Semantic processing based on eye-tracking metrics

ROBERT ANDREI BUCHMANN
Business Information Systems Dpt, Faculty of Economic Sciences and Busines Administration
Babe-Bolyai University
Str. Teodor Mihali, Nr. 58-60, Cluj Napoca, 400591
ROMANIA
robert.buchmann@econ.ubbcluj.ro

ALIN MIHAILA
Business Information Systems Dpt, Faculty of Economic Sciences and Busines Administration
Babe-Bolyai University
Str. Teodor Mihali, Nr. 58-60, Cluj Napoca, 400591
ROMANIA
alin.mihaila@econ.ubbcluj.ro

RADU MEZA
Journalism Dpt., Faculty of Political, Administrative and Communication Sciences
Babe-Bolyai University
str. G-ral Traian Moşoiu nr. 71, Cluj-Napoca, 400132
ROMANIA
mezaradu@yahoo.com

Abstract: - This paper proposes a framework for capturing semantics from eye tracking data during the process of text skimming/scanning by readers of electronic documents and HTML user interfaces. RDFa and HTML microformats are some of the easier ways proposed by the Semantic Web paradigm for embedding semantics in web pages. XSLT transformations or specialized parsers may easily convert such documents to RDF/XML semantic repositories. However, semantics do not usually have an absolute character. Although a variety of web 2.0 oriented ontologies and microformats have been widely adopted or even standardized (Dublin Core, FOAF, XFN etc.), in order to achieve semantic interoperability, there are scenarios in which user-relative semantics are especially important, such as in the development of customization engines (web session customization, recommender systems, targeted advertising), when a certain user must only share semantics with himself or with similar persons. Having the same web document, different readers would attach various semantics and relevance to the ideas, concepts or structural blocks of the document. Eye tracking is an emerging field with multiple applications in medicine, marketing, cognitive sciences and others, which allows the extraction of data regarding the eye activity of a user during human-computer interaction. Eye tracking data is valuable in measuring reading patterns, user experience and reflects the specific parts of an image or documents that attract the user's interest.

Key-Words: - eye tracking, semantics, text skimming, microformats

1 Introduction

For the purpose of this paper, we designate the term **text skimming** as the *efficiency-oriented* process of reading text from display devices, without passing each symbol of the text, but rather scanning its surface in search of visual cues such as keywords, headings, notes or just general text structure in order to identify its sections and relate them to the logic flow. The process of skimming is intended to create a superficial idea about the text and its purpose and sometimes it is enough for the reader who only wants to receive the general semantics represented by the text and not the text itself (journalism, aggregated data, product descriptions etc.). A particular type of skimming is **text scanning**, a more regular gaze movement pattern in search of a symbol or combination of symbols.

On the other hand, we have *effectiveness-oriented* reading, which is the proper reading process, covering every symbol of the text. Effective reading emphasizes text formalism and syntax and it's important in fields where text construction is more relevant than (or as relevant as) the semantics (literature, law etc.). In the context of Web 2.0 and electronic documents, text skimming tends to replace reading, as readers spend most of their time looking for relevant visual items, in order to find their "conceptual way" through the flow of information provided by a website [10].

Hypertext and the evolution of e-documents changed the way that humans are reading. Information accuracy and formal construction of the text may be sacrificed for faster semantic coverage.

Following Jakob Nielsen's recommendations for increased readability [8], web information should be presented in such formats and structures that decrease reading effort and still generate the knowledge represented by the text, in the reader's mind. Hypertext raised this awareness by introducing hyperlinks, which are practically a way of automating cross-references and, to a certain extent, creates a semantic network based on references.

Although HTML and XHTML have weak support for semantics, recent efforts from various communities (including W3C) promoted several solutions that meet both the requirements of Semantic Web and the simplicity of Web 1.0 tagging paradigm. Some of these, successfully employed by our research efforts are microformats[22], RDFa (RDF in attributes)[28] and, for conversion purposes, GRDDL (Gleaning Resource Descriptions from Dialects of Languages)[29].

Eye tracking studies are used to record the path taken by the users' gaze during the text skimming process (or other types of human-computer interaction).

The field of **eye tracking** originated in optical and neurological studies, providing tools that can be applied in a variety of interdisciplinary studies, in various fields such as:

- Medical and assistive technologies (regarding ocular mobility, infant perception, gaze-controlled user interfaces for people with disabilities etc.);
- Marketing (market research, visual design effectiveness) [16][17];
- Cognitive ergonomics and usability (as part of software quality) [1];
- Visual communication [14][15].

Mouse emulation through eye tracking permits the mouse cursor to be controlled by eye activity. This raises some technical challenges approached by many researchers in the recent years. The so-called *Midas touch* [4] means that a user cannot express intent of action through his eyes, but he has the potential to activate everything encountered in his gaze path. Some of the solutions employed by eye tracking system developers are:

- Dwell-time over a certain element, within certain time limits, longer than a fixation[4];
- Probabilistic selection of objects based on the eye movement pattern [12];
- Combining hand and eye movement, by using the keyboard for issuing any commands [18].

Our paper describes a framework for gathering relations between document elements, based on eye tracking generated metadata, and some side results observed during our eye tracking studies. The next section presents the tools, methodology and problem formulation, while the third one offers details regarding the proposed framework.

2 Instrumentation and Problem Formulation

2.1 Eye tracking tools

Most popular eye tracking suites record eye movement through videobased infrared oculography. Most of the commercial solutions provide both the hardware and the software necessary to store and analyze the data. Tobii technologies integrate the eye tracking receiver and server with a TFT monitor, eliminating any interference with the user experience [27]. Previous eye trackers were usually head-mounted or camera-focused, which limited head movement and natural user behavior. There are also some open source projects providing cheaper, less refined eye tracking solutions. Some of these are software-only, consisting in libraries for mapping images and video captured by webcams to eye tracking data structures. All of the tools used in the current state of our project were open source low cost or freeware, since the efforts are currently aimed at framework, metrics and methodology development, but acquisition of integrated commercial solutions will follow for wider studies with large sets of subjects. Some of the gaze calculation algorithms and libraries involved in our studies are:

- Starburst/openEyes [21][23]
- openGazer[26]
- TrackEye[24];
- ITU Gaze Tracker[25];
- OGAMA (analysis tool) [20].

Even if eye tracking solutions lack the standardization, commercial availability and ease of integration necessary to raise them at *domestic technology status*, we posit that the future will have eye tracking devices and drivers integrated with regular displays (already most laptops include a web cam). Of course, this will raise new debates regarding the balance *privacy versus customization*. TFT monitors with integrated eye tracking guarantee no interference between the system and the natural behavior of the user, while most of the current low-cost solutions are head mounted or based on a configuration of cameras/web cams which limit environment variables such as head movement or lighting conditions.

2.2 Methodology

Our current methodology involves recording both eye tracking and speech (narrative usage), in order to assign meaning to eye movements. A process of pre-selection and training of the subjects was necessary in order to avoid "attention deficit"-type behavior, as we expect such behavior to randomize data obtained from text skimming. Text skimming is sometimes imposed by putting time pressure on the subjects.

The subjects are involved in various strategies of text skimming, according to the purpose of reading:

S1: Entertainment; it is closer to effectives reading, targets both semantics and text formalism;

S2: Information acquisition; it is closer to scanning, may be considered a search technique, relies heavily on memory and prediction anchors: the reader stores mental cues based on the image of a key word and tries to match that image with the patterns in the text during a straight eye movement – usually horizontal, vertical or diagonal;

S3: Task processing (reviewing, filling a survey, a multiple choice test etc); balances the two strategies, relies heavily on document structures.

The subjects are 25 students recruited from our university, with an average age of 19 years. The gender breakdown is 10 males and 15 females. Two subgroups were defined, depending on their "digital literacy":

- Those who perform on-screen reading on a regular basis (daily): 15 subjects;
- Those who perform on-screen reading rarely: 10 subjects.

2.3 Guidelines for text-heavy documents

Two styles of text-heavy documents are involved in experiments:

- Traditional pyramid style of exposition – typical for research papers and reports: starting with introduction, review of similar ideas/documents, and closing in on the results and conclusions. This manifests as a general-to-specific argument, trying to draw facts/results from supposition/hypothesis;
- Inverted pyramid style of exposition – typical for journalism, blogging and entertainment-oriented documents, involving early hooks: conclusions briefed in the beginning, then detailed through semantically chained paragraphs. This manifests as a specific-to-general argument, trying to draw supposition from fact.

Documents and recordings are separated by scrolling (Page Down/Up) and hyperlink activation so that each different document image had its own attention map.

Tested documents follow certain guidelines suggested by Jakob Nielsen's [8] works for writing web documents:

- Concision, lack of verbosity. Nielsen even goes to stating that web content should be 50% shorter than having the same message published in hardcopy. This was related to low readability due to monitor resolutions compared to print resolution during the 90s but even with modern screens (in the context of increased mobility and miniaturization) it's important to have a clear idea if the text consumer will be focused on the formal aspects, on the rhetoric or on the message itself and, in the second case, to minimize the effort of acquiring the message semantics;
- Orientation towards scannability. Text scanning is the most focused type of skimming and it manifests during document searching. It usually involves a straight succession of eye movements (vertical or horizontal, more rarely on diagonal) over a succession of similarly formatted visual cues (headlines, list items, paragraph indentations, bolded words). Using proper highlighting of key concepts, meaningful headings and a clear structure of lists and paragraphs improves text scannability;
- Emphasizing the nature of hypertext. Hypertext differs from linear text flow in many ways, most these having to do with hyperlinks and how they implement references within document semantics. It's recommended that the paragraphs of a document should be clearly delimited and each focused on a single idea/concept, while containing hyperlinks towards documents detailing those concepts. The document formal structure should reflect a semantic structure;
- Increasing full-text readability. If the reader chooses to fully read certain parts of a document, his effort should be minimized by using optimal typesetting and color schemes: this means that font size should not be smaller than 10 and typesetting that lose typical shape when bolded or italicized are to be avoided. Text density should be lowered. This could be affected by antialiased characters, font ornation

(even serifs) or narrow fonts. Verdana is considered one of the most favored font by on-screen readability specialists, while Georgia is the most readable serif font. All caps text reduces reading speed. The color scheme should not use more than 4 colors with proper contrast, dark colors being recommended for text and desaturated colors for background. Numbers should be written as numerals;

- Increasing relevance of microcontent. Microcontent consists in small piece of text with high representativeness (headlines, subject lines, product descriptions etc.). Microcontent is often displayed out of context, as a list of hyperlinks, e-mails, bookmarks or other navigational elements. Even when displaying microcontent with its surrounding context, window area limits and the default on-line reader behavior makes it harder to perceive the surrounding context. Thus, microcontent should act as an abstract for larger content;
- Increasing relevance of nanocontent. Nanocontent consists in the first few characters read at each eye fixation. During text scanning, these might be the only characters that are read at all. Relevant nanocontent acts as a predictive cue for the rest of the message, possibly engaging an antisaccade (moving the eye away from a certain paragraph considered irrelevant). It is best to avoid wasting nanocontent on leading articles ("the","an","a"), generic formalities or figures of speech. The first two words of each headline, list item, hyperlink etc. should have enough semantic power and informative load to attract reader attention.

2.4 Semantic Tools

Microformats are the most simple way of annotating documents with metadata using strictly HTML tagging, thus providing the most concise and integrated semantic solution. Microformats usually exploit attributes such as CLASS and REL for defining semantics as an extension to their primary use (CSS formatting and hyperlink description). A limited set of names for classes and relations, destined for a certain context (geography, chemistry etc), forms a domain-specific microformat.

RDFa is a recently proposed W3C recommendation for embedding metadata using the structure of semantic triples directly in HTML/XHTML documents.

Traditional HTML provides a very limited range of semantic tags, such as META and LINK. Both of these are document-level (nested in the HEAD section) and provide metadata significant to browsers and search engines.

Later, RDF provided a framework for structuring semantics as triples, with the following components:

- Subject (a resource identified with URI, usually a document or a concept uniquely identifiable through a URI)
- Predicate (relationship, property or attribute of the subject, uniquely identified by an URI or, more conveniently, a compact URI prefixed by a namespace);
- Object (a resource related to the subject or the literal value of a property/attribute).

RDFa enriches HTML with the possibility of creating RDF triple structures for any part of a document, using certain attributes (inspired by META and LINK tags):

- ABOUT: specifies the subject resource, other than the current document or parts of it;
- REL, REV, PROPERTY: specify the predicate of triple, depending on its type (a relationship to another resource, a reverse relationship or a property with a literal value given by document content;
- HREF, SRC, RESOURCE: specify the object of a triple;
- CONTENT: specifies the object, when it overrides the tagged document content;
- DATATYPE, TYPEOF: specify the type of the literal object and the resource subject.

Example (with properties provided by the Dublin Core metadata set for description of documents):

```
<div about="paper.html">
        The paper
      <b property="dc:title">
      Eye tracking Semantics
      </b>
        is authored by
      <b property="dc:creator">
      RA Buchmann
      </b>
        and its abstract is available
      <a rel="abstract" href="abstract.html">
      here
      </a>.
</div>
```

2.5 Problem Formulation

We think that website adaptability and customization can be significantly improved by recording a conceptual semantic map drawn from the "attention map" created through human-computer interaction. Ideally this would apply to all elements presented on a user interface but the current scope of our studies is limited to the creation of a semantic map from an XML-tagged document. The tagging technologies involved were Topic Maps and RDF, with a preference for the latter, as RDF graphs are easier to build dynamically and transformations between microformats, RDFa and RDF are easily accomplished with languages such as XSLT or GRDDL (currently available for PHP through the PEAR package [19]). On another hand, topic maps seem more natural for large text documents covering recurrently certain topics and associations. [3]

Socially-oriented microformats (XFN) and ontologies (FOAF) inspired our proposed microformat (and its RDFa counterpart), aiming to describe parts of a document based on the user's attention map.

Key concepts and relations identified in the user's eye tracking data may be used by the customization engine of the web site (e.g.: a recommender system).

3 Problem Solution

3.1 The Proposed Framework

Eye tracking suites usually provide the hardware for recording the following states of ocular activity:

- Saccades– quick, simultaneous movements of both eyes in the same direction, or "ballistic jumps of the eyes serving to project specific locations of the scene onto the fovea"[17]. The saccades are the fastest movements in the human body and usually they temporarily suppress vision;
- Fixations-projections of the scene on the fovea, for visual processing, lasting 200-500 milliseconds [9];
- Pupillary diameter variation has been considered a relevant emotional response in literature. Recent studies [6] denied any correlation between the direction of the variation (dilation or contraction) and the valence of emotional response (positive or negative). However, it's been established that pupillary responses reflect variations in the central processing load [2];
- Blinking-sometimes associated with the saccades, usually detected and eliminated from the data.

Another state, the refreshing saccades (or microsaccades) provide retina stimulation for image refreshing and are linked to the constant state of vibrancy of the human eye. These are usually not recorded by eye trackers.

The combination of these states create the illusion of smooth eye movement.

The saccades are usually guided by anchors, which we classified with respect to the types of saccades indicated by [11]:

- Visual anchors, cues for visually guided saccades;
- Visual counter-anchors, cues for antisaccades (jumping away from a cue, in order to avoid visual contact);
- Memory anchors, cues remembered and expected to be found by the reader;
- Prediction anchors, cues expected to appear.

Reading a document from a screen may involve all types of anchors:

- Visual anchors are essential and take many forms:
    - Keywords emphasized through formatting;
    - Bulleted ideas;
    - Spaced paragraphs;
    - Headings
- Visual counter-anchors usually indicate text sections avoided by the reader; either the text is considered irrelevant by the reader, or it is avoided due to its straining effect on the eye (bad formatting and punctuation, insufficient paragraph spacing);
- Memory anchors are created while reading bulleted or numbered lists or during re-reading;
- Prediction anchors are especially involved in text skimming, when the reader expects to find a text section regarding a certain keyword.

Actually, text skimming is based on a combination between antisaccades and prediction saccades. The reader avoids certain parts of the text after he identifies them as being "irrelevant", usually based on nanocontent (he knows that a certain phrase or paragraph covers a certain concept or idea that he's not interested into or he thinks he already knows what it's about).

Most eye tracking studies aggregate data and represent it in an intuitive manner, using the so called heat maps (attention maps) and gaze trajectories (scan paths). These can be represented as graphs with valued (weighted) nodes and mapped on the surface that provides the visual stimulus (an image, a user interface, a document etc.). Semantic technologies such as RDF and Topic Maps are fit for the representation of semantic graphs.

The steps of the semantic mapping process are, as shown in fig. 1:

- The user skims over the document; his gaze is recorded by an eye tracker with mouse emulation;
- The mouse emulation permits event handling (MouseOver translates to EyesOver);
- Each fixation longer than a minimum threshold triggers the event and an asynchronous transfer of data towards the server;

- The server converts the eye activity, mapped on documents elements, to RDFa or microformat semantics, which are stored in the user profile repository;
- Subsequent sessions of the same user will request pages with dynamically embedded semantics according to his profile, in order to improve certain features such as searching or recommendations.

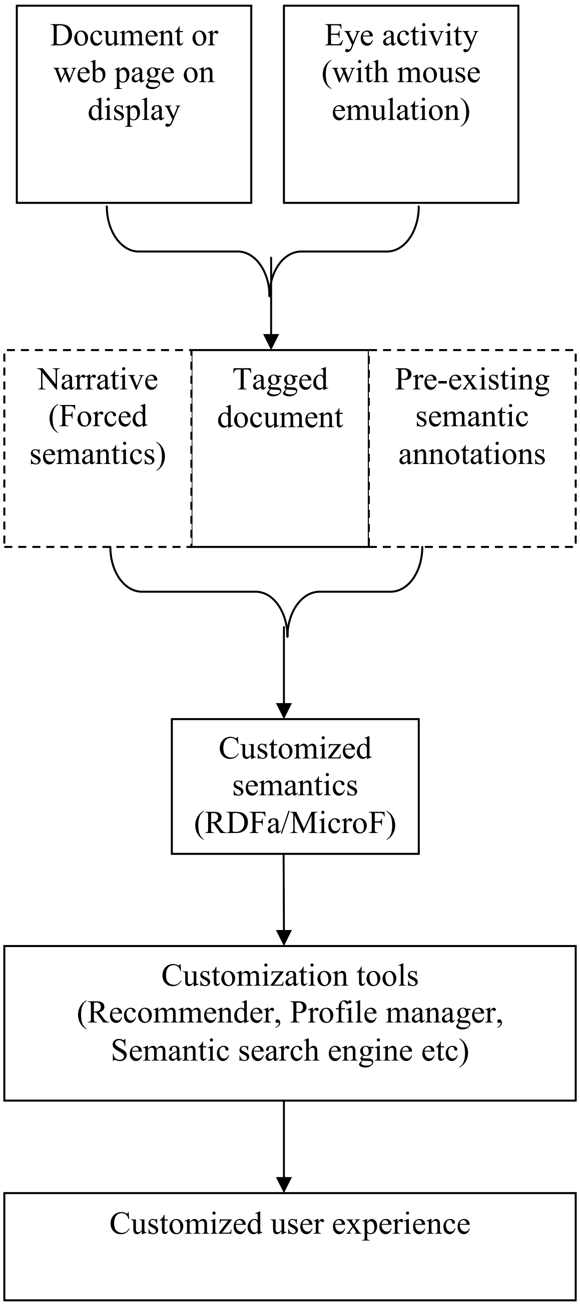

Figure 1. The steps of eye-tracking semantic mapping

The semantics may be defined using two methods:

a. Post-session semantics - extracted from the recorded narrative of the user experience; this wouldn't be feasible in a real-world fully automated application (although a form could be generated in order to obtain semantic tagging for the fixated elements);

b. Pre-session semantics - extracted from a superset of semantics initially annotated to the entire document. For example, general-to-specific relationships can be predefined based on the style of exposition (traditional or inverted pyramid).

The post-session semantics are much more complex, including subconscious associations between parts of the text: a document in which both the name of the writer Alice Sheldon and the location of Sunderland were fixated generated an association triggered by Lewis Carroll's *Alice in Wonderland*.

Eye tracking data extracted from the process of reading/skimming a web page might generate a semantic network customized for each reader's mindframe. The gaze trajectory or heat map generated during the process of reading/skimming is a graph which also reflects the way in which the reader's attention drifts between concepts reflected in the text. This data can be used for greater customization of user profiles/accounts, by tagging text parts with data reflecting the gaze activity of each user.

The keywords (subjects) in a text can be tagged with information such as:

- The gaze intensity (fixation, pupillary diameter);
- The connected concepts (the subject/object/hyperlink that made the reader to move its gaze over the tagged word, and/or away from the tagged word). This implements the so called *rhetoric of departure* and *rhetoric of arrival* (semantics regarding the two-way association determined by the existence of a hyperlink)[5][7].

The following figure and its attached code sample shows how a sequence of two attention maps would be represented in a simplified version of the proposed microformat:

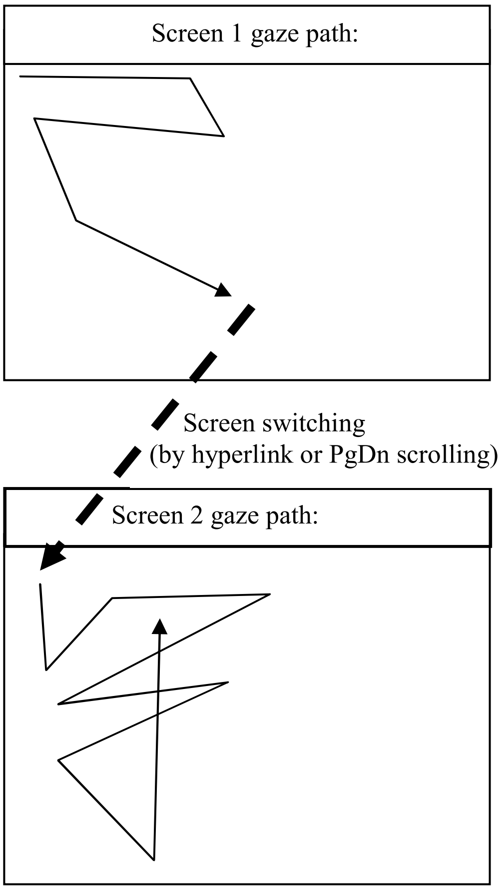

Figure 2. Two gaze paths connected by a screen-switching operation (only hyperlink or PageDown allowed)

The gaze paths would translate in the microformat reflected in the following code sample:

```
<div class="attention">
        <div class="user">Joe Smith</div>
        <a href="#scr1fix2" class="saccade">
        <div id="scr1fix1" class="fixated">
                <span class="description">
                 ...............
                </span>
                <span class="pupildiam">
                ................
                </span>
        </div>
        </a>
        <a href="#scr1fix3" class="saccade">
        <div id="scr1fix2" class="fixated">
                <span class="description">
                 ..............
                </span>
                <span class="pupildiam" >
                ......
                </span>
        </div>
        ...............................
        <a href="#scr2fix1"  rel="nextscr"
                        rev="prevscr">
        <div id="scr1fix6" class="fixated">
                <span class="description">
                ...............
                </span>
                <span class="rhet-arrival">
                .............
                </span>
                <span class="rhet-depart">
                .............
                </span>
                <span class="pupildiam">
        .       ......
                </span>
        </div>
        </a>

        <a href="#scr2fix2" rel="saccade">
        <div id="scr2fix1" class="fixated">
                <span class="description">
                ...............
                </span>

                <span class="pupildiam">
        .       ......
                </span>
        </div>
        </a>
.........
</div>
```

The microformat document would be created by the server based on eye tracking data received from the client and stored in a semantic repository for future use (related to the identity of the user).

In order to simplify eye tracking measurements, document scrolling is not allowed. However, using the PageDown key or activating hyperlinks are events that trigger the "screen switching", which resets the gaze path and defines a way of chain-linking gaze paths of consecutive screens.

The screen switching operations generate the departure and arrival rhetoric ("rhet-depart" and "rhet-arriva" metadata classes), while the REL and REV attributes establish the screen order. The description elements may include the nanocontent/microcontent that triggered each fixation.

The example shows an HTML-focused microformat intended for presentation processing, thus the saccades are defined as hyperlinks. A higher level of abstraction, more distanced from the presentation layer, is accomplished if the microformat is replaced with:

- an XML vocabulary based on ID-IDREF relationships and XLINK linking;
- an RDFa schema using a metadata element set defined by the proposed microformat, or even an ontology for eye tracking metrics.

The data exchange in a browser-server environment is easily accomplished by the AJAX mechanism, especially during reading documents, a process not as affected by network latency as the interactions with a web application. The AJAX mechanism provides ways of triggering client-server data transfers by a wide variety of events, other than activating hyperlinks or submitting forms, which were the main HTTP triggers in the pre-AJAX era.

One of the events that can trigger HTTP requests with POST or GET data attached is MouseOver. In an eye tracking environment with mouse emulation, this translates into an EyesOver event attached to key elements previously structured and identified within the document.

Once the eye tracking data is captured, there are two ways of sending it to the server:

- After each relevant fixation a data set is sent, describing the fixation (id of the fixated element, pupillary diameter, saccade source, semantic annotations);
- A more complex data structure (a JavaScript object or a JSON object representation) is incrementally created during the human-computer interaction, by adding all relevant data sets and chain-linking gaze paths during hyperlink activation. The data structure is sent at a later point in time.

JSON provides a string representation of objects, a more efficient alternative to XML based on delimiters rather than tagging:

- Comma-delimited pairs (key:value) nested between {...} represent an object;
- Comma-delimited values nested between [...] represent an array;
- JSON is recursively defined, thus the elements of an object or an array may consist in objects or arrays.

Exchanging JSON between the JavaScript client script and the web servers is widely supported by AJAX frameworks.

The versatility of the JSON data structure permits the emulation of metadata and even of RDF triples: property=value pairs (or predicate-object pairs) can be represented as key:value pairs and properly delimited and nested in order to subordinate them to a "subject" identifier.

A simplified example of a JSON structure containing eye tracking data would be:

```
attention={
user:Joe Smith,
saccade:{
        from:{
                id:scr1fix1,
                nanocontent:.....,
                pupildiam:........
                },
        to:scr1fix2
        },
saccade:{
        from:{
                id:scr1fix2,
                nanocontent:........,
                pupildiam:..........
                },
        to:scr1fix3
        },
..............
screenswitch:{
                from:{.........},
                to:{..........},
                rel:nextscr,
                arrival:.........,
                departure:.......,
                rev:prevscr
                },
saccade:{........},
.............
}
```

3.2 Side results

Also, several side results were indicated by our current studies:

- Text skimming effectiveness depends on the type of language in which the text is written. Text understanding through skimming was significantly higher if the text was written in a language where certain "logic operators" are better emphasized among other words. For example, the negation operator is better emphasized in English (*do not, does not, doesn't*) or French (*ne...pas*) than in languages where it consists in only one or two letters, like Romanian (*nu, n-*). This sometimes has a fundamental effect on semantic accuracy of text skimming, since a skimming reader sometimes (20% of the subjects) understands the negated message instead of the message itself (the test involved a multiple choice test with weak formatting and very little time to read the question and possible answers);

- By using visual cues such as bolding certain logic operators (negation, conjunction, disjunction), the accuracy of text skimming in Romanian was increased to the level of the other languages;
- 80% of the subjects skimmed through documents even when there was no time pressure;
- The "digital illiterates" employ word by word reading rather than skimming (8 of the 10 non-frequent readers) even under time pressure;
- However, "digital literates" are better at skimming hypertext and their scan-paths are more clearly defined, thus the customization potential is greatly increased;
- Inverted pyramid reading style was employed even in traditionally written technical papers (the "jumping to conclusion" strategy);
- Having a list of headlines to choose from and a limited time to choose, 88% chose a headline without leading articles ("the" or "a" or similar constructs in other languages).

3.3 Challenges for future development

Our work in progress meets several challenges, some approached by our research efforts, others to be solved by the progress made in eye tracking hardware development:

- The extension of the proposed microformat to an eye tracking ontology which would provide the metadata set for RDFa annotations. This would also include taxonomies for saccades, fixations, pupillary activity, nanocontent and microcontent;
- The definition of a gaze-oriented event framework which would eliminate the need of using mouse emulation;
- The design of user interfaces more appropriate to gaze control, which is currently very fatiguing and unstable (at least with shareware and freeware tools);
- Overcoming user limitation, since gaze tracking fails or is inaccurate for some participants, especially those who tend to partially cover their pupil with an eye lid and those wearing certain types of glasses/lenses;
- The development of robust drivers in order to integrate eye tracking with modern browsers, thus making gaze data directly available to the browser object model.

4 Conclusion

Eye tracking offers great potential for web customization. Currently web customization is mostly about detecting the delivery context (device detection, browser sniffing etc.) and analyzing the data provided consciously by the user (ratings, purchases, data created by social networking activities). Through eye tracking, the user provides better refined data regarding its interests with respect to the content provided visually by a web site. Text skimming is a process that can be analyzed in order to extract preferences, both for concepts and relations between concepts, even irrational and subconscious relations that only exist for a certain user. Our further studies will look for solutions to minimize user involvement in recording the data, specifically to eliminate narration and build a level of metaconcepts (RDFS) for modelling the eye activity during human-browser interaction.

5 Acknowledgement

This paper represents a result of the research project with the code 2443/2009 funded by the National Research Program IDEI, managed by Lect. Robert Andrei Buchmann.

References:

[1] Arh T., Blažič B.J., *A portal case study of usability testing – the SUMI evaluation approach of the EducaNext, WSEAS Transactions on Information Science and Applications*, Issue 2, Volume 5, February 2008, pp.175-181

[2] Beatty J., Lucero-Wagoner B., *The Pupillary System, Handbook of Psychophysiology*, Cambridge University Press, 2000, pp.142-162

[3] Buchmann R.A., *Rolul XML in interoperabilitatea sistemelor informatice pentru afaceri*, Ed. Risoprint, 2007

[4] Jacob R.J.K., *The use of eye-movements in human-computer interaction techniques: what you see is what you get, ACM Transactions on Information Systems*, 9(2), 1991, pp.152-169

[5] Landow G., *Hypertext 3.0: Critical Theory and New Media in an Era of Globalization*, The Johns Hopkins University Press, 2006

[6] Loewenfeld I.E., *The Pupil: Anatomy, Physiology and Clinical Applications*, Butterworth Heinemann 1999

[7] McGann, *Radiant Textuality: Literature after the World Wide Web*, Palgrave Macmillan, 2004

[8] Nielsen J., Loranger H., *Prioritizing Web Usability*, New Riders, 2006

[9] Rayner K., *Eye movements in reading and information processing: 20 years of research, Psychological Bulletin,* 124(3), 1998, pp.372-422

[10] Redish J., *Letting Go of the Words*, Morgan Kaufman, 2007

[11] Rommelse N.N., Van der Stigchel S., Sergeant J.A.,. *A review on eye movement studies in childhood and adolescent psychiatry. Brain and Cognition,* Vol. 63, Issue 8, 2008

[12] Salvucci D.D., Anderson J.R., *Intelligent gaze-added interfaces, Proceedings of the conference on human factors in computing systems*, ACM Press, 2000, pp.273-280

[13] San Agustin, J., Skovsgaard, H., Hansen, J. P., and Hansen, D. W. 2009. *Low-cost gaze interaction: ready to deliver the promises*, *Proceedings of the 27th international Conference Extended Abstracts on Human Factors in Computing Systems* (Boston, MA, USA, April 04 - 09, 2009). CHI EA '09. ACM, 2009, pp.4453-4458

[14] Sesin A., Adjouadi M., Ayala M., Barreto A., Rishe N., *Effective data conversion algorithm for real-time vision based human computer interface, Proceedings of the 6th WSEAS Int. Conf. on Software Engineering, Parallel and Distributed Systems*, Greece, February 16-19, WSEAS Press, 2007, pp. 121-126

[15] Sesin A., Ayala M., Cabrerizo M., Adjouadi M., *Jitter reduction in eye gaze tracking system and conception of a metric for performance evaluation, WSEAS Transactions on Computers*, Issue 5, Volume 3, November 2004, pp.1268-1274

[16] Wang W.C., Chen Y.H., Kao C.Y., *The analysis of store image concerning the visual point tacking of the consumers on viewing the window display, WSEAS Transactions On Business And Economics*, Issue 2, Volume 6, February 2009, pp.87-97

[17] Wedel M., Pieters R., *Eye tracking for visual marketing, Foundations and Trends in Marketing*, Vol.1, Issue 4., 2006, pp.231-320

[18] Zhai S., Morimoto C., Ihde S., *Manual and gaze input cascaded (magic) pointing, Proceedings of the SIGCHI conference on human factors in computing systems*, ACM Press, 1999, pp.246-253

[19] http://code.google.com/p/xmlgrddl/

[20] http://didaktik.physik.fu-berlin.de/projekte/ogama/

[21] http://joelclemens.colinr.ca/eyetrack

[22] http://microformats.org/

[23] http://thirtysixthspan.com/openEyes/

[24] http://www.codeproject.com/KB/cpp /TrackEye.aspx

[25] http://www.gazegroup.org/downloads /23-gazetracker

[26] http://www.inference.phy.cam.ac.uk/opengazer/

[27] http://www.tobii.com

[28] http://www.w3.org/TR/rdfa-syntax/

[29] http://www.w3.org/TR/grddl/