



ISSN Print: 2394-7500
ISSN Online: 2394-5869
Impact Factor: 5.2
IJAR 2017; 3(6): 872-874
www.allresearchjournal.com
Received: 01-04-2017
Accepted: 02-05-2017

Ahmed Ali
OIU, TU Comp. Science &
Technology, College of
Education and Science –
Khurma Taif University - Taif
Saudi Arabia

Seaf Eddin Fattoh
Emirate College for Science
and Technology, Dean of the
College Khurma Taif
University - Taif
Saudi Arabia

Sami Gorashi
OIU, TU Comp. Science &
Technology, College of
Education and Science –
Khurma Taif University - Taif
Saudi Arabia

Correspondence
Ahmed Ali
OIU, TU Comp. Science &
Technology, College of
Education and Science –
Khurma Taif University - Taif
Saudi Arabia

Automatic focusing on intended internet text (Nanotechnology case study)

Ahmed Ali, Seaf Eddin Fattoh and Sami Gorashi

Abstract

Automatic focus on the specific text on Web pages is important because it draws the user's attention to important information, this information may led to a tremendous amount of information required in the field of interest.

In this paper we have programmed a system that will automatically highlight certain words by area of interest of the user, although we have taken nanotechnology science as a case study to apply our system, our system is generic and can be used for other areas just by reconfigure the database, this system automatically without user intervention highlight certain words excite the attention of the user who shall click on the highlighted word to receive too much information in the area of his interest from dedicated database.

Keywords: Automatic search, text processing, internet text, intended text, focus on

1. Introduction

Focusing on some intended word or phrases is highly important due to many applications such as intention drawing, advertisement selection, etc.

There are many search engines available to users for use in their research and also there are many databases and special tools, but the search engines and databases or tools require the user to do something "such as typing a word or phrase" in order to find more information about them, we consider this is a first problem, the second one is there will be huge amount of results for this search that may be more broaden and confused.

There are also some applications in this area, such as CDAPubMed ^[1]

CDAPubMed is a platform-independent tool designed to facilitate literature searching using keywords contained in specific EHRs but this application also need an action from user unlike our proposed system.

Andrew S. Gordon, Qun Cao, and Reid Swanson ^[2] describe efforts to automatically capture stories from Internet weblogs by extracting them using statistical text classification techniques. They evaluate the precision and recall performance of competing approaches. They describe the large-scale application of story extraction technology to Internet weblogs, producing a corpus of stories with over a billion words.

In our system the user does not need to write a word or phrase to search, but the proposed system draws the attention of the user by highlighting a word or phrase indicating the importance of them.

Our proposed tool was designed by creating browser extension, this browser extension.

2. Background

2.1 Internet Text Highlighting

Highlighting text on a web page can help bring important information immediately to the reader's attention, in HTML5 highlighted text can be easily done by applying <mark> tag, by this tag you could also highlight one word or several words within a paragraph, to draw attention to specific words that are most important, also Highlighting text can be done by creating a class inside cascaded style sheet (CSS) file ^[3].

In figure (1) the code shows that there are three elements being defined. First the body background color is set to blue, a new class called "highlightme" with a yellow background has been created, and the paragraph tag will now have a white background.

After setting up the CSS code, if you wanted to use the "highlightme" class to highlight your text, you would use the tag and reference the CSS class, as shown in the example code below.

```
<style>
body { background-color:blue; }
.highlightme { background-color:#FFFF00; }
p { background-color:#FFFFFF; }
</style>
```

Fig 1

2.2 Browser Extension

2.2.1 What is a Browser Extension?

A browser extension is a plug-in that extends the functionality of a web browser. Some extensions are authored using web technologies such as HTML, JavaScript, and CSS. Others are developed using machine code and application programming interfaces (APIs) provided by web browsers, such as NPAPI and PPAPI [4]. Browser extensions can change the user interface of the web browser without directly affecting viewable content of a web page.

2.2.2 Browser Extension Functions?

Browser extensions are used for improving a browser's user interface, security or accessibility, blocking advertisements and various other features to make browsing the internet easier and more pleasant [5]. There are many types of extensions that can be used to control various aspects of browsing privacy and mitigate threats. For instance, they may prevent third parties from tracking the user's movements, block ads and scripts, or enforce good habits.

2.3 Design

The system proposed in this paper must meet some requirements to be functioning well, these requirements are as follows.

- The system should be capable of searching for nanotechnology keywords in the current web page automatically.
- Nanotechnology keywords must be highlighted by the system to draw user attention.
- The system must retrieve nanotechnology literature according to these keywords when they were clicked by the user.

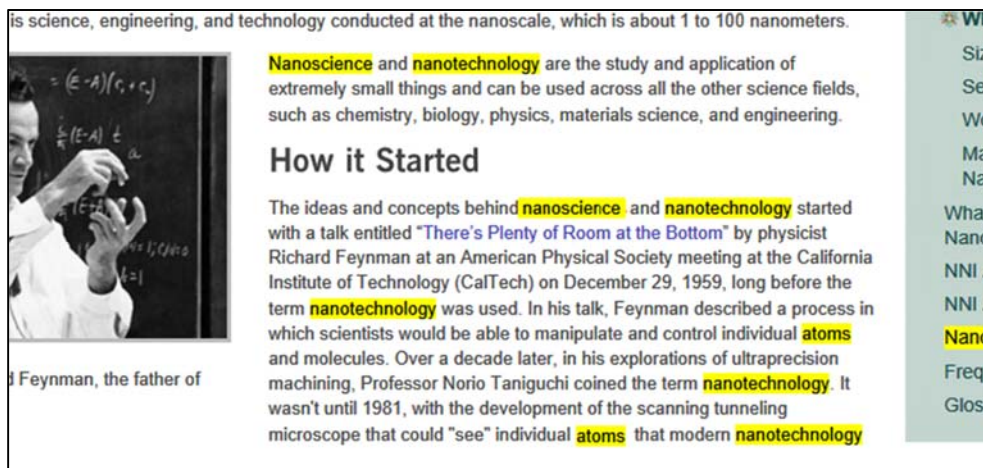


Figure 2: shows some keywords that highlighted by the system.

3. Implementation

3.1 Query generation

To generate a query through our system, the user may select one of the highlighted words or phrases in the current page by clicking it. For each selection, the system will search for the intended word or phrase in our huge database (NHD), this database prepared specially for nanotechnology literature, hence the user will obtain more complete information and knowledge about nanotechnology.

3.2 Architecture and technologies

A user must install our proposed tool on his browser, this tool is consist of two components and one database, one of the two components called Nanotechnology keyword finder (NKF), this component scans the user current page word by word to find any keyword that related to nanotechnology sciences then draw a user's attention to it, another component is called Nanotechnology Search Engen (NSE) which is responsible for nanotechnology information retrieval.

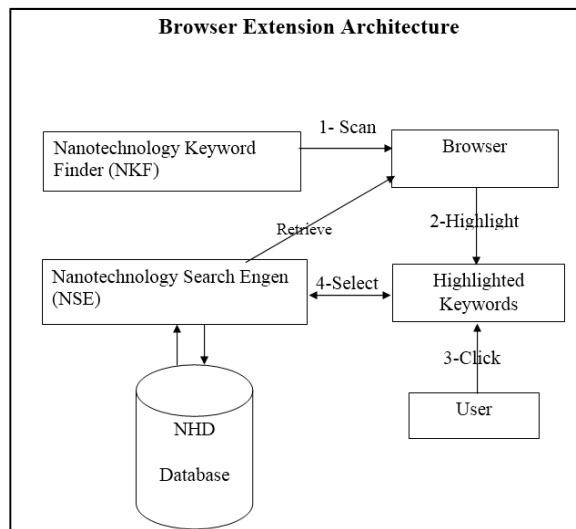


Fig 3

3.2 Query generation

To generate a query through our system, the user may select one of the highlighted words or phrases in the current page by clicking it. For each selection, the system will search for the intended word or phrase in our huge database (NHD), this database prepared specially for nanotechnology literature, hence the user will obtain more complete information and knowledge about nanotechnology.

4. Conclusions

This paper proposed a tool that can draw user's attention for existing important information on his interest in a browser automatically in real time without searching, this solution enables users to find large amount of information on their interest without typing any word of search.

Results from preliminary user studies indicate that the basic functionality of the extension is easy to use.

5. References

1. David Perez-Rey, Ana Jimenez-Castellanos, Miguel Garcia-Remesal, Jose Crespo and Victor Maojo. CDAPubMed: a browser extension to retrieve EHR-based biomedical literature.
2. Andrew S, Gordon, Qun Cao, Reid Swanson. Automated Story Capture From Internet Weblogs.
3. Computer Hope. [www.computerhope.com/issues/ch001391.htm].
4. Wikipedia. [https://en.wikipedia.org/wiki/Browser_extension].
5. Henry, Alan 2015. The Best Browser Extensions that Protect Your Privacy.
6. Google. [http://www.google.com].
7. Yahoo. [http://www.yahoo.com].
8. Bing. [http://www.bing.com].