MicroContent is Everywhere!!!

Defining MicroContent
MicroLearning 2006,
Innsbruck - 7 June 2006
Dr. Arnaud Leene
version 0.2

Intro
I would like to start with explaining the subject of MicroContent. This subject is central to this conference, so I will give you a head-start by explaining what MicroContent is. Maybe you are under the impression that MicroContent is something new, but it already with us for centuries. We just did not call it that way. The easiest way to define MicroContent is by illustration. So I will start with some examples.

Business Card

A good example is the business card, known to everybody. You all have a mental model for recognising the structure, the elements, the meaning, their use, of the things you see on a business card. It shows the name and contact information of a person.

What we are interested in, however, is the electronic equivalent of a business card. This electronic, computer, equivalent goes by different names and in my MicroContent Client it is called an address card. As you can see right away it shows much more details than it's physical counter part. Let us have a look at some details of my card. You see my name, my address, multiple e-mail addresses, my personal home-page, my image, etc. Again you are familiar enough with an address card to understand it's details.
This example allows me to illustrate some basic aspects of MicroContent. Each address card is called a MicroContent Item. You see that an Item has structure. This structure is called fields following database terminology. The address card has multiple fields, each with their own meaning. So you see my home email address, my work email address, etc. And a field can even have substructure, so my address has separate fields for street, postal code, city and country. This structure is a central characteristic of MicroContent. This address card is an example of a MicroContent Type. I will show some other MicroContent Types as we go along.

Some fields are present in most MicroContent Types, such as the title (in this case a combination of my first and last name) and a description (called Note in this case). Let me turn to another example, another MicroContent Type.
I guess that everyone recognises this MicroContent Item right away, even if you never cook. We see a title, a description (sometimes called directions). There is a list of ingredients, which is again a substructure consisting of multiple fields. The same is true for the preparation times.

We see a list of tags (called keywords here). Nowadays tags or becoming an integral part, a standard field, of each MicroContent Type.

The result of following the directions of the recipe is represented by an image of the prepared dish. This image is a MicroContent Item in it's own right. Usually it is only linked to a recipe, but it can be embedded as well. You see that the structure of this MicroContent Type is already more complex.
By the way the visual aspect of MicroContent, i.e. the way I show it to you is totally arbitrary. It is up to the viewer how he wants to see an image. A viewer can even decide to suppress certain fields. In the case of this recipe the rating that I gave to the recipe is not shown.

Let us look more closely at an image, another MicroContent Type.

Images

A photo or an image is a totally different kind of MicroContent Type. The image is really the main part of an Item. The image field, so to say, is the dominant field. It is what it is all about. Usually additional information is associated with an image, such as shown here. So we see a creation date, size, information on the exposure, etc. There might be titles, descriptions, ratings, etc. In the context of an image we might call this information metadata. An image has however structure just as in other MicroContent Types, the role of the fields is just different. What we should call metadata in the context of MicroContent is not quite clear.

Technically this metadata is embedded in the image file, i.e. in the format. And this points to another aspect of MicroContent: self-containing. The idea is that it is possible to manipulate MicroContent Items (moving, exchanging) without losing anything.
You might notice the size of this image and conclude that this not a small Item in terms of number of bytes. This seems to conflict with the word Micro. And it does if you look at the original use of the word. The word Micro is deceiving here, it should be about the focus of an Item. MicroContent Items should be about one thing, single messages, in this case the image of the ape. MicroContent has focus. MicroContent Items are indivisible. You can not cut a piece of MicroContent into smaller pieces, without losing its original meaning. We can not take a part of the image and still have an ape on the ground. If you do start cutting, you come in the realm of sampling, quoting or whatever you want to call it.

Audio

This MicroContent Item is an example to show that there does not have to be a visible part. The dominant field of audio, is audio. We have to listen to it. But this Item also has other fields such as the name of the song, the artist and the album. There might be an album cover and much more information. Again one might call this information metadata. Depending on the audio format used the metadata can be stored in the format as well.

There is a final characteristic of MicroContent that we did not discuss yet. How do you recognise a specific MicroContent Item. In the MicroContent Types that I showed, we use human inference. We know that a title of a recipe, points to a specific recipe; we know that the first and last name of an address card points to a specific address card. And in the case of an image, a title might not suffice, we might have to look at the image itself. This addressability is a key characteristic of MicroContent and even more so in the context of Internet.

MicroContent Characteristics

We have now seen enough examples of MicroContent Items of several Types that we can summarise the characteristics of MicroContent in general.

1. Focus - the most important characteristic addresses the term micro. MicroContent is about content that has recognisable focus. MicroContent is about a single subject, it has a single meaning. This characteristic has its roots in one of the original definitions of MicroContent as defined by Jakob Nielsen. In his original article he used it to describe the titles of documents. The definition of Anil Dash also talks about the focus and meaning of MicroContent. Anyway the size in bytes of MicroContent does not matter;

2. Self-contained - the introduction of RSS, in order to syndicate small pieces of text, led to the idea that MicroContent Items must be self-contained, it should be possible to pass Items around, so that they can used again. MicroContent Items can be mashed with other Items in order to create some new, something with new context;

3. Indivisible - for me MicroContent is indivisible. Items have only their original meaning;
meaning, when you use them as a whole. This is most clearly visible with a recipe. It is not possible to cut a recipe into pieces and still have something useful. Some people are working on services based on quotations in the form of paragraphs in order to recreate Roger Nelson's idea of transclusion;

4. **Structure** - a relative new realisation is that MicroContent has structure. When Dave Winer started with blogging, an Item consisted just of a piece of text. Although hidden fields (metadata) such as publication time and author were available. A MicroContent Item will consist of multiple fields and each field will have its own meaning and type. And the more structure the better;

5. **Addressability** - this is a characteristic which is much advanced by Marc Canter. It is the idea that one should be able to find back MicroContent Items. Each MicroContent Item should have a permanent link attached to it, also known as permalink. This permalink should never change and always point to the same piece of MicroContent, so that it always can be found. The idea of a permalink naturally comes from the web. On your computer you might use file names as a pointer to a specific MicroContent Item. Unfortunately we do not know yet, what a permalink should point to. Is it an archived web-page, as is often done at the moment? Or should it be some XML file, which only contains the Item? Can it be an anchor in a web-page? Or something different all together?

In this list of characteristics I presupposed that we talk only about computer data, although the characteristics of MicroContent also work for physical incarnations of MicroContent.

I evaded the creation of a formal definition for MicroContent. Different people still use different meanings for the same word, so we have to wait a bit until things stabilise. But in the mean time, feel free to contribute to the WikiPedia page on MicroContent.

**Wild MicroContent**

For the purpose of clarity I only showed some easy recognisable MicroContent Types. There are however many more MicroContent types. Maybe you missed the best known example, a blog-item. I did not use it as an example, as it has few structure. Other examples are bookmarks, resumes, events, reviews, sudoku puzzles, etc, etc. In fact you can see MicroContent Everywhere.

The application or service you use to manage your MicroContent determines the fields that are part of a specific MicroContent type. And every provider and developer decides on his own setup. Even this might not satisfy everyone. You might want to create your own MicroContent type. I call this Wild MicroContent.

I show here an example of a MicroContent Item, which type I created myself. It is a type that allows me to document sundials that I saw. I used pretty a pretty generic set of fields, such as a title, a description and keywords. To this I added a link to an image, I added a location, specified as longitude and latitude to indicate where I saw the sundial, I added address field, I added a creator field and I might add more as I play with it. And you can create your own MicroContent Types as well!!!!. Have a look
at Reger.com for weird types and place to create your own.

Standards

In order to be able to manipulate MicroContent we need standards. For many MicroContent types we are already enjoying such standards daily. Think here of images, audio and video. But also think of events and business cards. In the past standardisation committees did a lot of work to agree upon fields, formats and encoding.

But now that MicroContent is becoming popular this formal approach is no longer sufficient and we are starting to see grass roots efforts to define standards. RSS for blog Items is such a grass root standard defined by Dave Winer and is now hugely popular. Technorati with microformats is helping to organise a community to define standards as well. Application and Service builders define their own standards. MicroContent Types of all sorts pop-up all over the place. It seems that things are becoming a mess.

Maybe this is not so bad and just a new reality. We have to learn to live with it and find a new approach. The main problem of Wild MicroContent is that it hampers the im- and export of MicroContent. We will not be able to solve this problem by creating formal definitions. We have to look at solutions on other levels. One of the ideas is to use a standard for the definition of a MicroContent type itself. This can be done through the XML Structure Definition standard or more loosely through OPML. This works already within Reger.com.

Recognising MicroContent

In order to wrap this introduction to MicroContent up, I would like to have a look at some MicroContent in the wild. I would like to show you how Joi Ito, one of the fathers of MicroContent, is using MicroContent on his web-page.
The main MicroContent type we see here on this page are his blog items. These have a title, a description, a time-stamp (presumably creation date) and a tag (or category). These blog items can also be digested as RSS and Atom formats.

At the top we see his main categories (main, japanese, moblog, livejournal), these are also blogs, but he decided to separate them on different pages, and not just different tags. One can see such a page split as a special sort of tagging. On the lowest level one uses tags, at a higher level one uses categories, then one splits into separate pages and finally one creates multiple sites.

At the top we see a row of images. These are in fact bookmarks that point to an Image Item on his moblog. Instead of using text links thumbnails are used to present these bookmarks.

If we look at the sidebar we see mainly bookmarks to several other MicroContent types. At the top we see a (random) faceroll, which are thumbnail bookmarks of people that read his blog. These bookmarks point to the web-sites of these people. From a MicroContent point of view there is no difference between these bookmarks and the ones at the top. The only difference is the meaning: moblog versus faceroll. One can see this difference as implicit tags, which are unfortunately made explicit in
a MicroContent field.

Scrolling down we see the category/keyword/tag *Listening*. These are bookmarks to an artists-page at MusicBrainz that Joi Ito is currently listening to.

Scrolling down further we see his Technorati statistics, bookmarks to comments, weblog archive items, recent entries. We see a blogroll, which is a long list bookmarks of blogs of people he reads. We see a list bookmarks of web-pages that he recently added to Del.icio.us. And this list is also available in RSS-format. Below that a list of recent papers he has written. And further down we see a tag cloud, which are really bookmarks to specific searches on Google for a tag on his web-site and tag-pages on Technorati.

What should be clear from this analysis is that bookmarks (i.e. links) are a vital page of a web-page. But that they only have value when there is an explanation next to it in the form of a tag (or whatever you want to call it).

**Not so clear**

To confuse a bit I would like to finish with another MicroContent Type. I conveyed that the characteristics for MicroContent are clear: focused, structured, self-contained, indivisible and addressable. This example follows these characteristics. And I usually call it MicroContent. However look at more example and you will see that the focus can be very broad. We move into the grey zone between Micro- and Macrocontent.

**Web-services and products mentioned**

- Dave Winer - http://www.scripting.com/
- Joi Ito - http://joi.ito.com/
- OPML - http://www.opml.org/
- Reger - http://www.reger.com/
- RSS - http://blogs.law.harvard.edu/tech/rss
- Technorati - http://www.technorati.com/