



NORTHWESTERN  
UNIVERSITY

# Computer Science Lecture Series

## Toward the scalable integration of Internet information sources

Mr. Jaewoo Kang  
University of Wisconsin-Madison

12:30-1:30pm  
Wednesday, April 2, 2003  
Room 381 - Computer Science Dept.

### Abstract:

As the number of databases accessible on the Web grows, the ability to execute queries spanning multiple remote heterogeneous databases is becoming increasingly important. Two challenges in providing such a capability are (1) to discover the semantic correspondences between schema and data elements across the autonomous, heterogeneous information sources, and (2) developing query processing algorithms that work over data that arrives in a stream from a remote data source rather than data that resides on a local disk. Addressing the first problem, we introduced an automatic schema matching algorithm, "uninterpreted matching." With respect to the second problem, I will present a new optimization framework for continuous queries over unbounded streams, using a unique unit-time basis cost model.

For a complete calendar, see:  
[www.cs.northwestern.edu](http://www.cs.northwestern.edu);  
click on 'CS Seminars'

Join our e-mail list for notices of upcoming Computer Science lectures:  
send an e-mail with the word 'subscribe' in the subject line to:  
[cs\\_seminar@cs.northwestern.edu](mailto:cs_seminar@cs.northwestern.edu)