

Folksonomies: Categorization, Classification, or Disruptive User-centered Classification Technology?

Abstract

This poster captures early-stage research into the mechanics of folksonomies, alternately known as ethnoclassification, socioclassification (Star), or folk classification (Van der Wal). The chart, below, extends work by Jacob, 2004, in an attempt to situate folksonomic organization in relation to more traditional ideas of categorization and classification.

Elijah Wright
School of Library and Information Science
Indiana University, Bloomington

2007 Doctoral Research Forum
October 6, 2007

Comparison at a glance: Categorization, Classification, and Folksonomy

	Categorization	Classification	Folksonomy, Folk Taxonomy	Opportunities for future research
Process	Creative synthesis of entities based on context or perceived similarity	Systematic arrangement of entities based on analysis of necessary and sufficient characteristics	Utility-based; users assign tags to items ad-hoc. Balance between own and collaborative use.	How aware are users of social or collaborative needs when assigning tags (supposedly for their own use)?
Boundaries	Because membership in any group is non-binding, boundaries are "fuzzy"	Because classes are mutually-exclusive and non-overlapping, boundaries are "fixed"	Classes are emergent and ad-hoc - even unstable - rather than rigorously specified our bounded	Items may be multiply tagged; any pair of users may (not) tag items in the same way
Membership	Flexible: category membership is based on generalized knowledge and/or immediate context	Rigorous: an entity either is or is not a member of a particular class based on the intension of the class	Flexible: tag assignment based on the user's knowledge and perceptions of the world	Semantic / categorical equivalence - a car is clearly an automobile, but is Picasso art?
Criteria for Assignment	Criteria both context-dependent and context-independent	Criteria are predetermined guidelines or principles	Criteria are user and social context dependent	Shared contexts affected by both experience, knowledge, and cultural factors
Typicality	Individual members can be rank-ordered by typicality (graded structure)	All members are equally representative (ungraded structure)	Membership not equally representative, but rank order may differ by user	Typicality is complicated by socioclassification; ordering becomes highly variable
Structure	Clusters of entities; may form hierarchical structure	Hierarchical structure of fixed classes	Radically flat structure; non-hierarchical by definition	Flat structure does not mean that there is none; structure introduced by co-tagging of entities is less than obvious

Disruptive Classification Technologies

Folksonomy, alternately known as "folk classification", or socioclassification, or ethnoclassification, is a user-driven system of collaboratively generating categories or of sorting items into classes.

Folksonomy is also the output of such a collaborative process – the set of tags, labels, or taxa generated by collaborative or socio- or ethno- classification.

Research on folksonomies is a natural fit with other work on information organization; because folksonomy is oversimplified and breaks "the rules" that we typically load classification or categorization with, it forces us to re-examine our assumptions about the process, legitimate inputs to the process, or legitimate outcomes from a categorization or classification process.

Representations of aggregate input – driven by either unstructured or tag-based systems – are increasingly popular in Web systems. We may not be able to make use of the data provided by 300 individual users, but we may be able to do very interesting things with the fact that people prefer to tag a common resource in a similar manner.

Operationalizations:

An ontology is an implicit specification of a conceptualization (Gruber 1993) - a formalized way of reasoning about the world.

A taxonomy, by contrast, is an implementation of a classification system that organizes terms (taxa) into hierarchy that strongly imply relationships (proceeding from more general to more specific) between contained terms.

A metadata scheme, typically, is an implementation of an ontological and taxonomic scheme that allows for flexible application and description of documents or artifacts.

A folksonomy, in this context, is an applied categorical metadata scheme that allows end-users to generate free-form classifications by adding "tags" - effectively, taxa – to entities (or the representations of those entities) or social artifacts.

Points to ponder about folksonomy

- Works easily and flexibly with loose conceptual categories
- Prespecification (of classes, properties) is not necessary
- Allows precision without requiring it
- Structure of categories/classes is emergent, based on available data
- Individual categorical/conceptual bias is reduced, but collective bias is retained

Bibliography

Gruber, T. 1993.

Jacob, E. Categorization and classification: a difference that makes a difference. Library Trends, Winter 2004.