**#wedi**

Scientific coordinator: Students:

Dr. Buraga Sabin Corneliu Butnariu V. Vasile-Raducu

Deaconu C. Mihai

Sfirnaciuc Daniel-Ionut

**Documentation**

**#introduction**

Every day new data is added on the web. Along with this case a big problem occurs, the data is not well structured. In order to fix this one approach that came up was to divide data into different categories like government data, health related data, general knowledge, company information, flight information, places and many other. Doing this we have to admit that this data is interpretable only by humans. This data should be structured in a way that not only humans can read and interpret this data but also machines would also be capable to understand and perform more operations on the interpreted data.

The Semantic Web technologies help us to make the content of the web pages more interlinked and always available on the web. It is a system the enables to interpret and respond to complex human requests based on their meaning. This means that the relevant data sources should be semantically structured.

Some solutions to have data structured are Web Ontology Language (OWL), Extensible Markup Language (XML) and Resource Description Framework (RDF).

We noticed that many web pages started to include structured data into their HTML using markup standards such as **RDFa**, **microdata** and **microformats**.

Being aware with the fast changes on web data content for an user the need for a tool that provides a more human understandable representation of RDFa, microdata and microformats is imminent. With such a tool, the user can have in one place all the structured data.

**#similarapplications**

*hNews Reader* – is a Google Chrome extension for highlighting pages conforming to the hNews microformats standard. It displays the various hNews fields, required and optional, in a small popup window (Fig. 1).

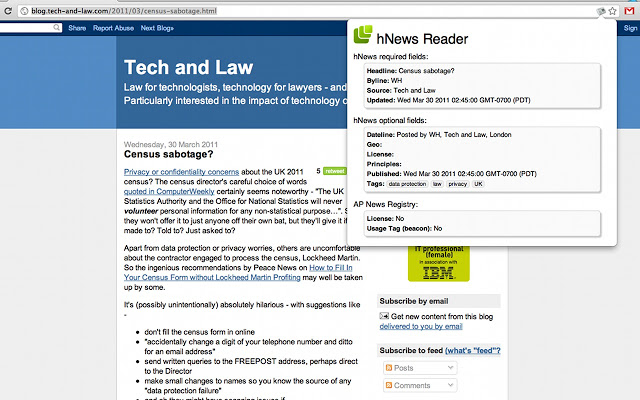


Fig. 1

*Microformats for Google Chrome* - an extension that supports hCard, hCalendar, hReview. With this extension you can export cards to \*.vcf files to add to Address Book or to your Google Contacts (Fig 2).

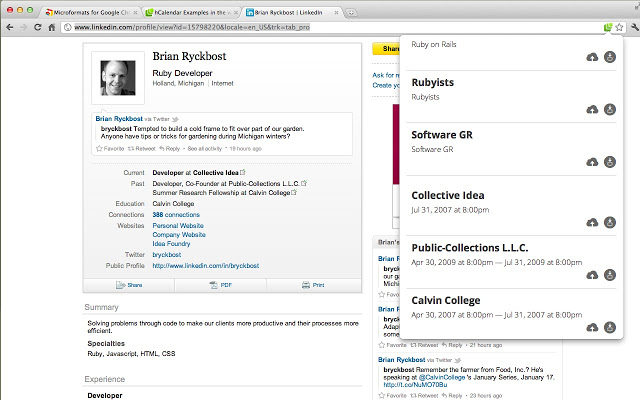


Fig. 2

*Tails Export* – is a Firefox extension used for showing and exporting Microformats. Currently it supports the following formats : hCard, hCalendar, xFolk and Rel-license (Fig.3).

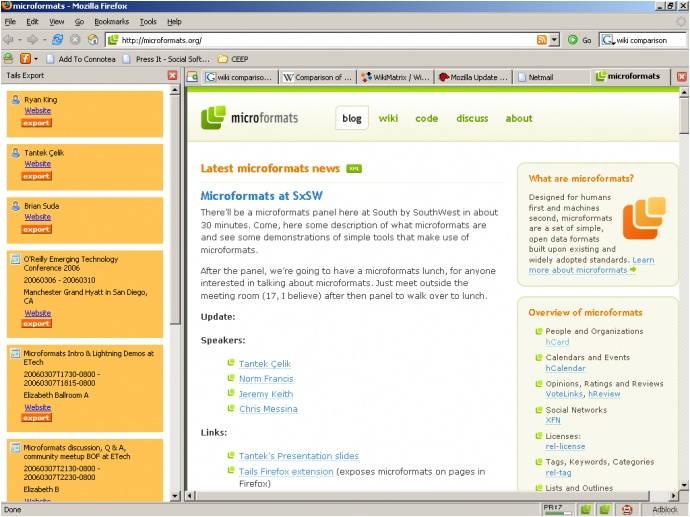


Fig. 3

*Operator*  - is a Firefox extension that lets you to combine pieces of the information on Web Sites with applications in ways that are useful such as Flickr with Google Maps, Yahoo! Local with your address book (Fig. 4).

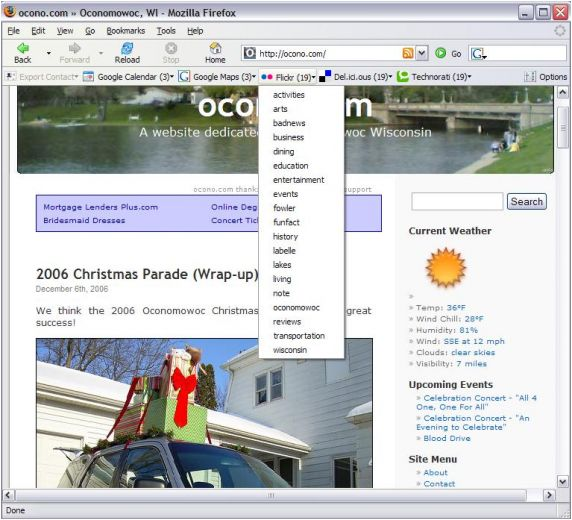


Fig. 5

**#systemarchitecture**

Wedi is a cross browser extension that allows users to read metadata from web pages. It consists of following components: WediOnAccess, WediParser, WediConverter, WediNotification, WediExporter and WediVisualiser.

The WediOnAccess component is invoked when the user access an URL. Also this component attaches global event listeners for interacting with the rest components and buttons from browser.

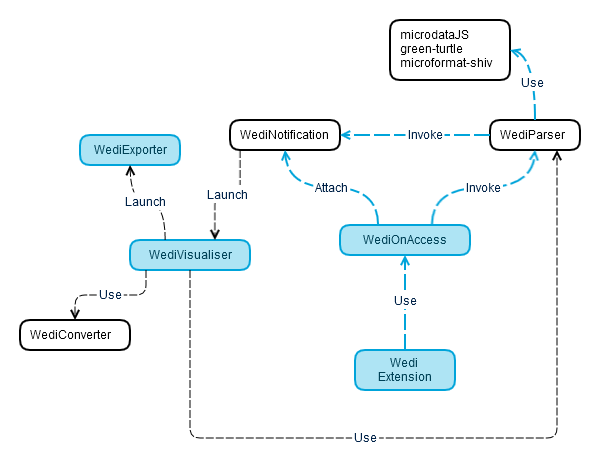
The WediNotification component is the component which notifies the user that into the current page the WediParser have found metadata.

The WediParser component is a wrapper for parsing libraries used in this project. This component ‘finds’ microformats, microdata and RDF metadata in current page content and returns an json object containing data extracted from page.

The WediVisualiser component will be the wedipage\_ui.html file that will show the visual representation of the microdata, microformats and RDFa formats. It will be a popup window and is invoked by user with a simple click on the notification icon.

The WediConverter component will convert the data.

The WediExporter component is responsible for exporting the content displayed in the WediVisualiser in XML format.



**#librariesandframeworks**

We use:

* [jQuery](http://jquery.com/) – javascript library.
* [microdataJS](https://github.com/foolip/microdatajs) – a jQuery plugin that provides and API for accessing Microdata inspired by the native DOM API.
* [green-turtle](https://code.google.com/p/green-turtle/) – an implementation of RDFa 1.1 for browsers.
* [microformat-shiv](http://microformatshiv.com/) – a light-weight cross browser javascript microformats parser

To pack the extension for different browsers we will use :

* [Kango framework](http://kangoextensions.com/kango.html)