers to 44,000 questions asked every day from 4.5 million daily visitors (F. Maxwell Harper et al., 2009, 2008; F.M. Harper, Weinberg, Logie, & Konstan, 2010; Nam et al., 2009). In fact, just the Yahoo! Answers Q&A community generates as many new Web pages each month as are contained in the English language Wikipedia (F. Maxwell Harper et al., 2009). In general, Q&A websites fall into three categories: Digital Reference Services (e.g., NY Public Library's "Ask Librarians Online"), Ask-An-Expert Services (e.g., The Madsoci Network or Scientific American's Ask-The-Brains), and Community Q&A Sites (e.g., Yahoo! Answers) (F. Maxwell Harper et al., 2008).

An Ask-A-Scientist social Q&A community like The Madsoci Network is not just another site to give yet another answer, but helps provide a way to *think critically* about scientific facts; this approach not only helps provide insight to students but builds techniques to engender lifelong learning. As a consequence of their design, learning (increased knowledge plus the ability to think critically about information) becomes an emergent property of these systems and can inform societal structures and communal knowledge.

In fact, social Q&A sites leverage the wisdom of crowds to go well beyond the limitations of search engines (F. Maxwell Harper et al., 2008). Social Q&A sites are thus also differentiated from question-answering systems which attempt to deconstruct a natural language question into information retrieval and analysis tasks automatically (Wren, 2011). The *raison d'être* for social Q&A sites is to allow users to approach questions that have many different perspectives or explore questions about which they want to think deeply and critically.

The general framework of such Q&A communities is usually composed of:

- Users: the information seekers who submit the queries
- Experts: the scientists or other domain experts who answer the queries
- Moderators: contributors who guide the question and answer flow, including triaging incoming questions, matching experts to new questions, evaluating answers for quality assurance, etc.

Q&A sites use many variations of this basic framework. For instance, different sites certify experts in different ways: The Madsci Network requires academic qualifications whereas Yahoo! Answers allows the users to also be the experts. In fact, different Q&A communities have different organizations, as well, with some sites not having any moderators or using the users to also help moderate the site. A typical dataflow for incoming questions to The Madsci Network is shown in Figure 1.

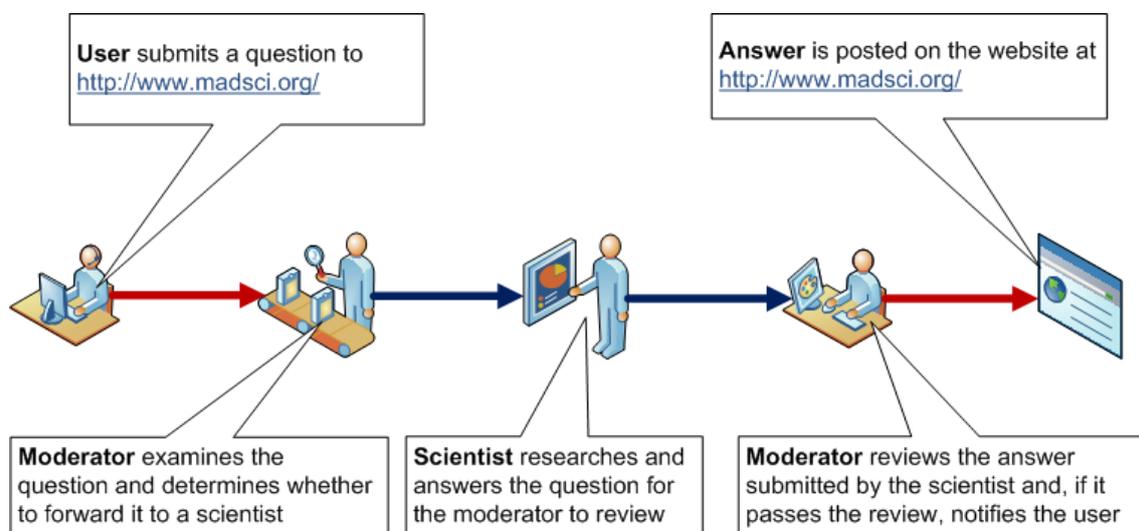


Figure 1: Overview of The Madsci Network Dataflow

3. Dataset

The Madsci Network makes its data available at <http://research.madsci.org/dataset/>. The Madsci Network maintains an online archive of more than 40,000 answered questions. These question/answer pairs are available from 1996-2011. An additional 110,000+ submitted questions from 1995 - present are maintained offline with associated meta-data. Full names and email addresses are not included. Access to partial or full components of the 150,000+ question dataset may be provided per request for research purposes only.

4. The Madsci Network Research Goals

The Madsci Network's ongoing research addresses three broad goals: to improve the efficiency and quality of Q&A sites, to explore the archives of these sites to identify social patterns and trends, and to create novel structures to support emergent eLearning in many different contexts. The direct exchange between scientists and students on The Madsci Network has already shown great promise for transforming how students learn and how societies cope with scientific and technological change. The inter-disciplinary nature of this enterprise has led to collaborations with experts from fields as diverse as Computer Science to Psychology to Science Education.

These collaborations in the analysis of over 10 years worth of data on The Madsci Network had surprising revelations, including a startling dominance of female contributions among K-12 students (contrary to offline situations). This female enthusiasm was observed in different countries, and had no correlation to the level of equality in those countries: Iranian girls, for example, turned out to be greater users of madsci.org than Iranian boys, although their direct environment does not necessarily promote such interest in science. This may indicate that the Internet and sites like The Madsci Network play a potentially empowering and democratic role as a free-choice science-learning environment, which is especially relevant to populations which are deprived of equal opportunities in learning formal science.

5. Acknowledgements

We would like to gratefully acknowledge support from NSF Award 1019343/Sub Award CIF-B-17.

References

- Baram-Tsabari, A., Sethi, R. J., Bry, L., & Yarden, A. (2006). Using questions sent to an Ask-A-Scientist site to identify children's interests in science. *Science Education*, 1050 – 1072. doi:10.1002/sce
- Baram-Tsabari, A., Sethi, R. J., Bry, L., & Yarden, A. (2008). Asking scientists: A decade of questions analyzed by age, gender, and country. *Science Education*, 93(1), 131–160. doi:10.1002/sce.20284
- Dawson, C. (2000). Upper primary boys' and girls' interests in science: have they changed since 1980? *International Journal of Science Education*, 22, 557.
- Etzioni, O. (2011). Search needs a shake-up. *Nature*, 476(7358), 25–6. doi:10.1038/476025a
- Ferrucci, D., Brown, E., Chu-carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., ... Prager, J. (2010). Building Watson: An Overview of the DeepQA Project. *AI Magazine*, 59–79.
- Forte, A., & Bruckman, A. (2006). From Wikipedia to the classroom: exploring online publication and learning. In *Learning sciences* (p. 188). International Society of the Learning Sciences.
- Harper, F. Maxwell, Moy, D., & Konstan, J. A. (2009). Facts or friends?: distinguishing informational and conversational questions in social Q&A sites. In *CHI* (pp. 759–768).
- Harper, F. Maxwell, Raban, D., Rafaei, S., & Konstan, J. A. (2008). Predictors of answer quality in online Q&A sites. In *SIG CHI* (p. 865). New York, New York, USA: ACM Press. doi:10.1145/1357054.1357191
- Harper, F.M., Weinberg, J., Logie, J., & Konstan, J. A. (2010). Question types in social Q&A sites. *First Monday*, 15(7-5).
- Kalyanpur, A., Murdock, J., & Fan, J. (2011). Leveraging Community-Built Knowledge for Type Coercion in Question Answering. *Knowledge Creation Diffusion Utilization*, 144–156.
- Nam, K. K., Ackerman, M. S., & Adamic, L. A. (2009). Questions in, Knowledge in? A Study of Naver's Question Answering Community. In *CHI* (pp. 779–788).
- Osborne, J., & Collins, S. (2001). Pupils' views of the role and value of the science curriculum: a focus-group study. *International Journal of Science*.
- Sethi, R. J. (2005). Using virtual laboratories and online instruction to enhance physics education. *JPTEO*.
- Wren, J. D. (2011). Question answering systems in biology and medicine--the time is now. *Bioinformatics (Oxford, England)*, 27(14), 2025–2026. doi:10.1093/bioinformatics/btr327