Humboldt Discoverer: A semantic P2P index for PDMS

Sven Herschel and Ralf Heese

Humboldt-Universität zu Berlin, Databases and Information Systems
Unter den Linden 6, D-10099 Berlin, Germany,
{herschel | rheese}@dbis.informatik.hu-berlin.de

Abstract. Recently, Peer Data Management Systems (PDMS) came into the focus of research as a natural extension to distributed databases in the peer-to-peer (P2P) context. A PDMS consists of a set of peers, each of which acts as an information integration component. Queries submitted to one peer are answered by local data as well as by data that is reachable along paths of mappings through the network of peers. It is necessary, however, to manually formulate these mappings between peers.

The Humboldt Discoverer overcomes this restriction by providing a semantic peer-to-peer index for the PDMS architecture. Using this index, a peer locates useful information sources that are not reachable by a conventional mapping path. The Humboldt Discoverer realizes a hybrid P2P infrastructure combining a distributed hash table approach with an unstructured P2P network, which indexes peer schemata by means of Semantic Web technologies. By indexing schema information and choosing the hybrid architecture approach we expect the index to be reasonably insensitive to both changes of the network structure as well as updates of the data instances.