databases consolidating masses of data. The substantial increase in access to information afforded by the Internet has strengthened the importance of being able to distinguish between similar documents and locate relevant documents. These issues of navigability, findability, and relevance, under the guise of information retrieval and information seeking, have been of importance to library and information science since its inception.

Classification and indexing through a hierarchical classification system or thesaurus are common methods of attempting to resolve this problem, using controlled vocabularies to rationalise natural languages, removing ambiguities, and consolidating similar items. A solidly designed classification system using terms and keywords appropriate to the context of the intended user can help to reduce the difficulty inherent in searching large document spaces for information.

While the creation of generic hierarchical classification systems or subject-specific taxonomies has a long history, the design of these classification systems has been left largely to professional intermediaries. Because of the increasing amount and specialisation of information collected and user requests for more fine-grained access, these systems can be too generic for user needs. And, while full text search can provide this fine-grained access to supplement controlled vocabularies, it tends to be at the expense of precision that results from use of differing terminology.

The rise of collaborative tagging systems suggests an alternative method for creating classification systems. In fact, such social bookmarking sites are being touted as a potential solution to the problems of scale inherent in the application of any controlled vocabulary to a large document set (Hammond et al. 2005; Mathes 2004; Morville 2005). It has also been suggested that user tags, combined with topic maps and tag clusters, have the potential to provide the benefits of a controlled vocabulary, which controls for terminological differences while still allowing the use of natural language vocabulary (Shirky 2005).

This paper reports on the results of an exploratory study of CiteULike (a social bookmarking service). It examines the relationship of collaborative tagging to classical classification and indexing by comparing the tags assigned to academic journal articles by users of the CiteULike bookmarking system to library descriptors assigned by intermediary indexers and author keywords assigned by authors to their own journal articles.