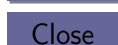
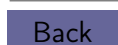


Social Information Retrieval

Sebastian Marius Kirsch
skirsch@moebius.inka.de



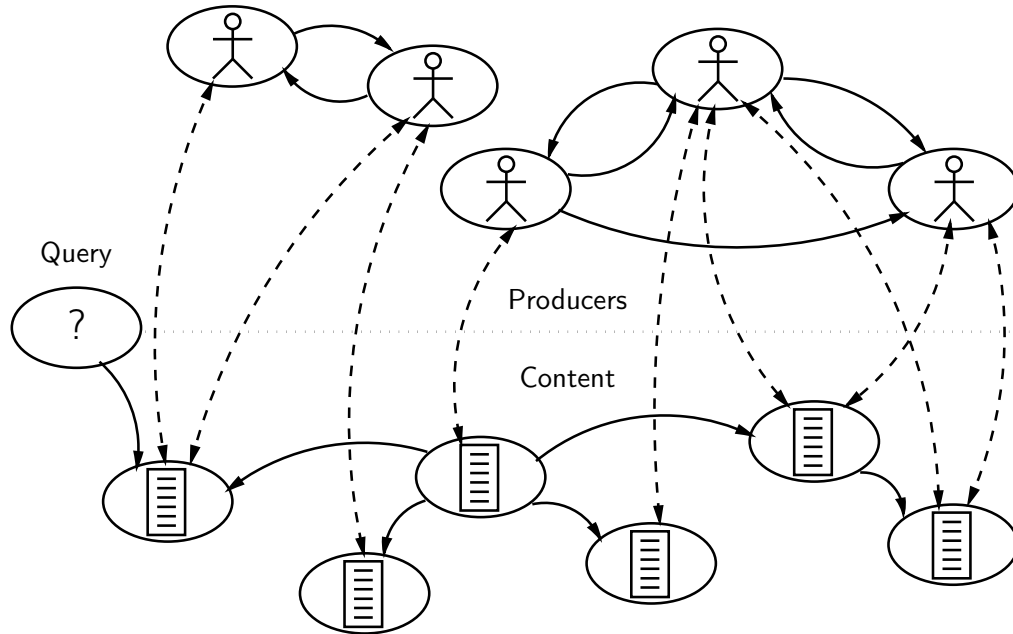


Social IR: Definition and motivation

- information production and information seeking are social activities (Wilson, 1981)
- incorporate information about authorship/social relationships
- applicable in settings where such information is available
- goal: improve search effectiveness
- social IR meets web retrieval in the context of peer-to-peer IR systems
- application: ISKODOR
- next step after personalization and collaborative IR



Model for social IR





Model and techniques

- model domain as associative network
- nodes for information producers, content, and queries
- proposed techniques:
 - spreading activation search (Anderson, 1983)
 - maximum flow communities (Flake et al., 2004)
 - baseline: vector space search with pseudo relevance feedback
- local algorithms
- scalable, implementable in a distributed system





Evaluation

- standard corpora not useful
- requirements for suitable corpora:
 1. authorship of individual documents is available
 2. social network can be inferred
 3. similarity or references between documents
- selected corpora:
 - 25 years of SIGIR proceedings
 - German Wikipedia
 - mailing list archive ('origami-l')
- evaluation setting: known-item retrieval



References

- John R. Anderson. *The Architecture of Cognition*. Cognitive Science Series. Harvard University Press, 1983. ISBN 0-674-04425-8.
- Gary William Flake, Kostas Tsioutsoulis, and Leonid Zhukov. Methods for mining web communities: Bibliometric, spectral, and flow. In Alexandra Poulou and Mark Levene, editors, *Web Dynamics*, chapter 4, pages 45–68. Springer Verlag, 2004. ISBN 3-540-40676-X. URL <http://research.yahoo.com/publications/4.pdf>.
- T. D. Wilson. On user studies and information needs. *Journal of Librarianship*, 37(1):3–15, 1981. URL <http://informationr.net/tdw/publ/papers/1981infoneeds.html>.

