

	<p style="text-align: right;">2nd Version Date : 19/06/2007</p> <p>TAXONOMY DIRECTED FOLKSONOMIES</p> <p>Integrating user tagging and controlled vocabularies for Australian education networks</p> <p>Sarah Hayman and Nick Lothian education.au Adelaide Australia</p>
<p>Meeting:</p>	<p>157 Classification and Indexing</p>
<p>Simultaneous Interpretation:</p>	<p>No</p>
<p style="text-align: center;">WORLD LIBRARY AND INFORMATION CONGRESS: 73RD IFLA GENERAL CONFERENCE AND COUNCIL 19-23 August 2007, Durban, South Africa http://www.ifla.org/iv/ifla73/index.htm</p>	

**TAXONOMY DIRECTED
FOLKSONOMIES**

*Integrating user tagging and controlled
vocabularies for Australian education
networks*

Sarah Hayman and Nick Lothian
education.au

*World Library and Information Congress / 73rd IFLA
General Conference and Council
Durban, South Africa, August 2007*



Abstract

What is the role of controlled vocabulary in a Web 2.0 world? Can we have the best of both worlds: balancing folksonomies and controlled vocabularies to help communities of users find and share information and resources most relevant to them?

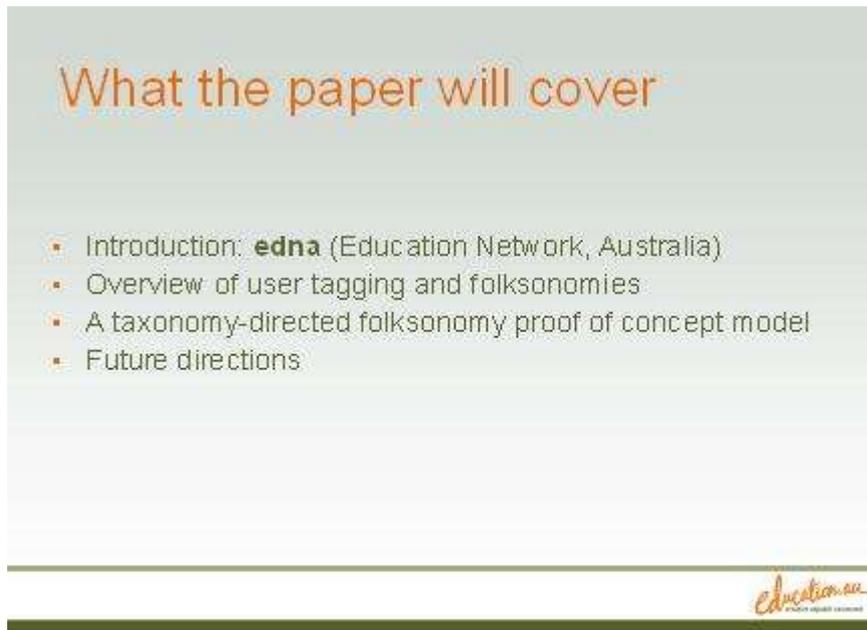
***education.au** develops and manages Australian online services for education and training. Its goal is to bring people, learning and technology together. **education.au** projects are increasingly involved in exploring the use of Web 2.0 developments building on user ideas, knowledge and experience, and how these might be integrated with existing information management systems. This paper presents work being undertaken in this area, particularly in relation to controlled vocabularies, and discusses the challenges faced.*

*Education Network Australia (**edna**) is a leading online resource collection and collaborative network for education, with an extensive repository of selected educational resources with metadata created by educators and information specialists. It uses controlled vocabularies for metadata creation and searching, where users receive suggested related terms from an education thesaurus, with their results. We recognise that no formal thesaurus can keep pace with user needs so are interested in exploiting the power of folksonomies.*

*We describe a proof of concept project to develop community contributions to managing information and resources, using **Taxonomy-Directed Folksonomy**. An established taxonomy from the Australian education sector suggests terms for tagging and users can suggest terms. Importantly, the folksonomy will feed back into the taxonomy showing gaps in coverage and helping us to monitor new terms and usage to improve and develop our formal taxonomies.*

*This model would initially sit alongside the current **edna** repositories, tools and services but will give us valuable user contributed resources as well as information about how users manage resources. Observing terms suggested, chosen and used in folksonomies is a rich source of information for developing our formal systems so that we can indeed get the best of both worlds.*

1 Introduction



The potential of the Internet to benefit teaching and learning was the catalyst for the creation of the company *education.au*, and the dynamic advances in information and communication technology (ICT) since then have driven the evolution of the company's capability and offerings. As Australia's national ICT in education agency, *education.au* develops and manages national online services for students and educators across all sectors of Australian education. The organisation explores and provides innovative technology solutions and aims to develop connections between the creative edge of new technology and proven practice. This has always been done in a context of collaboration, sharing and networks and therefore the emergence of user contributions to information organisation and management on the web, as exemplified in the development of user tagging and folksonomies, is of great interest. *education.au* is keen to explore ways in which this new technology and user practice can inform its own services. One potential model is under consideration as a proof of concept development for a project referred to as *myedna*, a personalised interface to *edna*..



edna (Education Network Australia) is Australia’s leading online resource collection and collaborative network for the education and training community. **edna** is a joint initiative of the state and territory governments and the Australian Government, through their education departments, to provide free news, resources, networks and online tools for educators. **edna** is managed by *education.au* limited and has been operating since 1996. Although **edna** is developed and funded by the Australian education and training sector it is free online for use by anyone interested, including international users. Ownership of groups may not be available to those outside Australia but groups may have international members and **edna** resources may be searched and retrieved worldwide; many of the tools on the site are also freely available.

edna identifies and links to online teaching and learning resources from Australian and international collections, extensive listings of national and international events; online groups, email discussion lists and newsletters; ICT innovations, tools and technologies, including learning objects, RSS feeds, wikis, blogs and podcasts. **edna** resources are carefully chosen and evaluated by specialised staff and are quality assured and up-to-date.

The current **edna** website provides links to all the above online resources, searchable, categorised in a browse structure and organised within Australian educational sectors. Each resource is also assigned subject index terms by expert indexers using an appropriate thesaurus. All users essentially have the same experience on the website at present and see the same interface to events, resources and news. When they log in to **edna** Groups or **edna** Lists, they see their own selection of services. The next stage of development for **edna** is a project known as *myedna* which will enable users to personalise the way content from **edna** and other providers is displayed. However a fully customisable personal learning journey *myedna* will also allow users to do considerably more than that.

We are keen to find a way to involve users in the collection, evaluation and organisation of resources, so that they can customise their usage of **edna**, share their experiences and tell us through their behaviour and decisions what it is they value and how they want to use it. This matches very closely many of the developments in the Web 2.0 world with the advent of user generated content, online social networks, user tagging and folksonomies.

2 *User tagging and folksonomies*



User tagging and folksonomies

The development of tagging

- Web 2.0 services are about social networking
- Tagging is central to Web 2.0
- Tagging is the adding of keywords to a digital object to categorise it
- Tagging can be seen as subject indexing without a controlled vocabulary
- Tagging is not new: what is new is that tagging is now being done by anyone

education

2.1 **The development of tagging**

The term Web 2.0, first used by Tim O'Reilly in 2004, describes a cluster of web-based services with a social collaboration and sharing component, where the community as a whole contributes, takes control, votes and ranks content and contributors. Web 2.0 services include social networking sites, wikis, communication tools, weblogs, social bookmarking, podcasts, RSS feeds (and other forms of many-to-many publishing), social software, and folksonomies. Central to this new Web is the idea of tagging — the adding of keywords to a digital object (e.g. a website, picture, audiofile or videoclip) to categorise it. This activity is effectively subject indexing but generally without a controlled vocabulary.

Tagging of course is not a new concept, especially to librarians, indexers and classification professionals. What is new is that the tagging is being done by everyone, no longer by only a small group of experts, and that the tags are being made public and shared. The development of the internet and the web, and of search

engines, led to users doing their own searching. In the Web 2.0 environment users are now also doing their own content creation and information management.

User tagging and folksonomies

The development of tagging, ctd.

- 28% of internet users have tagged or categorized content online (Pew internet survey, 2006)
- Tagging is used in a range of sites for many different types of resources, e.g.

- Tags:
 - Are meaningful to users
 - Can be added to already tagged resources
 - Can be public or private.
 - Can be grouped, shared, displayed, published and managed
 - Are often displayed in a "tag cloud"

education.au

The PEW internet survey of December 2006 (Rainie, 2007) found that 28% of internet users have tagged or categorised content online such as photos, news stories or blog posts. On a typical day online, 7% of internet users say they tag or categorise online content.

Tagging is used in a range of sites for many different types of resources. Tagging is done somewhat differently at different websites, but the following all use some type of user tagging:

- Blogs (Technorati: <http://technorati.com/>)
- Bookmarks (Delicious: <http://del.icio.us/>)
- Books (Librarything: <http://www.librarything.com/>, Amazon: <http://www.amazon.com/>)
- Emails (Gmail: <http://mail.google.com/>)
- Events (<http://www.goingtomeet.com/>)
- People (Tagalag: <http://www.tagalag.com/>)
- Pictures (Flickr: <http://www.flickr.com/>)
- Podcasts (Odeo: <http://odeo.com/>)
- Videos (YouTube: <http://www.youtube.com/>)

Even perhaps tagging of tags? (<http://tagtagger.com/>)

In user tagging, after an account has been created a user can apply a tag (or label, or keyword) to a resource; it may be a website, a photograph or video, or a record for a book as in Librarything.

The user chooses a tag that is meaningful to him or her. In most sites it has to be a single word – more about that later. A large number of tags can be applied (e.g. in Flickr the maximum number at the time of writing is 75). Once the tags have been assigned, they act as index terms and they may be public or private. When they are public, the tags together can all be searched by all users, creating a “folksonomy”.

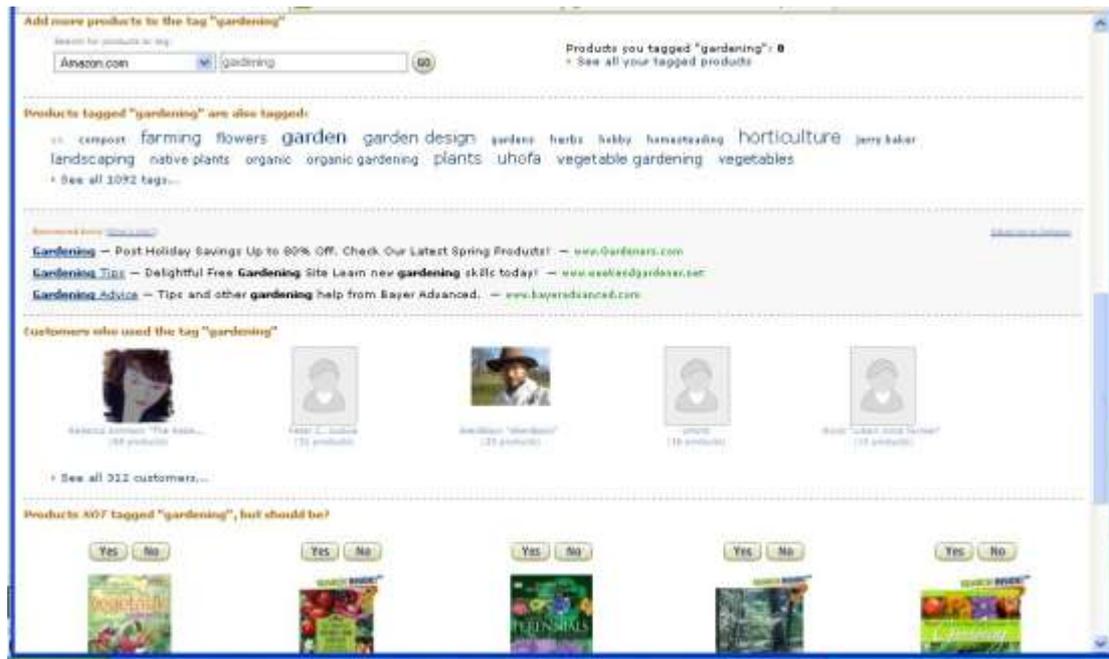
It is important to remember that users have complete freedom in the tags they choose and may assign tags for their own organising purposes, without regard to any other users who may wish to make use of them. Even if this is the case, there may still be valuable information in the collection of tags that develops. In many cases however, users are keen to share their tags and will choose tags that others have also used.

Users can add their own tags to already tagged resources. They may use a different word for the same concept or a broader or more specific word for a related concept. The aggregation of all the tags allows a site like Flickr to organise resources better for all users, and also informs the site owners about the popularity of tags and of resources. This can be described as a bottom-up rather than top-down building of categories.

Tags, once assigned, can be grouped, shared, displayed, published and managed in several ways. Typically tags are displayed in a “tag cloud” on many sites, where the graphical display indicates by size, font or colour how many times the tag has been used or how many resources have been assigned that tag. It is possible to see all tags assigned to a resource, all people who have used a particular tag, other tags that have been used for similar items, popular tags, recent tags etc. Del.icio.us allows users to manage their tags by bundling them and renaming them, and provides for a rudimentary hierarchical structure using the “/” between tags. Clearly, tagging is potentially a very powerful information management resource.

It is interesting to see what Amazon offers users in relation to tags. See for example this page on Amazon for the tag *gardening*, at <http://www.amazon.com/tag/gardening>





The page shows the name of the tag, how often it has been used and by how many people; it provides information about the first 5 of those 516 products tagged *gardening*, with an offer to see all 516 and for each one the opportunity to tag it or remove it; it shows items in this set that were recently given this tag; it shows the users their own tags and tagged products; it shows a tag cloud for products that have been tagged *gardening*; it shows customers who have used the tag; it suggests other products that could also have been tagged *gardening*.

The last item is interesting as an example of an element of control being added by the service to the collection management that users have done – a suggestion is made of other items that also could be assigned the same tag. Here Amazon's own classification is being used to identify and suggest other items on the same subject – but it is only the users who decide to add a tag.

The selection of options above also illustrates the three distinct entities in the world of tagging: the tag, the item being tagged and the person doing the tagging. They all have a separate existence but vital relationships to each other that can be exploited for information. Amazon has compiled and presented data about all three: the products tagged (and how often and how recently), the tag itself (noting which other tags have also been used with this tag) and the person doing the tagging (others who have also used the same tag). This richness of data has huge potential to provide multiple layers of information, and we discuss later in this paper how we aim to employ a similar architecture in our proof of concept model for tagging educational online resources.

2.2 Language about tagging



In an emerging field, names of new phenomena are not always clear and many different terms are currently used for the activity of tagging and the results of the tagging activity.

Tagging has come to be the most frequently used name for the action of applying a label to an item; however some sites do talk about labels and others enable tagging seamlessly so that it may not be apparent that it is being used.

A definition of tagging from Wikipedia:

“A tag is a (relevant) keyword or term associated with or assigned to a piece of information (like picture, article, or video clip), thus describing the item and enabling keyword-based classification of information it is applied to.”
(Wikipedia, 2007)

The word tag has also been used for some time in the context of HTML (hypertext markup language) where it refers to formatting codes used in HTML documents. HTML tags indicate how parts of a document will appear when displayed by browsing software.

The terminology about user tagging is still fairly fluid and many terms for the same phenomena are being used, often in slightly different ways, with debate starting about the exact usage and meaning of the terms. These terms currently include:

Collaborative tagging, shared tagging, user tagging, social bookmarking, collaborative bookmarking; folksonomies, tagsonomies, tagonomies, collabularies, tagosphere, folksonomic zeitgeist.

The term folksonomy was coined by Thomas Vander Wal in 2004 to signify what he called a “user-generated classification, emerging through bottom-up consensus”. It is a fusion of the words folk and taxonomy.

There is debate about the nature of these concepts and terms. Some writers have distinguished between a *folksonomy* (a collection of tags created by an individual for personal use) and a *collabulary* (a collective vocabulary). Other writers however use *folksonomy* to mean a collective vocabulary.

In this paper we use *tagging* to refer to labelling of web items, *user tagging* when that tagging is done by the user, and *folksonomy* to refer to the collection of user tags.

2.3 Different approaches to tagging



Elements to consider are: who does the tagging?; is it collaborative? is it intended for categorisation? what use is made of the tags?

Different sites use tagging in different ways. Tagging is not always done by users, as in the sites listed above. In some cases tagging is done by “experts” with the results still displayed as a tag cloud. An example of this approach is Surf the News.com (http://news.com.com/2243-12_3-0.html) where News.com editors tag each published story with one or more topics. Here the tagging is done not by users but by news editors. In the tag cloud, the most assigned topics are large and red and those used least are small and gray. The data is sorted by most-used topic to least-used topic. The data is shown for the last 30 days of stories. If a topic is not assigned within that time, the topic will not appear on the page. This is tagging but not folksonomy.

Another approach to tagging is to develop an application that will take RSS feeds and analyse the keywords in the feeds to produce tag clouds. An example of this was

NewsCloud (<http://www.revsys.com/aboutnewscloud/>), an application that took all of the RSS feeds from the Washington Post website and built a tag cloud from the keywords. Each story's full text was pulled from the website and indexed by these keywords. There were typically around 23,000 news stories and 71,000 keywords being indexed at any given time. This is an example of machine tagging.

A similar application designed for people to put on their own sites to create tag clouds is ZoomClouds (http://zoomclouds.egrupos.net/cloud/ZC_CNN/). A web site owner can use ZoomClouds to put small (or large) tag clouds on a website and control the look and feel of the tag cloud to provide visitors with a way to see what terms are more often mentioned in the website. Again, these are examples of tagging but not folksonomies.

Even when user tagging is the activity on the site, different users will be behaving differently. Some users choose tags they wish to make public and to share. Other users keep their tags private and simply use this as a way of storing and organising information for their own purposes. You can have a completely private area in Del.icio.us as a way of managing your own bookmarks simply so that you can access it from anywhere but without making it a social networking activity. You can even search other people's tags if you wish and still not share your own.

Other users may make their tags public but deliberately choose private words so that they know few others will use them and they can keep control over a manageable number of items tagged. For example, a recent blog posting stated:

“Tags are not for taxonomists. They are not for the creation of some enormous world-wide content-finding system. They are PERSONAL. It's nice that they can provide some benefit in a social context, but I know what *my* 'frob' tag specifies, and I frankly don't care what anyone else's does. When I want to find something, I use *search*.” (Feinberg, 2007).

In many cases tags are designed to be shared but only within a small group or already defined community which may have its own idiosyncratic use of language and wish to limit the number of users of the tags in order to control their own information access and management. This is a type of folksonomy but for a restricted group.

The *education.au* team working on the *myedna* project share a common invented tag, *myednapoc*, to inform each other about websites and resources of interest for discussion. (See <http://del.icio.us/tag/myednapoc/>).

Our view is that search facilities and collection management can be enhanced and enriched by noting and analysing what users' tags do. Purely personal ones tags, like the example *frob* above, and narrow interest group folksonomies would be excluded through low frequency but it can be of value to know what those users who do like to share their tags choose as the most meaningful words. We can thus obtain information about current linguistic usage as well as about topics of current interest.

Many tagging sites cater for multiple levels of interaction and tagging activity. Although private uses of this type of facility are also valid, it is, we argue, only when tags are publicly shared that a folksonomy develops.

2.4 Folksonomies

A folksonomy is essentially the name given to a collection of tags built up by the action of user tagging, effectively a user generated taxonomy as opposed to an authoritative hierarchical taxonomy like Library of Congress Subject Headings or a subject thesaurus. A key feature of a folksonomy is that tags may be reused many times, providing information about the popularity of the tags themselves (which synonyms come to be more popular over time) as well as information about emerging areas of interest.

The essence of folksonomies is that the tags allocated are chosen by the user. This is a fascinating new development for those of us who have been working in the field of controlled vocabularies and we are keen to see what terms users will choose for tags, how they will use them and how they will organise them. What new vocabularies and taxonomies will emerge from these clouds of tags developed and used by anyone who has an interest and inclination to do so?

There are many studies already emerging in this area. These include an examination of numbers of people using tags (Rainie, 2007), a project investigating the potential of folksonomy in academia (TNN, 2007) and a comparison of the use of tags in Librarything and Amazon that has generated considerable discussion (Spalding, 2007). A study analysing the structure of collaborative tagging systems found “regularities in user activity, tag frequencies, kinds of tags used, bursts of popularity in bookmarking and a remarkable stability in the relative proportions of tags within a given url.” (Golder and Huberman, 2006).

It is worth reviewing some features of folksonomies and comparing them to formal classification systems.

2.4.1 Folksonomies vs. formal classification

2.4.1.1 Benefits of folksonomies

User tagging and folksonomies

Folksonomies: benefits

- Multidimensional
- Meaningful words
- Meaningful concepts
- Shared
- Simple and straightforward
- Community development
- Large scale brings organisation
- Source of information about users

education
what's equal is shared

The following are characteristics of tagging and folksonomies that can be seen as beneficial features.

They are multidimensional: users can assign a large number of tags to express a concept and can combine them.

Users can use their own language: words that have meaning for them. These words are likely to be current and reflect local usage.

Users select concepts that have meaning for them as individuals and analyse items to highlight what is important to them.

Tags can be shared, creating knowledge through aggregation.

“We now have millions and millions of people who are saying, in public, what they think pages and images are about. That's crucial information that we can use to pull together new ideas and information across the endless sea we've created for ourselves.” (Rainie, 2007)

Instead of having to store an item in a single folder, it can be tagged with many different terms and each of these could be used to generate an instant collection (e.g. if a collection of photographs contains photographs with tags such as birthday, family, holiday, Europe, sub-collections can be readily assembled by searching for single tags or pairs.)

Public tagging has been described as having an altruistic appeal, allowing people to contribute to a shared knowledge base. Social tagging fosters the development of communities around similar interests and viewpoints.

Social tagging provides information to professional providers and managers of information about areas of interest and how they are being described. It is a new window on the way our users are thinking and can provide insight into their information needs and habits.

Tagging is very quick, simple and straightforward. Users can apply tags without formal training in classification or indexing.

Clay Shirky has identified a number of advantages of tagging systems, including:

“Market Logic: [...] where you deal with individual motivation, but group value.

User and Time are Core Attributes: [...] because you can derive 'this is who this link is was tagged by' and 'this is when it was tagged, you can start to do inclusion and exclusion around people and time, not just tags. You can start to do grouping. You can start to do decay.

Signal Loss from Expression: [...] in a world where enough points of view are likely to provide some commonality, the aggregate signal loss falls with scale in tagging systems, while it grows with scale in systems with single points of view.” (Shirky, 2005)

As folksonomies grow, the larger scale can bring some organisation into the tagging process; judicious users will evaluate tags and tend to use existing tags to assist with forming useful connections. Thus the folksonomy can develop its own tagging conventions through group consensus rather than an externally imposed and possibly dated formal system.

2.4.1.2 Disadvantages of folksonomies

User tagging and folksonomies

Folksonomies: disadvantages

- Poor choice and application, inconsistency, lack of objectivity
- Tags with personal meanings used
- Different terms for same concept
- Same term for different concepts
- Tags may change as trends evolve
- Different language issues
- Tags can be a mixture of types, genres, formats etc
- Rules not applied (eg hyphens, spelling, singular/plural)
- Tags only single word in some systems
- Tagging vulnerable to malicious practice
- Tags may over-represent dominant view
- No formal structure linking tags

education

It will be readily apparent that many of the features of folksonomies listed above as advantages can also lead to problems for effective classification and information management.

The simplicity and ease of use of tagging can result in poorly chosen and applied tags. While it could be argued that this is a necessary feature of user tagging and insignificant, nevertheless, the following issues need to be considered.

Tags can be applied at different levels of specificity by different users (or even by the same user at different times) e.g. the tag *cats* may be used in one case and *animals* or *pets* in another. Or the tag *Kitty* may simply be used.

Different terms may be used for the same concept (again by different users or by the same user – users will not necessarily be consistent). So *felines* may be used for some items and *cats* for others.

A person searching for pictures of cats will have to use many different terms to be sure of finding all items.

Tags with personal meaning only are frequently used (example on Flickr: *viewfrommywindow*). This tag on its own is of virtually no use to anyone else.

Conversely, the same term can be used for different concepts. Typically, no information about the meaning of a tag is provided (although some systems, del.icio.us being one, do allow tag descriptions; see <http://del.icio.us/help/tagdescriptions>). The word *play* could occur in an educational resource collection in the drama context or the games context. The word *tag* itself has more than one meaning. Without even considering the issue of other languages,

English itself has a huge number of words with multiple meanings. Vocabularies have been built for specific communities where the meanings chosen are appropriate for that context. To some extent this will also apply in user tagging on the internet, but even within communities there can be ambiguities of meaning.

Uncontrolled tagging can result in a mixture of types of things, names of things, genres and formats. Many of these problems can arise even with specialist indexers, for example using *video* as a subject heading when the item is a video, when it should only have that subject heading if it is *about* videos. If it is already difficult for people to comply with requirements such as these, it will be far more difficult to have precision when there are no indexing guidelines other than those developed by individual users for their own practice and unlikely to be made explicit.

Regular indexing and cataloguing rules such as singular vs plural forms, use of hyphens and spelling conventions are not established in a folksonomy.

People's choice of tags may change as new trends evolve — e.g., it is likely that *blog*, *weblog*, *blogs* and *blogging* will all be used for the same concept.

Many systems only allow single word tags. It may be difficult to assign terms to complex concepts using only a single word and running two or more words together is difficult in many ways – the resulting words will be highly idiosyncratic and difficult to read and to search with precision.

Social tagging systems are vulnerable to spam and malicious practice.

A more subtle issue is that people may behave differently (consciously or unconsciously) when tagging other people's items as opposed to their own. The objectivity of a professional indexer is not necessarily a feature of social tagging.

Another high level concern is that over time tags may come to represent a dominant view, discouraging usage of less popular concepts (and terminology) which become disproportionately overwhelmed by the majority. Users will tend to use popular tags and may not realise that there is a more precise term available for their concept, or may be subtly discouraged from creating their own tags. Zeldman writes: "Network effects being exponential, what is immediately mildly popular quickly becomes artificially very popular, while what has yet to become popular never will be." (Zeldman, 2005)

Overall, social tagging and tags are uncontrolled and tags are not connected to each other by a reference structure, which in formal systems is used to link related terms and narrower or broader terms. The creation and application of tags by users who are not experts in information management leads to the problems described above.

However there are also clearly great benefits in user tagging and folksonomies, especially in the richness, currency, relevance and diversity of the terms used, and the collections of resources created. It is important to try to retain those qualities in any attempt to control folksonomies.

2.4.2 Combining folksonomies and formal classification

User tagging and folksonomies

Combining folksonomies and formal classification

Some possibilities:

- New ways to harvest tagging
- Intersection of tags with social networks: rating tags
- Hybrid systems

Some examples:

- Penntags (academic library where users tag items in the catalogue and share tags) <http://tags.library.upenn.edu/>
- Ann Arbor District Library (public library catalogue where users rate, review discuss and tag items) <http://www.aadl.org/catalog>



Is it possible to combine the two approaches and gain benefits from both? Some attempts have been made already and a few are mentioned here in a consideration of some future developments for social tagging. We then discuss our own model: the taxonomy-directed folksonomy for the *myedna* proof of concept.

In some ways it is too early to tell what will be the results of the Web 2.0 developments and in what directions social tagging and folksonomies will go. One prediction by David Weinberger about social tagging:

“Because it's useful when there's lots of information and the information is truly meaningful to individuals, it'll be adopted more and more widely. But we're also going to invent new ways to harvest tagging. Flickr, for example, is already able to cluster photographs by subject with impressive accuracy just by analyzing their tags, so that photos of Gerald Ford are separated from photos of Ford Motor cars. We'll also undoubtedly figure out how to intersect tags with social networks, so that the tags created by people we know and respect have more ‘weight’ when we search for tagged items. In fact, by analyzing how various social groups use tags, we can do better at understanding how seemingly different worldviews map to one another.” (Weinberger, 2007)

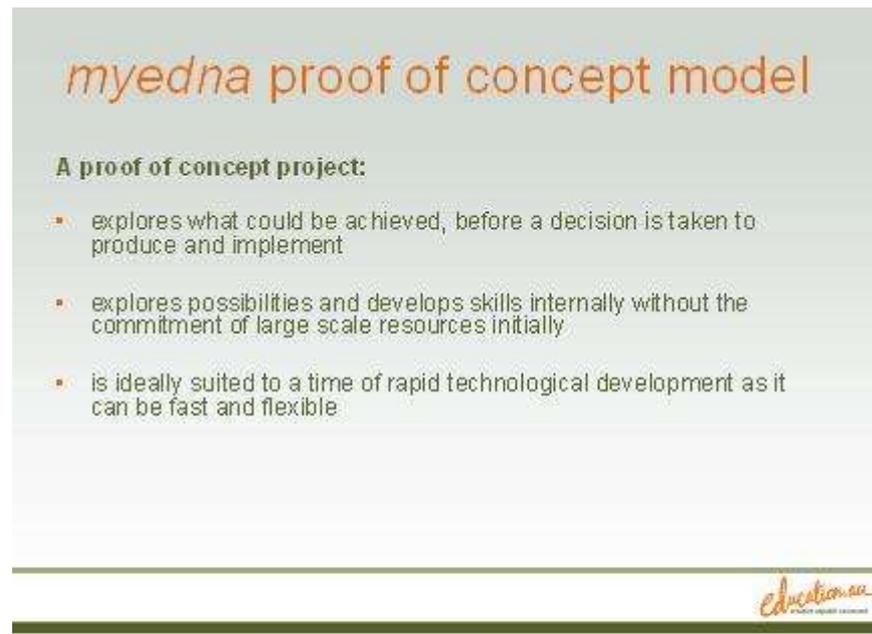
Another development, described in Wikipedia:

“Although ‘tagging’ is often promoted as an alternative to organization by a hierarchy of categories, more and more online resources seem to use a hybrid system, where items are organized into broad categories, with finer classification distinctions being made by the use of tags.” (Wikipedia, 2007)

In the public library sector, Ann Arbor District Library (AADL) has developed a set of social networking tools called the SOPAC, integrated into the library catalogue. It gives users the ability to rate, review, comment on, and tag items. (Blyberg, 2007)

3 A taxonomy directed folksonomy proof of concept model

3.1 Nature of proof of concept



As mentioned above, *education.au* aims to develop a *myedna* service where users can contribute, customise, manage and share their own resources. As a new development in the Web 2.0 era this will aim to make the best use of collaborative technologies and philosophies. We have begun this process by developing a proof of concept model to test the concept, develop the skills of the team and explore possibilities without being limited by our capabilities at this early stage.

A proof of concept model is a model that explores what could be achieved, before a decision is taken about whether it is produced and implemented. It explores possibilities and develops skills internally without the commitment of large scale resources initially. It is ideally suited to a time of rapid technological development as it can be fast and flexible.

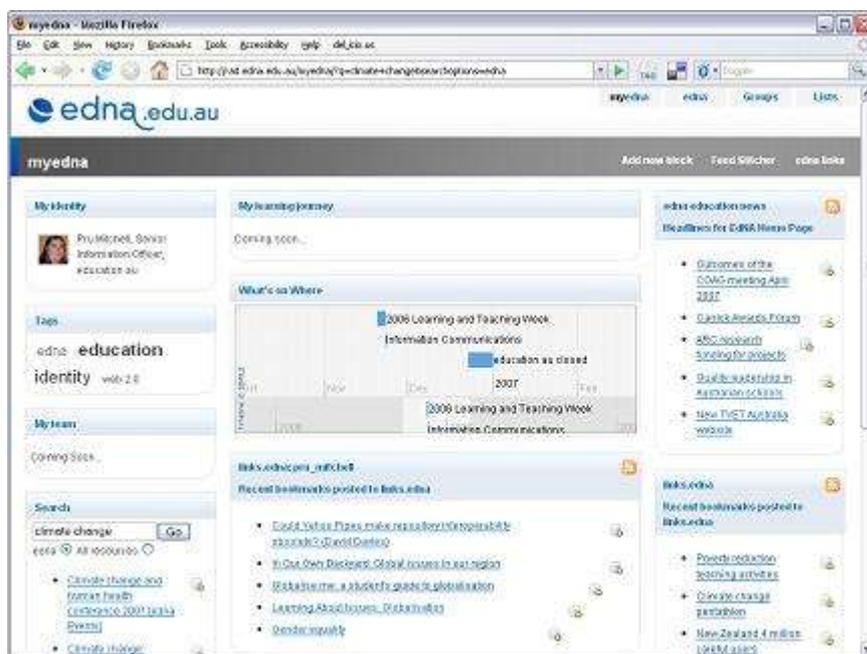
“[It is] a high risk, high trust and low governance project that creates a conceptual solution for the client. Focus solely on the concept. Forget security. Forget firewalls. Forget QA etc. What is delivered is a bare-knuckled prototype which demonstrates how it could solve the issues your client deals with. The project has no limits other than Time (approx 3 months) and Resource a small team working part time within a small financial budget.

Concluding each POC month is a ‘Show and Tell’ to an audience of sponsors, users and internal teams - each one being more public than the last.

Project Outcomes:

- Innovation: new ideas and new approaches from more directed conceptual thinking.
- Better starting position - much closer to the solution suited to the user needs.
- Better position to ask really pointed questions and address issues such as accessibility, security, privacy and quality.
- Agile in response to ideas and feedback particularly with shorter feedback times” (Cotton, 2007)

3.2 Key features of the *myedna* proof of concept



The aim is to develop a service based on the notion of sharing learning or “Watch me learn”. It will be a personal learning space that will:

- record a learning journey
- accommodate all types of learning: formal, informal, “a-ha moments”, work-related
- provide an online space organised by a person to meet his or her needs
- remind users what they have learned
- allow people to share information, resources, stories and narrative (in many formats) with multiple audiences
- record assets only once, and reuse them, catalogue, edit, comment and share for different contexts.

The interface will be customisable by the user to reflect individual needs, interests, preferences and whims. As a function provided by a national education service, it

should be a public space so that each person's personal learning journey will add to the community's knowledge and learning. It will be a portal containing a range of features, one of which will be a facility to "stitch" RSS feeds together. It will provide for storage of documents, images and other files that users will be able to access from anywhere any time.

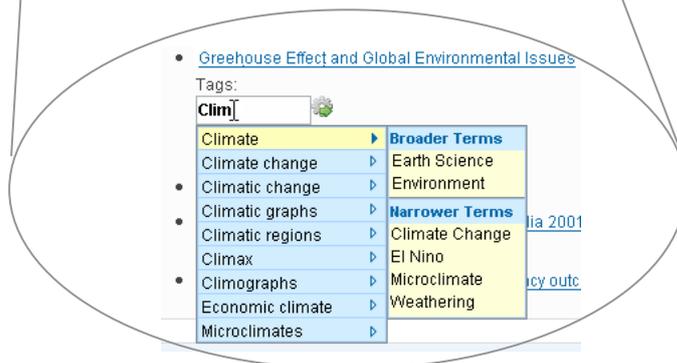
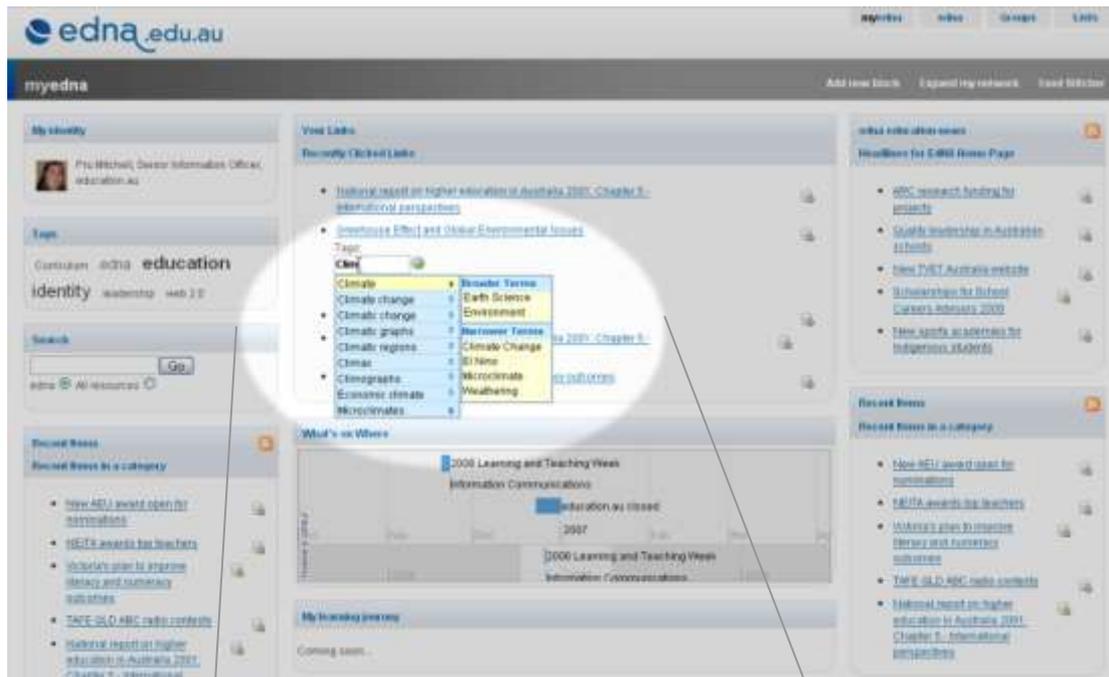
It will also include a section called **edna Links** where users can collect, store and label information about favourite sites as bookmarks using tags. Our proof of concept model for this area includes the feature we have named the Taxonomy-directed Folksonomy.

3.3 The taxonomy-directed folksonomy

Tagging was described above as a type of subject indexing, performed by users, without a controlled vocabulary. In our *myedna* proof of concept model we are aiming to combine user tagging with a controlled vocabulary and harness the best of both worlds. A recent article about the concept of Library 2.0 stated:

"tags and standardized subjects are not mutually exclusive. The catalog of Library 2.0 would enable users to follow both standardized and user-tagged subjects; whichever makes most sense to them. In turn, they can add tags to resources. The user responds to the system, the system to the user. This tagged catalog is an open catalog, a customized, user-centered catalog. It is library science at its best." (Maness, 2006)

This is similar to the philosophy underlying the proposed *myedna* development.



In the **edna Links** area of the *myedna* portal, users will be provided with a box in which to enter their own tags for resources. As they type a tag, they will be prompted by a thesaurus, which will suggest terms that match the term they have entered.

The algorithm for matching terms is still under development. We are exploring ways to mesh taxonomy based functionality with the ease of use that users are familiar with from a social tagging context. Taxonomy based functionality which we are looking at includes:

- exact match, suggestions for broader or narrower terms, or prompt with a term which contains the user's term in its scope note
- choice of thesaurus or one may be supplied to fit a user's signon if he or she is from a particular educational sector.

We are using the Schools Online Thesaurus (ScOT) - used for schools for this project. **edna** makes use of a number of specialist education thesauri in addition to ScOT, including VOCED - a vocational education research thesaurus and ATED (Australian thesaurus of education descriptors) – particularly suitable for higher education. As

each of these thesauri adheres to the controlled vocabularies standard, ANSI/NISO Z39.19-2005 (National Information Standards Organization, 2005), it would be technically feasible for users to be given a choice of thesauri for tagging.

Users may still choose to use their own terms. Tags will be collated over time and a tag cloud produced and displayed. The tags in the clouds will come both from user tags and from tags selected from thesauri. This collection of tags will be a folksonomy that has been directed by a taxonomy.

A possible future development would be to give users information about the tags as they are about to choose one – so they not only see related tags but also scope notes or guidelines for usage. This would give further taxonomic direction to the folksonomy.



The folksonomy thus created will generate valuable information.

Information about the tags

it will indicate which thesaurus terms are useful to our users
it will indicate new terms for existing concepts that should be considered for our thesauri (either as preferred or non-preferred terms)
it will indicate new concepts and suggest terms for them

Information about the items tagged (resources)

it will indicate which items are considered of value by our users simply because of the number of times they have been tagged, and, if ratings are included, which are valued by valued taggers

Information about the people doing the tagging

it will indicate what tags and items a person has used, and each person will have a profile about their learning journey

In our current model for **edna Links** we propose to display results for each of the three facets or dimensions we have identified: Tags, Resources and People. If you select a tag you will see all items assigned that tag and all people using that tag. If you select a resource you will see all tags assigned to that resource (as a tag cloud), all people who have tagged that resource, and ratings given to that resource. If you select a person you will see all tags they have used and all resources they have tagged.

We have also included the facility for users to make comments on resources, tags and people, rate resources and contribute to threaded discussions.

Guy and Tonkin have written:

“One missed area of opportunity is that of more discussion tools through which users can share reasons for tagging things in a certain way. At the moment there is little discussion on folksonomy sites about the appropriateness of tags. Most of the sites do not offer the opportunity to provide actual text feedback, though some allow you to change other users' metadata.” (Guy and Tonkin, 2006)

In this proposed model we have built in the opportunity for users to: identify, bookmark and evaluate resources of interest to their community; choose tags from an appropriate existing formal taxonomy; suggest new tags; comment on the tags; comment on the resources; find other users with similar interests; discuss any of the above.

3.4 Future directions



Future directions

- myedna* is a project that aims to develop a model for a culture of active user participation
- The taxonomy-directed folksonomy is one of the key features that will allow us to make that aim a reality
- We hope to move from a proof of concept to a fully developed facility and build a taxonomy-directed folksonomy over time populated with real data.

education.au

myedna is a project that aims to develop a culture of active user participation enabling our organisation to adapt itself, interact with our community and respond to the needs of our users through a participatory cycle of feedback, service development and reevaluation.

The taxonomy-directed folksonomy described above is one of the key features that will allow us to make that aim a reality. It is our hope that it will move from a proof of concept to a fully developed facility so that we can test it with our users and build a taxonomy-directed folksonomy over time populated with real data. The information we collect about the terms our users choose or create, and suggestions they make about the terms, will inform future formal taxonomy developments, especially in areas of new terminology and concepts, such as elearning. The information we receive from them about resources will enable us to enhance our collection development and management to ensure up-to-date and extensive coverage.

Our anticipated outcomes are:

- Educators will be able to manage their educational resources with tags meaningful to them and others
- A greater number of resources will be identified and shared by the wider **edna** community
- There will be more consistent categorisation of resources
- Resources identified by the community will come to light for full cataloguing by *education.au* specialist indexers
- The **edna** community will form fluid social networks and discussions around tags and resources.



In this project we have aimed to make the best possible use of emerging technology and current practice in the Web 2.0 environment by exploring how we can unite it

with best practice in formal classification and information management to improve outcomes for the Australian education community.

4 References

Anderson, Paul, *What is Web 2.0? Ideas, technologies and implications for education*, JISC, 2007 (<http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf>)

Ann Arbor District Library (<http://www.aadl.org/catalog>) (Accessed April, 2007)

Beckett, Dave, 2006, “Semantics through the tag”, XTech 2006: “Building Web 2.0” — 16-19 May 2006, Amsterdam, The Netherlands (<http://xtech06.usefulinc.com/schedule/paper/135>)

Blyberg, John, “AADL.org Goes Social” post to *blyberg.net* (<http://www.blyberg.net/2007/01/21/aadlorg-goes-social/>) (January, 2007)

Cotton, Tom, “The really shiny part of my job - Proof of Concepts” post to *Tom Cotton's 8109*, March 2007, (<http://blogs.educationau.edu.au/tcotton/>)

Fienberg, Jonathan, *Poundbang* blog comment, 2007, (<http://poundbang.in/2007/02/24/measuring-efficiency-of-tagging-with-entropy/#comments>) (Accessed April, 2007)

Golder, Scott A. and Huberman, Bernard A, *The structure of collaborative tagging systems*, 2006 (<http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf>)

Guy, Marieke and Tonkin, Emma, “Folksonomies: tidying up tags”, *D-Lib magazine*, January 2006, v. 12, no. 1 (<http://www.dlib.org/dlib/january06/guy/01guy.html>)

Lothian, Nick, 2007, *Nick @ Education.au: Nick's Education.au Weblog* (<http://blogs.educationau.edu.au/nlothian/>)

Maness, Jack, 2006, “Library 2.0 Theory: Web 2.0 and Its Implications for Libraries”, *Webology*, Volume 3, Number 2, June, 2006 (<http://www.webology.ir/2006/v3n2/a25.html>)

National Information Standards Organization, 2005, *ANSI/NISO Z39.19-2005 Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies* (http://www.niso.org/standards/standard_detail.cfm?std_id=814)

Porter, Joshua, “The Del.icio.us lesson”, post to *Bokardo.com* (<http://bokardo.com/archives/the-delicious-lesson/>) (May 2, 2006)

Rainie, Lee, 2007, *28% of Online Americans Have Used the Internet to Tag Content*, Pew Internet and American Life Project (http://www.pewinternet.org/pdfs/PIP_Tagging.pdf)

Shirky, Clay, 2005, "Ontology is Overrated: Categories, Links, and Tags", in *Clay Shirky's Writings About the Internet*,
(http://www.shirky.com/writings/ontology_overrated.html)

Spalding, Tim, 2007, "When tags work and when they don't: Amazon and LibraryThing" post to *Librarything*,
(<http://www.librarything.com/thingology/2007/02/when-tags-works-and-when-they-dont.php>) (February, 2007)

TNN -Tags / Networks / Narrative, blog (<http://www.hum.dmu.ac.uk/blogs/tnn/>)
(Accessed April, 2007)

University of Pennsylvania, *Penntags*, (<http://tags.library.upenn.edu/>) (accessed April, 2007)

Vander Wal, Thomas, *vanderwal.net* (<http://vanderwal.net/>)

Weinberger, David, 2006, comment to *Joho the Blog*,
(http://www.hyperorg.com/blogger/mtarchive/penntags_when_card_catalogs_me.html
1)

Zeldman, Jeffrey, "Remove Forebrain and Serve: Tag Clouds II",
(<http://www.zeldman.com/daily/0505a.shtml>) (4 May, 2005)