

Semester: Summer 2010
MS Student: Sourabh R. Joshi
Chair: Dr. Sridhar Radhakrishnan
Thesis Title: INFORMATION SHARING ON THE WEB: ARCHITECTURE,
POLICIES AND COMPLAINTS

ABSTRACT:

In the past decade, Internet has reformed the way we look at information sharing. This transformation is brought about by the technological developments as well as active participation of the web users. The Internet paradigm of information sharing is still a developing concept. In theory, Information should be treated as intellectual property of the producer of information. But this concept is not very well recognized by the Internet community due to the open source nature of the web. This is a cause of many problems surrounding information sharing on the web. In the recent years, active members of the web community have been trying to introduce mechanisms and models to implement order and control over the situation. To mention a few of them, Robot Exclusion Protocol (REP) and the Automated Content Access Protocol (ACAP) are the most popular and widely used of them. In the first phase of this study, we studied and critically analyzed the available solutions and surveyed all the issues surrounding information sharing on the web. We have developed a new protocol for information sharing on the web that takes into account new requirements and salient features of the existing protocols. This new model uses a policy based approach. The policies provide the owner with control over the dissemination and use of shared information. Policies can be thought of as rules and restrictions that dictate the will of the owner. The scope of these policies covers who gets the information, how he gets it, how can he use/share it and other such regulations. We have also extended previous works gathering meta-information about the website to enable search engines to crawl better. Previous work in this area, like Sitemap and Resource of resource (ROR) has an extensive implementation base. But these standards have not been revised and updated for some time now. The second phase of our study deals with implementing compliance in the system. Compliance is to verify that the policies prescribed are being adhered to. This new feature increases the level of trust of the participants in the model, which is necessary. In our model, we deal with the two different types of interactions demanding different compliance mechanisms solutions. The verification strategy based on log auditing is the basis of these mechanisms. The variations are only in the strategies of how the logs are collected. When trying to verify an interaction between two indirect participants it becomes important to detect a fraud and find the cheaters. Thus the log collection strategies are designed to find the cheater thus making the compliance mechanism more dependable. When different parts of the study are combined together we get a model that deals with a whole range of problems in the information sharing paradigm.