



FP7-SME-1  
Project no. 262289

## **HARMOSEARCH**

Harmonised Semantic Meta-Search in  
Distributed Heterogeneous Databases



## **Microformat Usage and Specifications**

Due date of deliverable: 2011-07-30  
Actual submission date: 2011-07-30

Start date of project: 2010-12-01

Duration: 24 month

Project funded by the European Commission within the Seventh Framework Programme

X

## Project Acronym: **HARMOSEARCH**

**Project Title:** Harmonised Semantic Meta-Search in Distributed Heterogeneous Databases

**Grant Agreement:** 262289

**Starting date:** December 2010    **Ending date:** November 2012

**Deliverable Number:** Microformat 1.1

**Title of the Deliverable:** Microformat Usage and Specifications

**Lead Beneficiary:** CPR

**Task/WP related to the Deliverable:** WP 4, Task x.1x

**Type (Internal or Restricted or Public):** Public

**Author(s):** Sabine Schneider, Adriano Venturini

**Partner(s) Contributing:** eCTRL, CPR

**Contractual Date of Delivery to the CEC:** July 30, 2011

**Actual Date of Delivery to the CEC:** July 30, 2011

## PROJECT CO-ORDINATOR

Company name: [X+O]  
Name of representative: Manfred Hackl  
Address: Hamburgerstrasse 10/7, A-1050 Vienna, Austria  
Phone number: +43-676-842755-100  
Fax number: +43-676-842755-599  
E-mail: manfred.hackl@xpluso.com  
Project WEB site address: [www.harmonsearch.org](http://www.harmonsearch.org)

## TABLE OF CONTENTS

<b>1 INTRODUCTION</b> .....	<b>4</b>
<b>2 MOTIVATIONS</b> .....	<b>5</b>
2.1 WHY YOU SHOULD ADD MICROFORMATS .....	5
2.2 GOOGLE CONTENT SEARCH .....	6
2.3 PUBLIC FACEBOOK EVENTS.....	7
2.4 NORMAL BROWSERS .....	9
<b>3 HOW TO MICROFORMAT YOUR PAGES</b> .....	<b>10</b>
3.1 SIMPLE TAGGING (VEVENT) .....	10
3.2 MORE ADVANCED APPROACH (VEVENT & VCARD) .....	10
3.3 USING AN ABBREVIATION (INSTEAD OF DIV OR SPAN).....	12
<b>4 LIST OF FIGURES</b> .....	<b>13</b>

## 1 INTRODUCTION

Microformats are small patterns of HTML used to identify special and common kinds of published information like people, places, events, contact data or reviews. They can also be seen as simple markup to extend human-readable data on web-pages used for software-based indexing and search. Microformat adds semantic to web pages, which otherwise contains unstructured information.

With Microformats automated services can “understand” the content of a web page, for example if it is dealing with an event, when and where this event will be. This allows the user to be able to perform a more accurate search, and it enables other software systems to search and perform reasoning on it.

Data can be exchanged, translated and queried among different organizations.

In the following sections, the motivations for enhancing (your) web pages with microformat are described and the specifications for annotating events with a standardized microformat are consequently provided.

## 2 MOTIVATIONS

### 2.1 WHY YOU SHOULD ADD MICROFORMATS

There are various motivations to decide to enhance web pages with a microformat structure.

When using microformats, semantic markups containing some structured information about the content of a web page are directly embedded in the HTML file, and are therefore refreshed automatically as soon as the content of the web site changes.

With respect to provide semantic information via XML, one obvious advantage against XML enrichment is that an XML document has to be maintained *in addition* to the usual web site, and is not refreshed automatically in case if the web site owner changes the content.

This fact, that information is not synchronized automatically, contradicts with the common known principle of reusing data, and not rewriting it. Therefore it might be a good alternative to use microformats instead of semantic XML information, as in this case the information written in the HTML code is automatically reused as semantic information, while in the case of microformats it has to be rewritten entirely.

Some other advantages and benefits of microformats are briefly mentioned in the following.

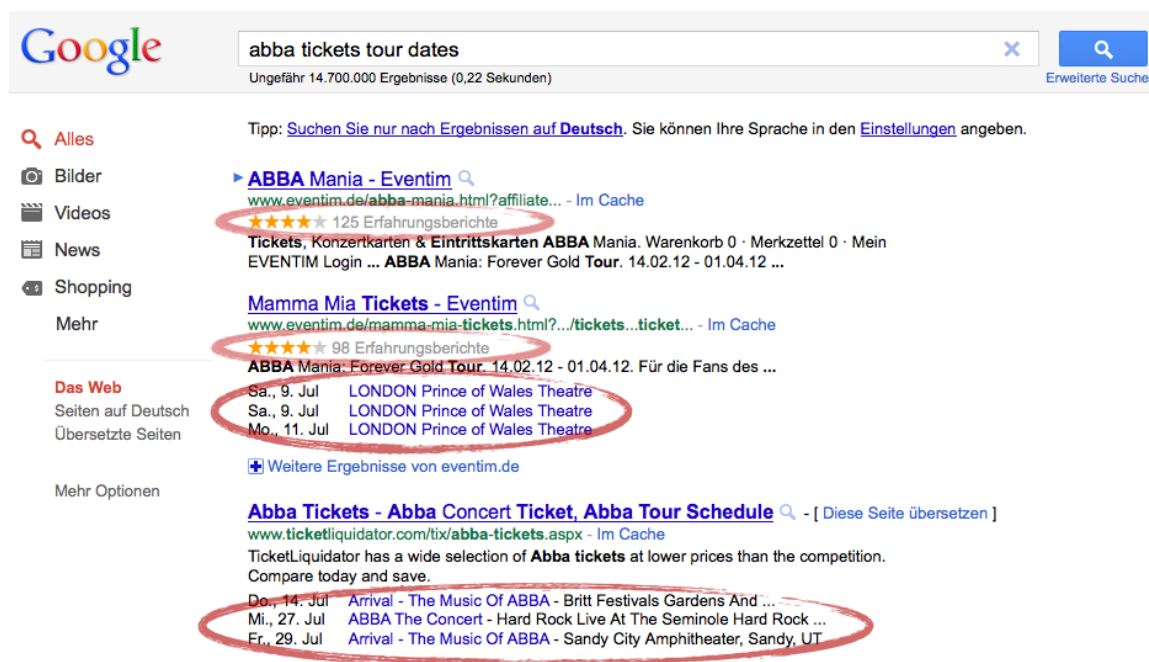
## 2.2 GOOGLE CONTENT SEARCH

Another motivation that made so many web sites in the recent past decide to implement microformats, is that - two years ago - Google has added a functionality to its search mechanism in order to search for microformat hCard and vCalendar entries on sites, and to show them in a special way in the list of search results, as well as on Google Places sites.

As already decades ago some people started to build an XML schema that basically should describe the entire world, or even the universe, Google's idea of adding some semantic information to search results might be a first (larger) step in this direction.

In this way, Google users have the possibility to obtain a structured summary of the main information contained in the search results as so called „rich snippets“: They get more detailed information about title, date and place of an event, which enables them to better distinguish which search results are of interest for them and which are not. This makes it easier for them to decide on which link in the list of search results to click.

Microformats are easy to implement, as well as a very effective SEO strategy.



The screenshot shows a Google search for "abba tickets tour dates". The search results include several rich snippets. The first result is "ABBA Mania - Eventim" with a star rating of 4.5 and 125 reviews. The second result is "Mamma Mia Tickets - Eventim" with a star rating of 4.5 and 98 reviews. The third result is "Abba Tickets - Abba Concert Ticket, Abba Tour Schedule" with a star rating of 4.5 and 12 reviews. The snippets provide detailed information about the events, including dates and locations.

Google

abba tickets tour dates

Ungefähr 14.700.000 Ergebnisse (0,22 Sekunden)

Erweiterte Suche

Tip: Suchen Sie nur nach Ergebnissen auf Deutsch. Sie können Ihre Sprache in den Einstellungen angeben.

Alles

Bilder

Videos

News

Shopping

Mehr

Das Web

Seiten auf Deutsch

Übersetzte Seiten

Mehr Optionen

ABBA Mania - Eventim

www.eventim.de/abba-mania.html?affiliate... - Im Cache

★★★★★ 125 Erfahrungsberichte

Tickets, Konzertkarten & Eintrittskarten ABBA Mania. Warenkorb 0 · Merkzettel 0 · Mein EVENTIM Login ... ABBA Mania: Forever Gold Tour. 14.02.12 - 01.04.12 ...

Mamma Mia Tickets - Eventim

www.eventim.de/mamma-mia-tickets.html?.../tickets...ticket... - Im Cache

★★★★★ 98 Erfahrungsberichte

ABBA Mania: Forever Gold Tour. 14.02.12 - 01.04.12. Für die Fans des ...

Sa., 9. Jul LONDON Prince of Wales Theatre

Sa., 9. Jul LONDON Prince of Wales Theatre

Mo., 11. Jul LONDON Prince of Wales Theatre

Weitere Ergebnisse von eventim.de

Abba Tickets - Abba Concert Ticket, Abba Tour Schedule

www.ticketliquidator.com/tix/abba-tickets.aspx - Im Cache

TicketLiquidator has a wide selection of Abba tickets at lower prices than the competition. Compare today and save.

Do., 14. Jul Arrival - The Music Of ABBA - Britt Festivals Gardens And ...

Mi., 27. Jul ABBA The Concert - Hard Rock Live At The Seminole Hard Rock ...

Fr., 29. Jul Arrival - The Music Of ABBA - Sandy City Amphitheater, Sandy, UT

## 2.3 PUBLIC FACEBOOK EVENTS

Social network sites like Facebook added this format to all public events. These sites of public Facebook events have now the usual microformat vevent (and vcard) format. In the following there are two examples, one having, and one missing the vcard location details.

Google is indexing them.

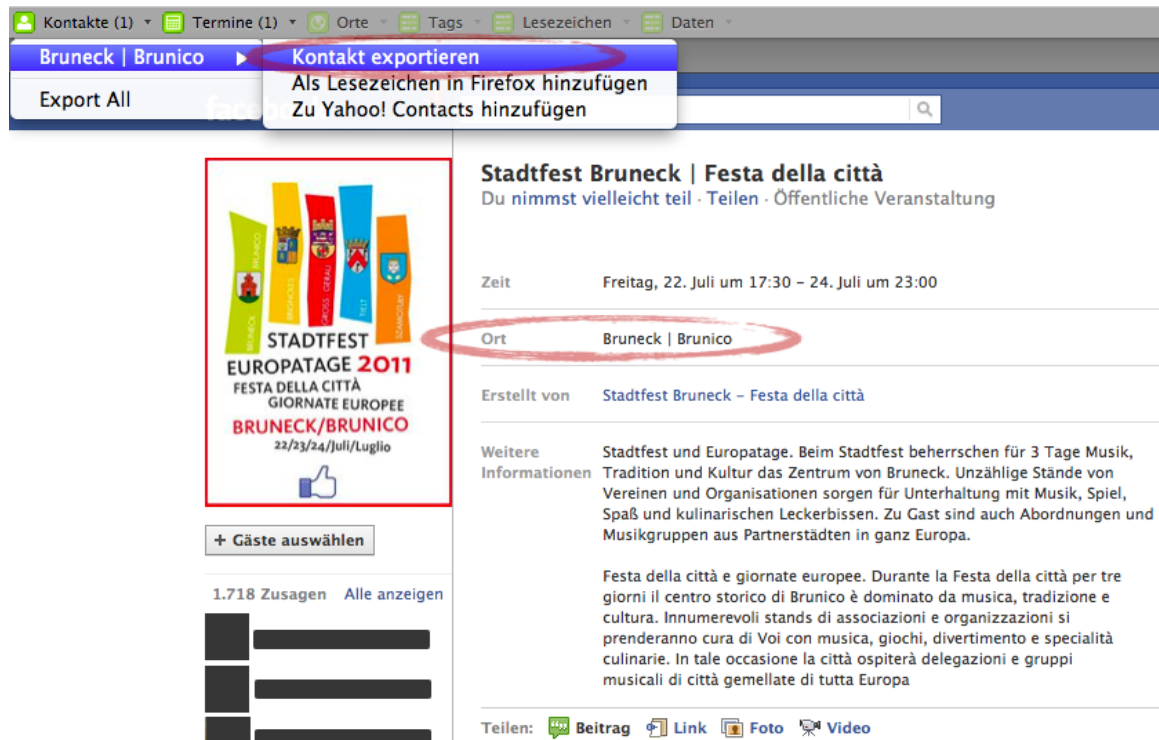
Screenshots showing these events are shown below.

```
<body class="vevent ...">
...
<span class="profileName fn fsxl fwb">REFERENDUM 12.-13. Juni</span>
...
<div>
  <span class="dtstart">
    <span class="value-title" title="2011-06-12T07:00:00"> </span>
    Sonntag, 12. Juni um 07:00
  </span> -
  <span class="dtend">
    <span class="value-title" title="2011-06-13T15:00:00"> </span>
    13. Juni um 15:00
  </span>
</div>
...
<div class="description summary">
  Damit das Referendum gelingt, muessen mindestens 25 Millionen
  Menschen waehlen gehen!<br>Teilt diese Einladung mit FreundInnen,
  mobilisiert und verbreitet die Informationen! <br>... am 12. und
  13. Juni 4 JA fuer die DEMOKRATIE!<br><br>- JA zum Stoppen der
  Nuklearwirtschaft<br>- JA das Wasser gehoert uns allen (2 mal)<br>-
  JA gegen die Prozess-Abwesenheit von Regierungsmitgliedern
  (legittimo impedimento)<br><br><br>Per far si che il Referendum
  passi bisogna superare il Quorum! Quindi è necessario che votino
  almeno 25 milioni di persone !<br>...Il 12 e il 13 Giugno bastano 4
  SI per la DEMOCRAZIA !<br><br>Le Domande :<br>- Vuoi annullare la
  privatizzazione sull'acqua ? (2 QUESITI - 2 SI ! )<br>- Vuoi
  eliminare il ricorso all'energia nucleare? SI !<br>- Vuoi
  sbarazzarti del legittimo impedimento ? (Vuoi che la legge sia
  uguale per TUTTI ? ) SI !
</div>
...
</body>
```

Another example, containing also some address information:

```
<body>
...
<span class="... fn ...">ASV Vetzan - Schnals</span>
...
<div>
  Freitag, 20. Mai .
  <span class="dtstart">
    <span class="value-title" title="2011-05-20T20:30:00"> </span>
    20:30
  </span> -
  <span class="dtend">
    <span class="value-title" title="2011-05-20T23:30:00"> </span>
    23:30
  </span>
</div>
...
<div class="location vcard">
  <span class="fn org">Sportplatz Vetzan</span>
  <div class="adr"></div>
</div>
```

```
...  
<div class="description summary">  
  Letztes Heimspiel der Saison  
</div>  
...  
</body>
```



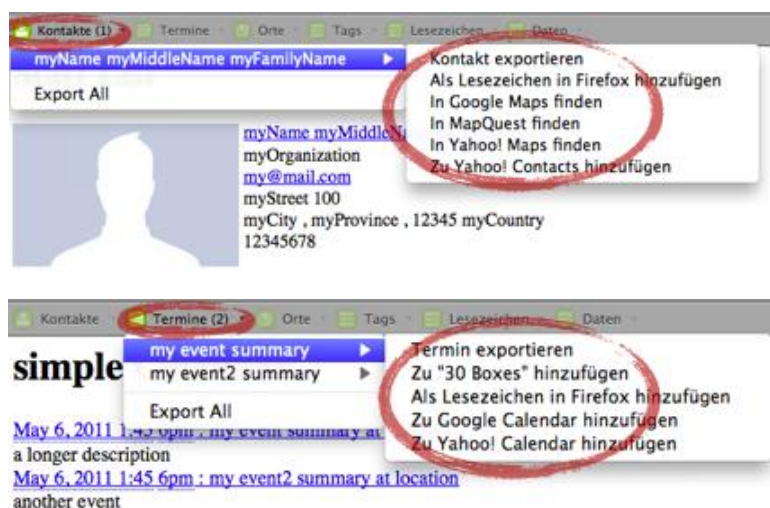


## 2.4 NORMAL BROWSERS

In addition to the benefits when indexing webpages containing microformat enriched data, it also provides benefits for regular users browsing those websites.

For all major web browsers (Internet Explorer, Firefox, Chrome, Safari) plugins exist to support the use of microformatted sites (like Operator...), while Firefox 3 even comes with built-in support for microformats.

Those plugins or built-in-functionalities allow to list all microformatted data and give quick access to commonly used functionality on those data: exporting contact data to the address book, showing address information on maps, adding events to a calendar, ...



## 3 HOW TO MICROFORMAT YOUR PAGES

In the following, a brief introduction into the microformats used for events called hCalendar, or vevent, is given by showing some examples and by explaining how these formats could be used to structure events on HTML pages.

### 3.1 SIMPLE TAGGING (VEVENT)

A simple event could be described in the following way in HTML code:

```
<div>
  <a href="http://www.google.com">
    Important meeting in office 3.60 on
    May 6, 2011, 1:45 pm until 6 pm:
    Discussion of current problems...
  </a>
</div>
```

The same event in standard hCalendar format looks like this:

```
<div class="vevent">
  <a href="http://www.google.com" class="url">
    <span class="summary">Important meeting</span> in
    <span class="location">office 3.60</span> on
    <span class="dtstart">
      <span class="value-title" title="2011-05-06T13:45:00" />
      May 6, 2011, 1:45 pm
    </span> until
    <span class="dtend">
      <span class="value-title" title="2011-05-06T18:00:00" />
      6 pm
    </span>:
  </a>
  <div class="description">Discussion of current problems...</div>
</div>
```

### 3.2 MORE ADVANCED APPROACH (VEVENT & VCARD)

If the location is known in more detail, it should be pointed out explicitly in the microformat, for the same event, if the location would not only be „office 3.60“, but e.g.:

```
<div>
  <a href="http://www.google.com">
    Important meeting on
    May 6, 2011, 1:45 pm until 6 pm:
    Discussion of current problems...
    The meeting will be in SomeBuilding, office 3.60
    Company Street 1a
    1234 Company City, REGION
  </a>
</div>
```

Then the hCalendar event could be specified in more detail:

```
<div class="vevent">
  <a href="http://www.google.com" class="url">
    <span class="summary">Important meeting</span> on
    <span class="dtstart">
      <span class="value-title" title="2011-05-06T13:45:00" />
      May 6, 2011, 1:45 pm
    </span> until
    <span class="dtend">
      <span class="value-title" title="2011-05-06T18:00:00" />

```

```

        6 pm
    </span>:
</a>
<div class="description">Discussion of current problems...</div>
The meeting will be in
<span class="location vcard">
  <span class="fn org">SomeBuilding, office 3.60</span>
  <span class="adr">
    <span class="street-address">Company Street 1a</span>
    <span class="postal-code">1234</span>
    <span class="locality">Company City</span>,
    <span class="region">REGION</span>
  </span>
</span>
</div>

```

Another way (if the full name of the organization should not be mentioned in the location tag, but only the address) would be to not specify the spans of the vcard and of the organization name, and to use a span called „location adr“ as outer tag of the location.

If additional information about the address or the organization describing the location is given, this tag can also be extended further. As a simple example, consider the following HTML snippet containing such information:

```

<div>
  LOCATION:
  
  <strong>SomeBuilding, office 3.60</strong>
  Main Building of AVeryImportant Company
  Company Street 1a,
  1234 Company City, REGION
  THIS COUNTRY
  Email: info@some-company.org
  Phone: 1234 56789
  On a map: lat: 40.0000000; long: -70.0000000
</div>

```

Extending the HTML code with proper microformats gives the following:

```

<div class="vcard">
  LOCATION:
  
  <strong class="fn">SomeBuilding, office 3.60</strong>
  <span class="title">Main Building</span> of
  <span class="org">AVery Important Company</span>
  <span class="adr">
    <span class="street-address">Company Street 1a</span>,
    <span class="postal-code">1234</span>
    <span class="locality">Company City</span>,
    <span class="region">REGION</span>
    <span class="country-name">THIS COUNTRY</span>
  </span>
  Email:
  <span class="email"> info@some-company.org </span>
  Phone:
  <span class="tel">1234 56789</span>
  On a map:
  <span class="geo">
    lat:
    <span class="latitude"> 40.0000000 </span>;
    long:

```

```
<span class="longitude">-70.000000 </span>  
</span>  
</div>
```

### **3.3 USING AN ABBREVIATION (INSTEAD OF DIV OR SPAN)**

If e.g. the geo-information (longitude and latitude), which is of interest for the microformats structure, should not be shown to the user, the geo tag could be exchanged by the following abbreviation (abbr) tag:

```
<abbr class="geo" title="37.386013;-122.082932"> </abbr>
```

The same is true for any other tags which shouldn't be shown to the user, but which provide relevant semantic information.

## **4 LIST OF FIGURES**