

Snips & Spaces: Managing Microlearning

Christian Langreiter, Andreas Bolka
synerge development services

Abstract: Radically lowering "barriers to publish", wikis and weblogs are rapidly gaining acceptance as simple and hassle-free ways to share and link information in a community of interest (or overlapping communities of interest). Based on a working definition of microlearning as learning from microformats, we discuss the characteristics of both formats and outline problems that may arise in a microlearning context. We propose that by combining both formats to form an integrated whole, those problems can be largely solved. This is complemented by a description of several aspects of Vanilla, a system based on this idea.

1 Microlearning

We understand microlearning primarily as learning from microcontent¹ - from "*small pieces, loosely joined*" (Weinberger 2002).

Microlearning as a term reflects the emerging reality of the ever-increasing fragmentation of both information sources and information units used for learning, especially in fast-moving areas which see rapid development and a constantly high degree of change.

While in the past a single authoritative work (or even a single authoritative teacher) may have been all that was necessary to sufficiently acquaint oneself with a given topic of interest, this is increasingly untrue, especially as the necessity to (quickly) learn (a lot) extends into almost everyone's work life.

Books, magazine articles, a multitude of web resources (like online books, tutorials, encyclopedias, forum and weblog postings, emails and comprehensive teaching material collections as produced by MIT's OpenCourseWare² project or the Connexions³ effort hosted at Rice University) form essential ingredients of the source mix of almost any non-institutionalized learning effort - and, increasingly, of many institutionalized efforts as well.

Fragmentation of sources has both positive and negative aspects. From a producer's standpoint, information fragments are much easier to create than larger works.

¹ As defined by Nova Spivack:

http://novaspivack.tyepad.com/nova_spivacks_weblog/2003/12/defining_microc.html
(accessed Jun, 2005)

² <http://ocw.mit.edu/> (accessed Jun, 2005)

³ <http://cnx.rice.edu/> (accessed Jun, 2005)

Furthermore, disaggregated content - theoretically - can be re-aggregated to optimally suit an individual learner's preferences (instead of the needs of an idealized common denominator). The other side of the coin is that a significant fraction of the consolidation and organization effort is shifted towards the learner.

It will increasingly be the task of microlearning management systems to assist the learner (or group of learners) to consolidate information gleaned from such disparate sources into a coherent whole. We see personal knowledge mapping as enabled by combined wiki/weblog software as a first step in that direction.

Based on our working definition of microlearning, a microlearning management tool should assist an individual learner or a group of learners in

- stating a plan of learning,
- representing recognized concepts and the relations between them,
- attaching relevant information to those concepts,
- and, ideally, guide further exploration.

2 Weblogs

2.1 Introducing Weblogs

In only a couple of years, weblogs have gone from fringe to mass phenomenon. The format of reverse-chronological short articles with lots of links has existed since the earliest days of the web, but only in 1999 the terms "weblog" and "blog" gained traction; one of the effects was a surge in self-awareness (some might say self-obsession) of the then small community of webloggers.

Theoretically, weblogs can be created without dedicated tools at all. A simple text editor suffices. However, dedicated weblogging software like Blogger, Manila and Antville emerged quickly, relieving publishers of routine chores like uploading HTML pages to a server and manually moving articles to archive pages. Those tools played an essential role in the popularization of the format. In Rebecca Blood's words: "[...] *the bandwagon-jumping turned into an explosion*" (Blood 2000).

Soon, the simple format began to evolve, and today features like accepting comments from visitors, "permalinks" and trackbacks are almost universally supported. Some of those concepts warrant further discussion.

Permalinks ("permanent links") emerged when people began to refer to other people's postings. When linking only to the frontpage of a weblog (as was common in the beginning), the referenced posting might have disappeared into the archive already when a reader followed said link (often providing crucial context necessary to understand the referencing posting). Permalinks allow authors to address individual postings (or even smaller units of information like paragraphs) permanently, whether they are still visible on the front page or not. This allows readers to track cross-weblog threads of discussion.

Trackbacks enable authors to notify other weblog owners of the fact that they referenced one of their postings.

Another noteworthy development which originated within the weblog community are "really simple syndication" standards like RSS and Atom. Those were developed primarily in order to make the consumption of a large number of weblogs feasible in the first place (Robert Scoble⁴: "*But, remember, I read 1000 blogs [...]*"), relieving the reader of having to surf through a long chain of weblogs again and again - only to find out that on many, nothing new had appeared since the last visit.

In the context of education, many lecturers have embraced weblogs as a natural communication medium providing everything from slide sets, supplemental links touching on relevant issues discussed in the last lecture to administrative like cancellation announcements.

For additional background on weblogs, the interested reader is referred to Blood (2000), Winer (2003) or the German article by Praschl (2001).

2.2 Problems with Weblogs

The biggest problem of weblogs in the context of microlearning management is that in many cases, the potentially valuable and relevant archived body of information is not used to full effect.

Remembering, finding and referencing archived postings are not particularly well-supported activities in most weblog publishing systems, the root cause being the limited facilities commonly provided to relate and organize individual items.

Apart from the format-native organizational principle of time (e.g. monthly archives), most current weblog publishing systems either

- do not support categorization at all (Blogger),
- only allow a single category per posting (Antville) or
- make creating new categories comparatively cumbersome, with the predictable result that the set of said categories often remains relatively small and static.

Overly coarse-grained categorization in turn results in overloaded category overview pages, severely limiting their usefulness.

Recently, ad-hoc tagging as known from del.icio.us and Flickr has been implemented for weblog publishing systems as well⁵. If current practice as observed on del.icio.us is an indicator, it can be expected that this scheme will result in significantly finer-grained categorization.

Another issue is that representing lexical knowledge – information about companies, products, people, books, etc. and the relations between them – does not fit well with the mainly narrative-focused approach traditionally encouraged. This limits the usefulness of weblogs as personal knowledge mapping tools.

⁴ <http://radio.weblogs.com/0001011/2004/11/28.html> (accessed Jun, 2005)

⁵ See e.g. "Jerome's Keywords" by Jerome Lavigne: <http://vapourtrails.ca/wp-keywords> (accessed Jun, 2005)

3 Wikis

3.1 *Introducing Wikis*

A wiki is a web application that enables any user to easily add to and modify a collection of interlinked snippets of information. We argue that three aspects are fundamental to a wiki:

- Snips and spaces
- Effective support for content creation and modification
- Ridiculously easy linking

3.1.1 Snips and Spaces

A wiki is generally understood to be a web site where any user is allowed to add, modify or delete a significant fraction of the content present. The following definitions from two of the most important sources regarding wikis capture that common ground very well:

"A Wiki [...] is a web application that allows users to add content, as on an Internet forum, but also allows anyone to edit the content."
— Wikipedia⁶

"[A wiki is a] freely expandable collection of interlinked Web 'pages,' a hypertext system for storing and modifying information – a database, where each page is easily editable by any user with a forms-capable Web browser client."
— Leuf & Cunningham (2001)

Beyond this general agreement, terminology used to describe and discuss wikis is ambiguous. The term wiki itself is widely used to refer to both

- a particular web site that embraces the wiki fundamentals – as in "the C2 wiki" and
- a particular web application that can be used to maintain such web sites – as in "the Vanilla wiki".

We use the term "space" to refer to the former. A space is a web site (supported by wiki technology) that consists of a collection of pages or snips. We strongly prefer the term "snip" for two reasons: it disambiguates discussion between web-pages (pages) and wiki-pages (snips) and it aptly depicts the typical nature of wiki-pages in a personal learning context – small snippets of information related to a specific concept.

3.1.2 Effective Support for Content Creation and Modification

The "effective" here refers to two things: content management facilities must be easy to use and easy to discover. Discoverability (and therefore a low entrance barrier) is typically achieved by placing an "edit" button prominently on each page (see Figure 1). Snip creation facilities vary among wiki implementations, some enforce a policy that new

⁶ <http://en.wikipedia.org/wiki/Wiki> (accessed Jun, 2005)

snips ought to be created only when the (not-yet existing) snip is linked to from some other snip – wikis typically present a “create” link in those situations (see Figure 2). This policy stems from the belief that all wiki snips should be interlinked – no insular snip shall exist. Other implementations simply provide an additional “create new snip” button.

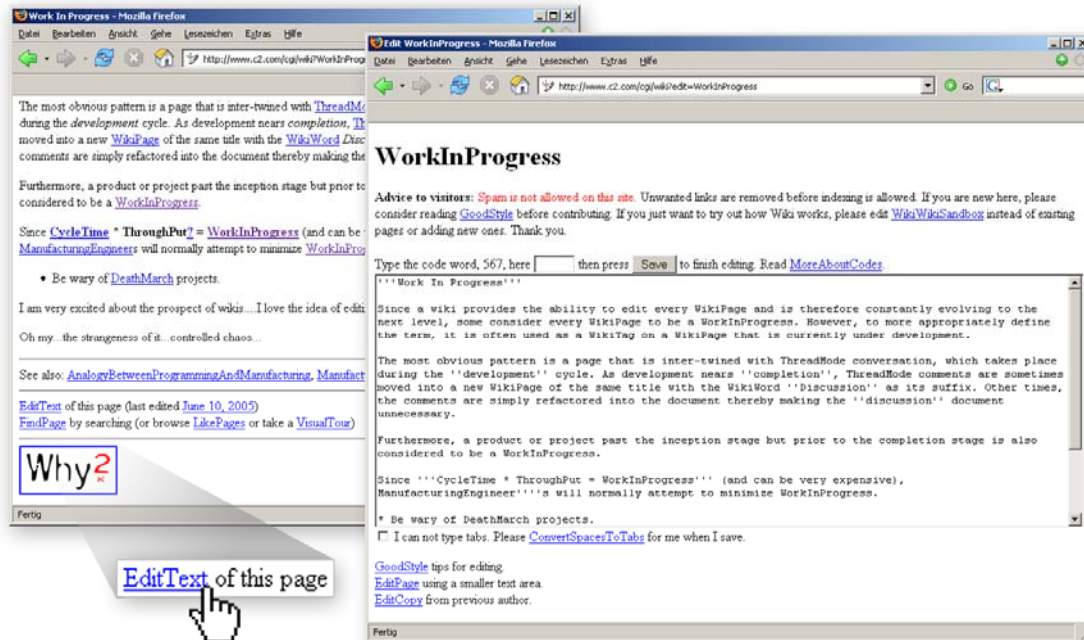


Figure 1: The infamous "edit" link

To facilitate easy editing, a simple markup language for snip content is usually provided. No standard has emerged yet, so wiki implementations differ substantially in the markup capabilities provided – but some functionality is common to most. Important, obviously, are means for linking to other snips. Traditional wiki implementations like the C2 wiki⁷ use the CamelCase convention for linking, under which snips are named by joining capitalized words (a snip on the Portland Pattern Repository which discusses this convention is named JoinCapitalizedWords, for example). As this convention is certainly counter-intuitive for wiki novices, we will not discuss it and the associated advantages and problems here.⁸

More recent implementations abandoned the CamelCase convention for something less idiosyncratic. Snips can be named freely and links to other snips can be established by surrounding the name of the target snip with minor markup like two rectangular brackets (MediaWiki), asterisks (Vanilla) or similar. So to link to a snip named “Microlearning 2005” one would simply write `[[Microlearning 2005]]` or `*Microlearning 2005*` (cf. the right part of Figure 2).

⁷ The first Wiki, created by Ward Cunningham in 1996 - <http://www.c2.com/cgi/wiki> (accessed Jun, 2005)

⁸ The following places provide an introduction to the discussion:

- <http://www.c2.com/cgi/wiki?JoinCapitalizedWords> (accessed Jun, 2005)
- <http://en.wikipedia.org/wiki/CamelCase> (accessed Jun, 2005)

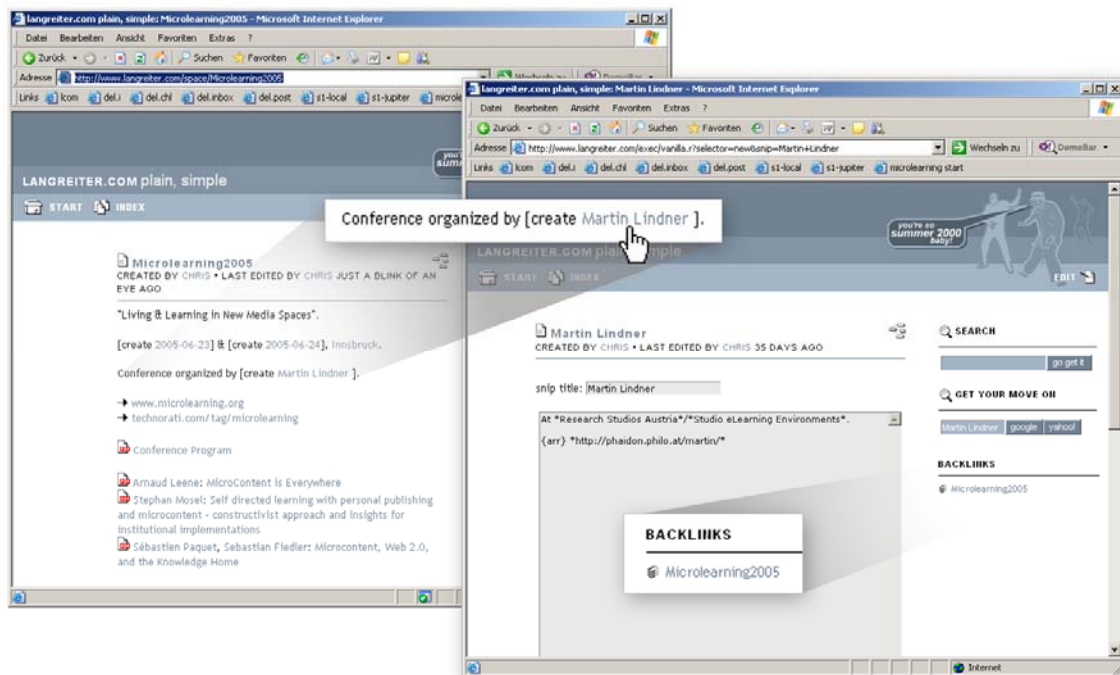


Figure 2: A create link, link markup, backlinks

Beyond link markup, typically various text formatting options are supported. Capabilities encountered range from simple things like boldface or italics formatting to support for laying out complex tables or even creating diagrams. Furthermore, some implementations allow the use of full HTML for advanced users, others provide a WYSIWYG editor.

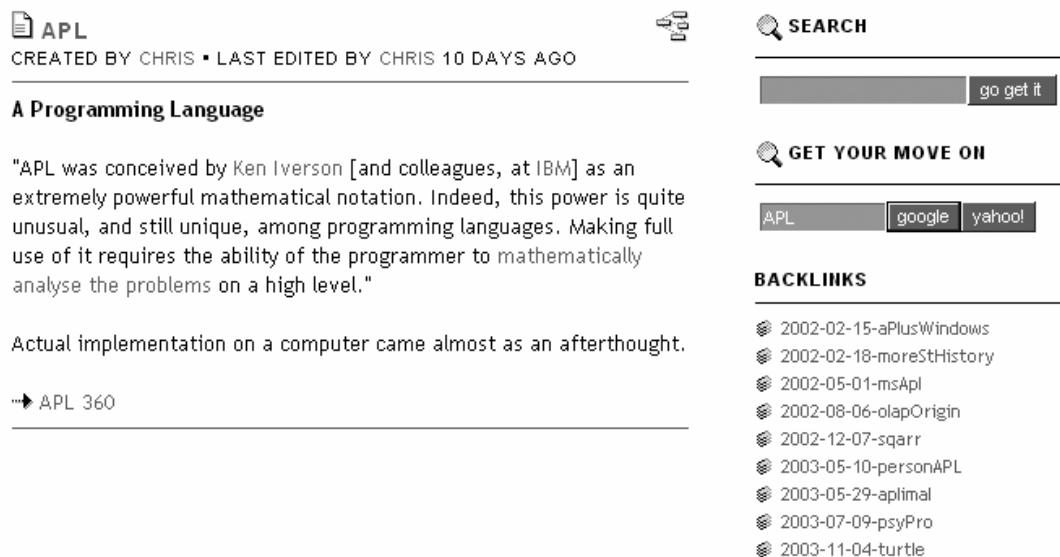
3.1.3 Ridiculously Easy Linking

While the “anyone can edit everything” functionality is well-understood and prominently featured as a fundamental wiki principle in almost every article discussing wikis, we argue that another aspect is at least as important. Over a decade after the creation of the web, we got used to interlinked information. The key to wikis is that they make fundamental hypertext ideas so ridiculously easy to access that applying them becomes second nature. Creation and maintenance of links is a no-brainer in a wiki. Links – and the functions enabled by them – are central to the power of wikis.

“We should work toward a universal linked information system”
– Berners-Lee (1989)

The importance of the link was obviously appreciated by Tim Berners-Lee in his seminal work that led to the creation of the World Wide Web. But wikis are able to provide functionality that goes beyond what unidirectional HTML links can support. As a wiki space is a closed, self-contained entity, links between snips are fully bidirectional, i.e. not only outgoing links of any snip are known, but also the incoming links (cf. Figure 3).

Those are usually referred to as “backlinks.” We will expand on the power of backlinks in Sect. 4.



The image shows a screenshot of a wiki page for "APL". The main content area includes a title "A Programming Language", a paragraph describing its history and use, and a link to "APL 360". The sidebar on the right contains a search box, a "GET YOUR MOVE ON" section with search engine buttons for "google" and "yahoo!", and a "BACKLINKS" section listing various dates and links.

APL
CREATED BY CHRIS • LAST EDITED BY CHRIS 10 DAYS AGO

A Programming Language

"APL was conceived by Ken Iverson [and colleagues, at IBM] as an extremely powerful mathematical notation. Indeed, this power is quite unusual, and still unique, among programming languages. Making full use of it requires the ability of the programmer to mathematically analyse the problems on a high level."

Actual implementation on a computer came almost as an afterthought.

→ [APL 360](#)

SEARCH

go get it

GET YOUR MOVE ON

APL google yahoo!

BACKLINKS

- 2002-02-15-aPlusWindows
- 2002-02-18-moreStHistory
- 2002-05-01-msApl
- 2002-08-06-olapOrigin
- 2002-12-07-sqarr
- 2003-05-10-personAPL
- 2003-05-29-aplimal
- 2003-07-09-psyPro
- 2003-11-04-turtle

Figure 3: Backlinks displayed in a sidebar

“Intertwingularity is not generally acknowledged – people keep pretending they can make things deeply hierarchical, categorizable and sequential when they can’t. Everything is deeply intertwined.”
— Nelson (1974)

It is through the clever use of hypertext features a wiki achieves its flexibility. Links can be used to establish almost any kind of structure over the underlying content snippets, and indeed, even multiple structural views can be provided in parallel.

Classical hierarchies pose a barrier to quick content creation. Unless something fits perfectly into one of the available drawers, attention of the content creator is drained by the ever-nagging question “now where the heck shall I put this?” Using bidirectional linking, a hierarchical structure can easily be established. The original Portland Pattern Repository (the “C2 wiki”) employs backlinks to categorize snips⁹, a pattern later adopted by Wikipedia¹⁰. We will briefly describe how this works in Sect. 4.

“A Wiki, a free form, unstructured space, is like a room. A room can be used for anything. For meetings, for sleeping, for swimming (provided it has a pool), but a room is just a space. And to the people using the room, that space becomes a place. A place to situation themselves, a place to call home, a place to organize themselves, or a place to orient from.”
— Meatball¹¹

⁹ cf. <http://www.c2.com/cgi/wiki?CategoryCategory> (accessed Jun, 2005)

¹⁰ cf. <http://en.wikipedia.org/wiki/Wikipedia:Category> (accessed Jun, 2005)

¹¹ <http://www.usemod.com/cgi-bin/mb.pl?WikiAsRoom> (accessed Jun, 2005)

The three fundamental principles of the wiki result in an extremely flexible medium, aptly described on Meatball wiki by analogy to a room.

Contemporary wiki implementations supplant those basic features with a slew of useful extensions. While the totally open nature is appropriate in some situations, personal wikis or wikis supporting small project teams often require access control, a feature found in almost every modern wiki implementation. Another widespread piece of functionality is snip revision control, which not only aids against vandalism in totally open wikis but also supports cooperative document creation or similar tasks in project scenarios.

Full-text search with results sorted by relevance is a natural and important complement to the hypertextual “browsing” typically used to retrieve information from a space. Especially useful in wikis targeted at personal learning support or smaller project teams is an attachment feature, where arbitrary files can be attached to a snip. This effectively extends the wiki to also become a document repository; current implementations are, however, typically rather basic.

A list of “recent changes”, i.e. pages that were created/modified during a certain time period, helps users in watching the activity on a wiki. “Cross-pollination” from weblog publishing tools resulted in the availability of the recent changes list as RSS feed or similar formats. Further, comment systems in the style of those found on weblogs or web forums are incorporated into some wikis to replace the discussion method used in traditional wikis, where users wanting to discuss a certain paragraph in a snip simply append their contribution indented right below the paragraph in question.

Plug-in mechanisms and/or macro facilities allow wikis to be used as small-scale web application development frameworks and provide powerful tools for experimentation. In Vanilla, for example, most of the features (snip attachments, full-text search etc.) are implemented as plug-ins on top of a slim wiki “core”.

3.2 Problems with Wikis

Traditional wikis not enriched with weblog functionality suffer from a set of problems which are mostly inherent in the format (“problems of format”). Furthermore, wikis in general also suffer from a “problem of perception”. We use this to describe the phenomenon that wikis are widely perceived as exclusively collaborative tools. This perception obviously hinders adoption for personal microlearning purposes and can only be overcome by further research and proper communication of the possibilities of wikis in this context.

A major problem with the format is the lack of guidance. Using the “Wiki as Room” analogy mentioned before, people confronted with an empty room do not immediately appreciate what can be possibly made of that room. In another situation, a person entering a multi-floor building is certainly not aware of “what’s going on” in that building.

While helping users with initially getting acquainted with a corpus of content is typically achieved through “entrance” or “guided tour” snips, users who actively want to watch the progress on a given wiki are usually left with only the recent changes lists. In active wikis

where lots of different threads are developing simultaneously those recent changes lists can quickly get overwhelming. Voß (2005) measured the change rate of the English Wikipedia at an average of 16 changes per minute and while this is certainly an extreme example, it makes obvious the fact that even far less active wikis cannot be completely monitored by a single user. Traditional wikis lack effective means for providing a condensed, edited summary of what is happening inside the space.

"[People] either fall in love with the use of hypertext on Wikis or they run away screaming"
— Angeles (2004)

People trying to establish a personal wiki are often intimidated by the “empty space” they are initially confronted with. Questions like “what should I put into it?”, “how should I name things?”, “how can I structure my thoughts?” result in a considerable entrance barrier. Those questions arise because a wiki can be used for almost everything (at least that is how they are typically promoted). This naturally leads to confusion – what is “everything”? Those who already drank the wiki kool-aid are regularly baffled by those problems – they have already found their personal nail to use the wiki-hammer for.

But even advanced users are confronted with problems inherent in the format. A wiki has no natural or obvious place where information snippets that are to be recorded for later reference (clippings) can be stored. Every page stored in a wiki must be named, so there is a certain burden on the user to come up with proper names. But some pieces of information simply resist attempts to be named at the moment they are recorded – as in situations where the user has no or limited knowledge of the field of information the piece is related to. Traditional wikis have no practical facility to handle ephemeral notes and observations.

And finally, even a feature as fundamental as bidirectional linking is underrepresented and underused in traditional wikis. Usually, backlinks are not displayed along with the content of a snip but are only accessible through an explicit request by the user (for example by clicking on the snip title). Once requested, only the backlinks are displayed, so the user has to switch between content and context. This severely reduces the value of backlinks as “context providers”.

4 Combining Wikis and Weblogs

We argue that some weaknesses of the two formats – wikis and weblogs – can be overcome simply by combining and integrating them. This claim is supported by several years of experience with using, developing and observing the use of the first combined wiki/weblog tool, Vanilla¹². Such combined wiki/weblog systems are also commonly referred to as “Blikis” or “Wikilogs”.

Vanilla was originally developed by Christian Langreiter as a testbed for experiments in web application design in late 1999 and has since been developed by Christian Langreiter, Andreas Bolka and a group of contributors. The original inspirations were exposure to the Swiki wiki implementation and the release of Userland’s Manila content

¹² <http://www.vanillasite.at/> (accessed Jun, 2005)

management system, which was one of the first CMSes to embrace weblog functionality as conceptual cornerstone.

It soon became obvious that both weblog postings and wiki snips are fundamentally similar. Both are small snippets of information (hence the term snip), and by adopting a simple naming convention (YYYY-MM-DD) for weblog snips, a couple of lines of plug-in code ("dynasnips" in Vanilla parlance) would be sufficient to appropriately render weblogs in the familiar reverse-chronological order.

A significant portion of common wiki functionality can be reframed and understood as a "dynamic view" onto the contents of a space: recent changes are a limited (filtered) view sorted by the date of last modification. An index is a full view, sorted lexicographically. Backlinks are a view of snips which link to the current snip. Full-text search results are an ad-hoc view defined via user-provided criteria. And, ultimately, a weblog is nothing but a limited view over snips following a certain naming convention, sorted reverse-chronologically.

It should be noted that this paradigm of decomposing information into small, easily consumed, easily (re-)arranged units and then proceeding to provide dynamic views upon such collections is rapidly gaining favor. On the decomposition side this is evidenced by the enthusiasm surrounding microcontent formats/microformats (hCalendar etc.). Dynamic views as supplement (and in many cases, replacement) to the rigid, hierarchical folder structures of desktop operating systems, mail clients and other "productivity" software have recently been implemented in mainstream software like Apple's iTunes ("Smart Playlists"), MacOS X Finder ("Smart Folders") and various email clients ("Virtual Folders", e.g. in Novell Evolution, Mozilla Thunderbird or TheBat!).

4.1 Benefits of Adding Weblog Functionality to Wikis

Lower barrier of entry: When introducing wikis and weblogs to new project teams, it has been observed again and again that of the two formats, weblogs provide the lower barrier of entry. Usually it takes a while before opportunities to use wiki functionality are recognized; the value of weblogs as project communication tool however is intuitively obvious upon first contact.

By starting with weblogging in a learning context, the situations where a user wants to gather background information related to certain concepts naturally arise. In a tightly integrated system like Vanilla, the writer of a weblog entry only has to surround a reference to the concept to be expanded on with asterisks – after storing the weblog entry, a link will be displayed that allows to immediately create the wiki snip.

It is through this combination of easy-to-grasp weblog functionality and the seamless availability of wiki functionality that a smooth progression towards using the full power of the wikis is enabled.

Narrative glue: Weblogs are often used to provide contextualizing narrative to an effort. In a combined wiki/weblog, this narrative can serve as the glue between otherwise disconnected information fragments. This also helps to overcome the problem of overwhelming recent changes lists by providing a human-edited view, a window into the wiki space. Jon Udell (2001) precisely describes the value of such a "storytelling" facility.

“[Weblogging is] a powerful new way to tell stories that refer to, and make sense of, the documents and messages that we create and exchange in our professional lives.”

— Udell (2001)

Short-term memory: The problems related to storing ephemeral information in a sensible way are effectively eliminated by using the weblog as “dump” for interesting news and information snippets with appropriate links to the rest of the space. Through those links, these small notes are interconnected and therefore retrievable through a variety of means (full-text search, browsing) or even automatically visible through a permanently visible backlinks list.

An interesting effect arises out of the fact that a weblog view typically lists only a certain number of postings (e.g. the most recent 10 postings or the postings of the last month). As such postings are visible on the front page for a limited amount of time only; the weblog can be regarded as the “short-term memory” of a wiki. Ephemeral weblog postings stay around for a few days and then vanish (metaphorically speaking) into the subconsciousness of the space.

4.2 Benefits of Adding Wiki Functionality to Weblogs

Obviously, the place for storing lexical knowledge that is missing in a weblog is provided by the supplemental wiki functionality. As soon as weblog entries (“weblog snips”) are linked to other snips (“wiki snips”), a number of interesting synergetic effects can be observed.

The aforementioned problems of remembering and referencing weblog entries are alleviated. The archived body of information in weblog entries stays readily accessible – backlinks from non-weblog snips to weblog snips contextualize weblog information, which in turn leads to serendipitous remembering of postings. Referencing a weblog posting (which is nothing but a snip following a certain naming convention) is as easy as referencing any other snip.

In contrast to many other wiki implementations, backlinks are always visible in Vanilla. In combination with the weblog functionality, a special tool called the “backlink browser” easily allows to track contact with certain topics over time (see Figure 4).

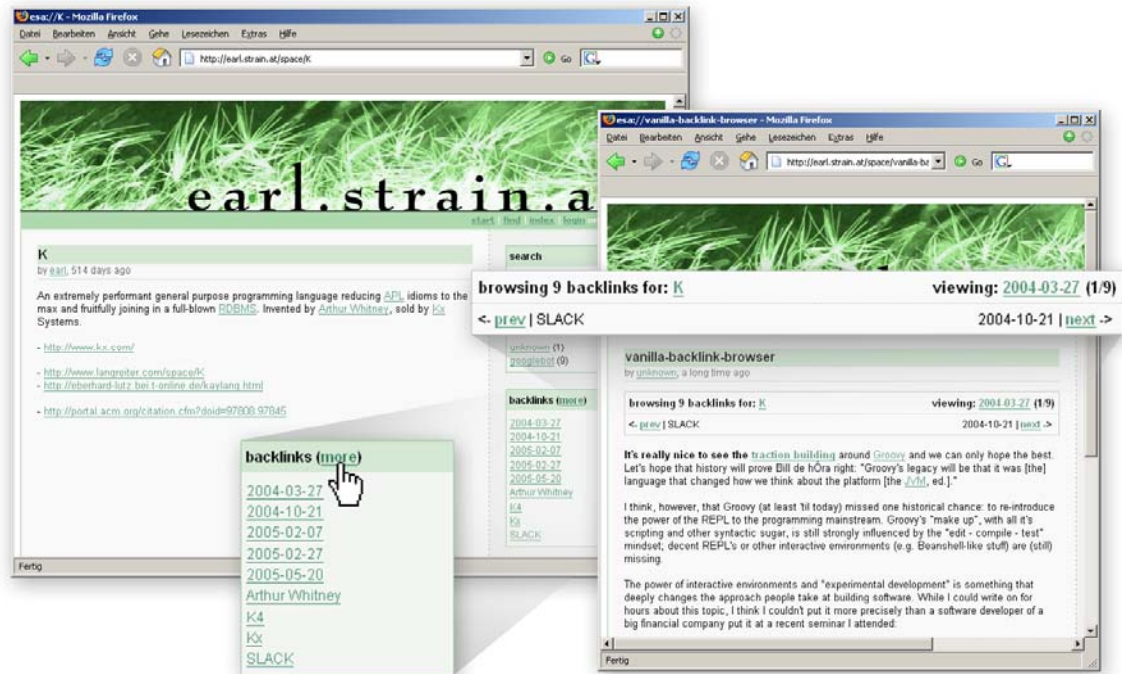


Figure 4: The backlink browser

Generally, bi-directional links provide a rather surprising flexibility in building all kinds of information structures. Categorization can be maintained by simply linking “content” snips to “category” snips. For example, a snip about “K” would link to a “Category Programming Language” snip. The backlinks of a category snip then automatically form the list of all snips belonging to a category. Even “tagging”, as has recently become popular, can be subsumed by this functionality with ease: no one restrains the author from placing multiple links to different “category” snips. Add sweet user-interface sugar consisting of powerful pivoting and browsing functions and have your delicious, tag-enabled wiki served ice-cold.

5 Future Directions

5.1 Profile-driven Information Filtering

Interest profiles have been used to filter and personalize large information collections for a long time, with varying degrees of success. Two axes along which profile-driven filtering systems can be categorized are:

- Mode of data collection - implicit (side-effect of primary user action, like buying a book) or explicit (solicitation of ratings).
- Nature of data to be filtered - system-supplied or user-supplied.

Amazon.com regards an individual customer's purchasing and viewing histories as interest profile; the data is therefore generated as side-effect of the primary actions of looking at product details and purchasing products. Through item-to-item collaborative

filtering (Linden, Smith & York 2003), product recommendations are generated. Users cannot reference products not already in Amazon.com's database.

Findory¹³ is a "personalized newspaper" service being developed by Greg Linden, one of the prime architects of Amazon.com's recommender systems. By reading news and/or weblog articles, an interest profile is implicitly built up. It is possible for users to add new syndication feeds.

del.icio.us¹⁴ is a "social bookmarks manager" having rapidly gained in popularity over the past year. The central organizing metaphor is "tagging". An arbitrary number of tags (lightweight keywords) can be assigned to every bookmark posted (popular tags: "web", "design", "music", "css", "java"). What makes this service interesting from an information filtering perspective is the multitude of possibilities to filter bookmarks - by user, by tag or even by a combination of tags. This allows users of the service to discover other people interested in the same topics (and, prerequisite, using the same tags). As in the case of snips and spaces, we can interpret the set of all tags used by a user as an interest profile.

Google Personalized is an experimental search service based on technology originally developed at Kaltix (Haveliwala, Kamvar & Jeh 2003). It requires the user to explicitly select from several dozen domains of interest before search results can be personalized.

The totality of all snips contained in a space can be regarded as an extraordinarily fine-grained interest profile of the individual or group contributing, conditioned on the specific context the space was created and used for. As the main components of such a profile are subjectively relevant concepts represented by individual snips, we propose the term concept-centric interest profiles. When combined with data gathered through usage tracking (logged edit/view events), the profile can be seen as an accurate time-dynamic model of equally time-dynamic user interest.

Concept-centric interest profiles can be used to

- seed automatically repeated, long-running web queries ("agents") using search engine APIs,
- filter high-volume streams of information (world news, RSS feeds or even emails),
- automatically generate (and then weight) combined-term queries,
- discover people with similar interest profiles.

Discovery of people with similar interest profiles (people matching) could be implemented based on simple concept intersection counts, maybe allowing for some fuzziness in spelling. More sophisticated implementations could use a variation of the well-known TF/IDF algorithm (cf. Baeza-Yates & Ribeiro-Neto 1999) to weight individual concepts. In this case, concept counts over all spaces examined should determine term frequency, whereas concept counts local to a given space should contribute to what is usually referred to as document frequency. Similar to the effect when filtering document collections, concepts everyone is interested in would therefore be discounted, whereas rare concepts would contribute heavily to the final similarity score.

¹³ <http://findory.com/> (accessed Jun, 2005)

¹⁴ <http://del.icio.us/> (accessed Jun, 2005)

For example, if it is observed that people who first were interested in a given concept cluster A often progress to map out or otherwise show interest in concept cluster B, the system could present that information to people currently showing interest for concepts in cluster A.

In an educational context, both fellow learners currently dealing with similar topics and potential tutors competent in the area of interest could be discovered in such a fashion.

It should be noted that a similar concept-centric interest profile can be extracted from traditional weblogs as well. This, however, requires the application of sophisticated natural language processing techniques such as key phrase/concept extraction, etc. - for a good introduction see Manning & Schütze (1999). Furthermore, the precision of selecting subjectively valid relevant concepts manually can be (and usually is) much higher than when utilizing automated techniques.

As many more people are using weblogs than wiki/weblogs at the moment, however, the potential value of deriving interest profiles in such a fashion could nevertheless be substantial - despite the higher complexity. As informally captured in Metcalfe's "law" and shown heuristically in Odlyzko & Tilly (2005), the value of a network increases with the number of participants:

"[...] if we have to select a rule [...] that is concise and yet captures many of the key features of communication networks, then we feel that our $n \log(n)$ formula fits the available evidence and is supported by reasonable heuristics."

— Odlyzko & Tilly (2005)

The distinguishing feature of the proposed approach is its inherent simplicity - when using a wiki/weblog, the profile comes for free.

5.2 Automated Syllabus Construction

Above we observed that while microcontent reduces the effort to create and therefore the barriers to publish, it shifts the job of organizing disparate pieces of information into a coherent whole to the learner. For microlearning in a decentralized setting to realize its full potential, the emergence of either organizing intermediaries or powerful tools for semi-automatic organization of materials is a necessity, so that learners can invest most energy in learning - instead of spending disproportionate amounts of time hunting down the proverbial herd of chunk-sized cats.

In a network of wiki/weblog spaces, we can observe publishing and (to a certain degree) consumption patterns, leading to time-varying interest profiles as described in the latter section. By analyzing a sufficiently large number of such profiles, patterns of dependency among topic clusters can be inferred from frequently appearing similar subsequences of the complete event/activity sequence. By linking those sequential dependency patterns to a measure of individuals' learning performances, it will be possible to discover preferential topic exposition sequences; those in turn could then be used to either directly guide new learners or as valuable input for syllabus designers.

A precondition for such a scenario to be realistic at all is critical mass, in terms of both the number of participating persons as well as in volume and level of detail of individuals' learning journaling.

6 Conclusion

We have shown that the combination of weblogs and wikis is, especially in a microlearning context, more than the sum of its parts. Complementing each other, the primarily associative nature of wikis and the more ephemeral, journal-like nature of weblogs allow for and motivate both the constant inflow of new information as well as the gradual build-up of a body of more permanent, lexically structured information, which serves as a meaningful structural backbone for this kind of personal knowledge store. The immense flexibility hypertext systems gain through the power of the link assists in building structures that support multiple means of re-discovering information.

As avenues for further research, we sketched the possibilities of observing learning behaviour to derive material sequencing hints, eventually contributing to automated syllabus construction, as well as regarding the set of concepts and relations recorded as an extraordinarily rich and fine-grained interest profile to be used for calibrating information filtering systems or matching like-minded persons.

7 References

- Angeles, M. (2004), 'Using a Wiki for Documentation and Collaborative Authoring', LLRX.com.
URL: <http://www.llrx.com/features/librarywikis.htm> (accessed Jun, 2005)
- Baeza-Yates, R. & Ribeiro-Neto, B. (1999), *Modern Information Retrieval*, Addison Wesley Longman.
- Berners-Lee, T. (1989), Information Management: A Proposal.
URL: <http://www.w3.org/History/1989/proposal.html> (accessed Jun, 2005)
- Blood, R. (2000), 'Weblogs: A History and Perspective', Rebecca's Pocket.
URL: http://www.rebeccablood.net/essays/weblog_history.html (accessed Jun, 2005)
- Haveliwala, T., Kamvar, S., Jeh, G. (2003). An Analytical Comparison of Approaches to Personalizing PageRank, Technical report, Stanford University.
URL: <http://citeseer.ist.psu.edu/haveliwala03analytical.html> (accessed Jun, 2005)
- Leuf, B. & Cunningham, W. (2001), *The Wiki Way*, Addison Wesley Longman.
- Linden, G. D., Smith, B. & York, J. (2003), Amazon.com Recommendations: Item-to-Item Collaborative Filtering', *IEEE Internet Computing* 7(1).

MICROLEARNING 2005 DRAFT

- Manning, C. D. & Schütze, H. (1999), *Foundations of Statistical Natural Language Processing*, The MIT Press.
- Nelson, T. H. (1974), *Computer Lib (& Dream Machines)*, Distributors.
- Odlyzko, A., Tilly, B. (2005). A refutation of Metcalfe's Law and a better estimate for the value of networks and network interconnections. (Preliminary version, March 2, 2005)
URL: <http://www.dtc.umn.edu/~odlyzko/doc/metcalfe.pdf> (accessed Jun, 2005)
- Praschl, P. (2001), 'Was ist eigentlich ein Weblog', jonet – Das Journalistennetz.
URL: <http://web.archive.org/web/20021229002617/http://www.jonet.org/modules.php?name=News&file=article&sid=17&mode=&order=0&thold=0> (accessed Jun, 2005)
- Udell, J. (2001), 'Telling A Story', Tangled in the Threads, Byte.com, May 2001.
URL: <http://udell.roninhouse.com/bytecols/2001-05-24.html> (accessed Jun, 2005)
- Voß, J. (2005), Measuring Wikipedia, *in* 'Proceedings of the 10th International Conference of the International Society for Scientometrics and Informetrics'.
- Weinberger, D. (2002), *Small Pieces Loosely Joined*, Perseus Books.
- Winer, D. (2003), 'What makes a weblog a weblog?', Harvard Weblogs.
URL: <http://blogs.law.harvard.edu/whatMakesAWeblogAWeblog> (accessed Jun, 2005)