

# What Do They Do With *My* RDFa?

Presented on

W3C Tag 2010

2010-09-15, Berlin

by Klaus Birkenbihl, Head World Offices, W3C

# RDFa in the wild

You may have heard that many big players on the Web support RDFa

You don't hear so often which vocabularies they support

You don't hear so often how they use RDFa

So the question arises: (how) can I benefit from it? (not being a specialist on SW).

This short talk gives a few examples how RDFa is used these days by Facebook, Google and Yahoo

They all support “data on the Web”. Usually beside RDFa there is Microformats, Microdata, eRDF or data delivered by a Web service supported

# RDFa and **facebook**

(3)

*Facebook* uses RDFa for its *Open Graph Protocol*

A very simple (and flat) vocabulary

Only one subject per Web page

All values (objects in RDF terminology) are literals (strings)

@prefix og: <<http://opengraphprotocol.org/schema/>>

Properties of *OGP*:

og:title, og:type, og:image, og:url,  
og:description, og:site\_name, og:latitude,  
og:longitude, og:street-address, og:locality,  
og:postal-code, og:country-name, og:email,  
og:phone\_number, og:fax\_number, og:upc, og:isbn

# RDFa and facebook

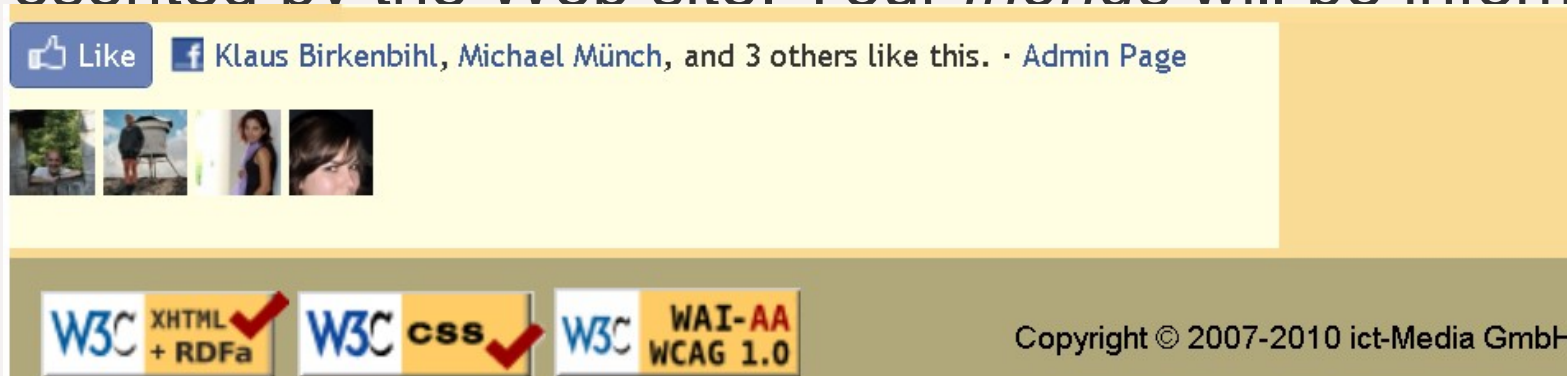
(4)

Allowed values for `og:type`:

activity, sport, bar, company, cafe, hotel, restaurant, cause, sports\_league, sports\_team, band, government, non\_profit, school, university, actor, athlete, author, director, musician, politician, public\_figure, city, country, landmark, state\_province, album, book, drink, food, game, movie, product, song, tv\_show, article, blog, website

Websites within the *Open Graph* will be treated like social objects within Facebook when registered

A *Like* button allows you to indicate that you like the subject represented by the Web site. Your *friends* will be informed



# RDFa and **facebook**

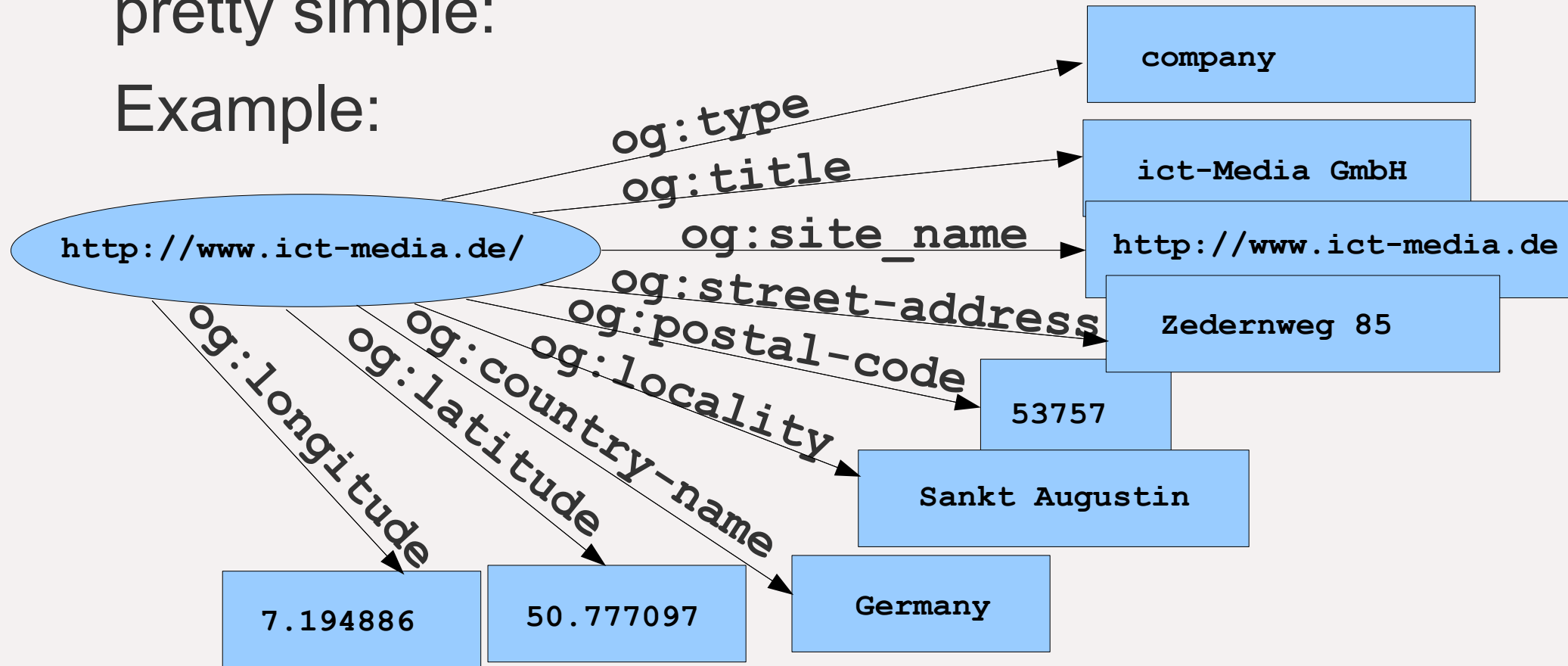
Data provided on a Web site can be used in Facebook applications in the same way as data from Facebook internal social objects is used

The page becomes a node on Facebook's social graph

# RDFa and **facebook**

The RDFa graph for Facebook looks (today) pretty simple:

Example:



# RDFa and Google

Google is introducing what they call *rich snippets*.

[Drooling Dog Bar B Q - Colfax, CA](#)  
★★★★☆ 15 reviews - Price range: \$\$  
Drooling Dog has some really good BBQ. I had the pulled pork sandwich, .... Drooling Dog BBQ is a great place to stop at on your way up the hill to Tahoe ...  
[www.yelp.com/biz/drooling-dog-bar-b-q-colfax](http://www.yelp.com/biz/drooling-dog-bar-b-q-colfax) - 75k - [Cached](#) - [Similar pages](#)

“With rich snippets, webmasters with sites containing structured content—for example, review sites or business listings—can label their content to make it clear that each labeled piece of text represents a certain type of data ...”

Currently only for review sites and social networking/people profile sites. Other types of content in the future

Moving forward with caution

# RDFa and Google

(8)

Formats: microdata,  
microformats, or RDFa.

Started with its own  
vocabulary

[http://rdf.data-  
vocabulary.org/](http://rdf.data-vocabulary.org/)

Meanwhile supports  
other popular  
vocabularies like foaf or  
vCard

Still sort of experimental  
– but with a high  
potential

## Google search preview

The following errors were found during preview generation:

- Insufficient data to generate the preview.

## Extracted Rich Snippet data from the page

Organization

Address

street-address = Zedernweg 85  
locality = 53757 Sankt Augustin  
region = Germany

Geo

latitude = 50.777097  
longitude = 7.194886

url = <http://www.ict-media.de/>

name = ict-Media GmbH

tel = +49 2241 396415

VCard

**Warnings** : RDFa tree has unknown type 'VCard'.

Address

street-address = Zedernweg 85  
locality = Sankt Augustin  
postal-code = 53757  
country-name = Germany

geo

latitude = 50.777097  
longitude = 7.194886

url = <http://www.ict-media.de/>

tel = +49 2241 396415



## SearchMonkey roles:

Developers develop search applications

Site owners provide data (e.g. using RDFa)

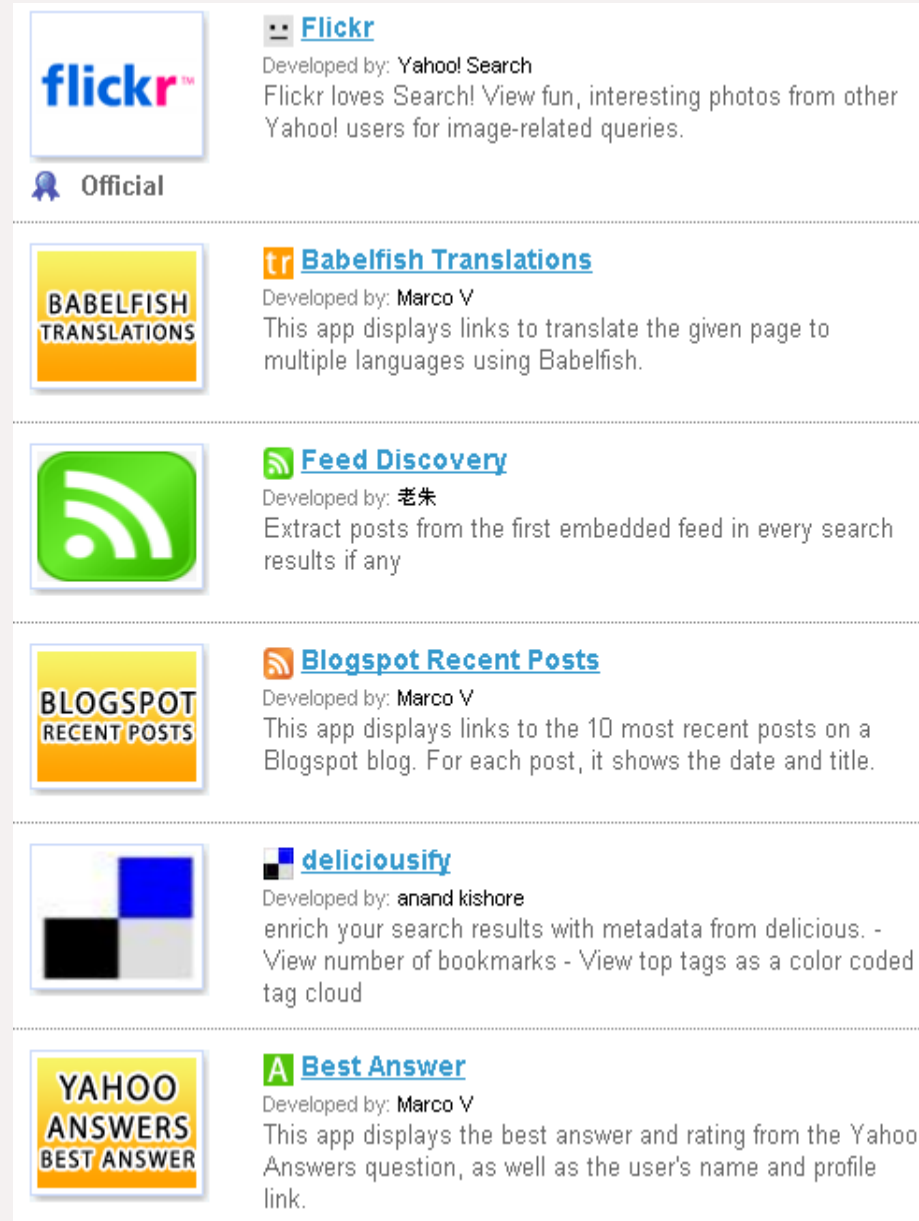
Users can register for certain SearchMonkey applications

All applications are sort of Open Source (Yahoo licence)

The SearchMonkey application gallery holds a lot of examples.

Developers can submit their applications for the Search Gallery

Developers can also offer the use of their application on their Website



The screenshot displays a grid of search applications. Each application card includes an icon, a title, the developer's name, and a brief description of the application's functionality.

- Flickr**: Developed by: Yahoo! Search. Flickr loves Search! View fun, interesting photos from other Yahoo! users for image-related queries. (Official)
- Babelfish Translations**: Developed by: Marco V. This app displays links to translate the given page to multiple languages using Babelfish.
- Feed Discovery**: Developed by: 老朱. Extract posts from the first embedded feed in every search results if any.
- Blogspot Recent Posts**: Developed by: Marco V. This app displays links to the 10 most recent posts on a Blogspot blog. For each post, it shows the date and title.
- deliciousify**: Developed by: anand kishore. enrich your search results with metadata from delicious. - View number of bookmarks - View top tags as a color coded tag cloud.
- Best Answer**: Developed by: Marco V. This app displays the best answer and rating from the Yahoo Answers question, as well as the user's name and profile link.

# YAHOO! SearchMonkey

SearchMonkey applications are pieces of PHP that define how – for a set of URIs – the search result are displayed.

SearchMonkey gives special support for some data “Objects”.

Data can be specified using popular vocabularies like *vcard*, *Dublin Core*, *foaf* and private ones like

<http://search.yahoo.com/searchmonkey/>

SearchMonkey applications go beyond use of data within the page (e.g. you can enhance a result with URLs for machine translations etc.)

Example search result with custom application:

[ict-Media GmbH](#)  
[万维网 \(cn\)](#) | [Das Web \(de\)](#) | [The Web \(en\)](#) | [La Web \(es\)](#)

- Straße: Zedernweg 85
- Ort: 53757 Sankt Augustin
- Land: Germany
- Telefon: +49-2241-296415

 [www.ict-media.de/index.html](http://www.ict-media.de/index.html) - [Cached](#)



SearchMonkey > Start Overview > Object Finder

## Testing Markup



### DATA NOT CACHED YET

The data stored in the Yahoo! Search Index doesn't match the data just retrieved from the live version. This is normal if you just changed your page — simply wait for us to recrawl and reindex your page. The average update time is two weeks for most URL. Here is how your page [currently appears in Yahoo! search results](#).



### MULTIPLE OBJECTS

You have multiple objects on your page. Unfortunately, SearchMonkey currently only supports displaying one object per page, so no enhanced result will display for this URL. You can remove all objects except for the one you wish to display, or wait for SearchMonkey to support displaying multiple objects.



Event

#### RECOMMENDED

<b>vcal:summary</b>	Some Steps from the Web to a Semantic Web
<b>vcal:dtstart</b>	2010 -05-07

#### OPTIONAL

<b>vcal:location</b>	New Delhi, India
----------------------	------------------

#### MESSAGES

<b>Info</b>	Data Not Cached Yet This object data is not cached in the Yahoo! Search Index. This is normal if you just changed your page — simply wait for us to recrawl and reindex your page.
-------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



Event

#### RECOMMENDED

<b>vcal:summary</b>	The "what" and "how" of W3C
---------------------	-----------------------------

Example from <http://www.ict-Media.de/talks.html>

Included from: <http://www.w3.org/Talks/>



Local

#### RECOMMENDED

<b>vcards:tel</b>	+49 2241 396415
<b>vcards:fn</b>	ict-Media GmbH

#### OPTIONAL

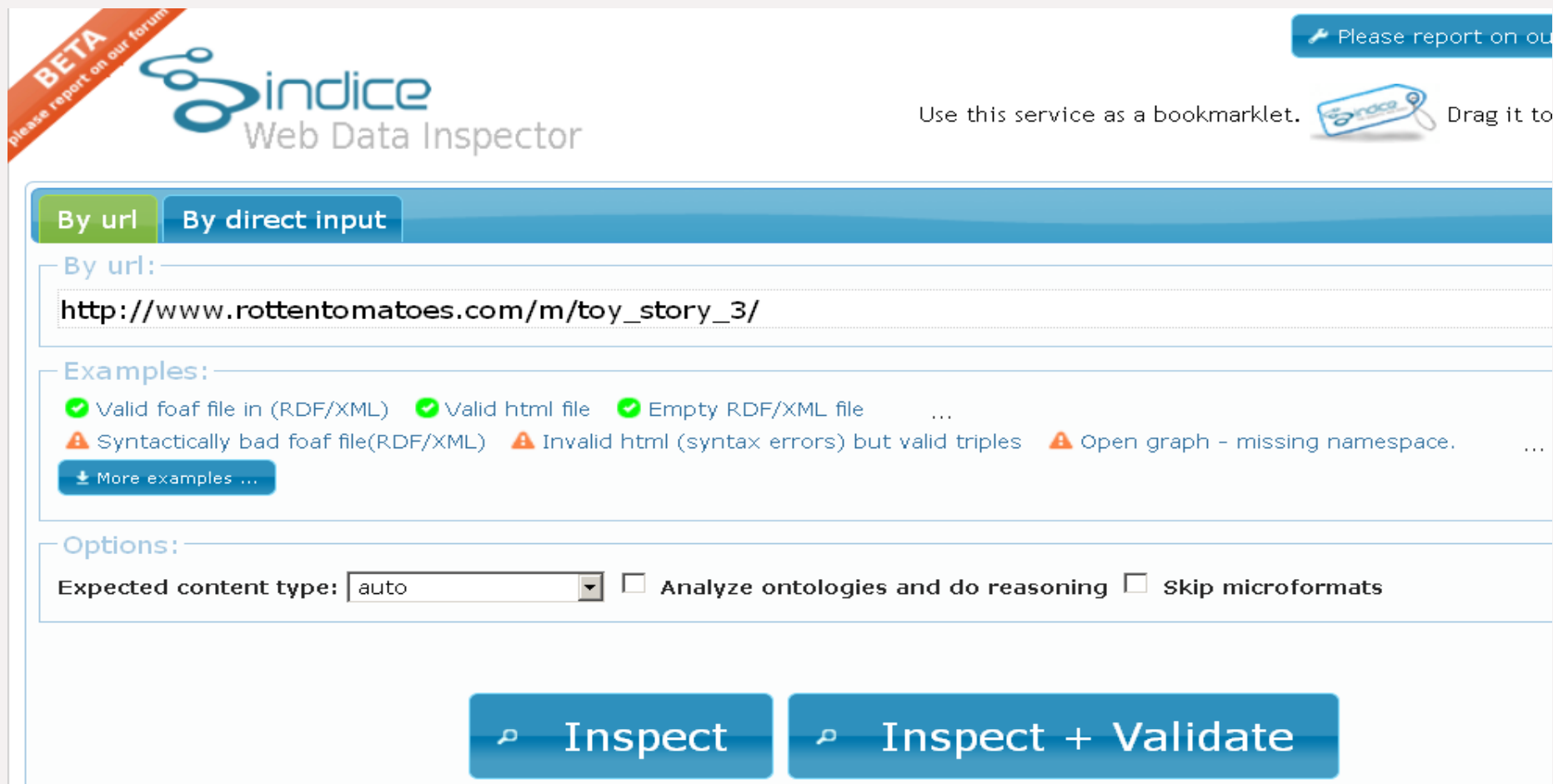
<b>vcards:street-address</b>	Zedernweg 85
<b>vcards:locality</b>	Sankt Augustin
<b>vcards:postal-code</b>	53757
<b>vcards:country-name</b>	Germany

from <http://www.ict-Media.de/>

# A Semantic Search Engine

There are many tools on the Web to play around with Web pages containing RDFa.

A nice one: the Sindice Inspector:



# The original Question

How can I benefit from it?

I have to provide my RDFa many times:

Once for each of their vocabularies (which actually many pages do!)

Some smart OWL code could probably help to bridge the gap between the most rather simple vocabularies but

Today they will not interpret OWL

# Conclusion

We are rather at the beginning

In terms of how RDFa is used by the big players

In terms of who can benefit from it

In terms of the effort to provide RDFa for all of them ...

# Nevertheless ...

Not only data users (like search engines) but also information providers use more and more RDFa on their Web pages

Famous RDFa providers: NewsWeek, TESCO, O'Reilly Catalog, Best Buy, Public Library of Science ...

There are many tools around for developers. (RDFa checker and RDFa parser)

Many Wikis, Blogs or CMSs can easily be configured to support RDFa



Slides are available at:  
<http://www.w3.org/2010/Talks/0915Berlin-KB/>  
in OpenDocument Presentation Format and PDF  
in English