

MIT9904-08 Haystack: Per-User Information Environments

David R. Karger and Lynn Andrea Stein



Project Overview



- Electronic infrastructure to triangulate knowledge
 - an individual
 - of collegial communities
 - of the world at large
- Independent, interacting information repositories, each customized to its individual user
 - Automated data gathering
 - Active observation of user activity
 - Adaptation to individual query needs
 - Inter-haystack collaboration

MIT9904-08 Haystack: Per-User Information Environments

David R. Karger and Lynn Andrea Stein



Progress Through June 2000

- Integrated structured queries
 - Structured properties (author=) using RDBMS
 - Unstructured (free text) properties using IR system
 - Hybrid queries combine structured & unstructured aspects.
- Implement and evaluate learning for query adaptation
 - Ranking, relevance feedback, and query expansion
 - Improved query behavior
- Paper to Haystack
 - Scan/OCR kiosk
 - Seamless integration paper documents into Haystack.
- Current version archives thousands of documents.

MIT9904-08 Haystack: Per-User Information Environments

David R. Karger and Lynn Andrea Stein



Research Plan for the Next Six Months

- Back End
 - More robust transaction management
 - Enhanced fault tolerance
 - Rollback
 - Versioning of data to allow access to prior system states
 - Kernel Re-modularization
- User Interface
 - Object Oriented
 - Hints-based
 - Integrate customized display with Haystack data model