Inside-In Search: an alternative for performing ancillary search tasks on the Web

Ricardo Cava, Carla M. D. S. Freitas
Instituto de Informática
Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre – RS, Brazil
{racava,carla}@inf.ufrgs.br

Eric Barboni, Philippe Palanque, Marco Winckler
ICS-IRIT
Université Paul Sabatier
Toulouse, France
{barboni, palanque, winckler}@irit.fr

***Abstract*— Some of the search tasks users perform on the Web aim at complementing the information they are currently reading in a Web page: they are ancillary search tasks. Currently, the standard way to support such ancillary searches follows an inside-out approach, which means that query results are shown in a new window/tab or as a replacement of the current page. We claim that such inside-out approach is only suitable if users really want to dissociate the search results from the Web page they were reading. In this paper we propose an alternative approach, called "inside-in", where query results are displayed inside the Web page next to the keyword that motivated the user to launch an ancillary search. In order to demonstrate the feasibility of our approach we have developed a tool that embeds an egocentric information visualization technique in the Web page. This tool supports nested queries and allows the display of multiple data attributes. The approach is illustrated by a case study based on ancillary searches of coauthors from a digital library. The paper also reports some preliminary results obtained with an experiment conducted with remote users.**

Keywords— Visualization techniques; Web interaction techniques; information retrieval; ancillary Web queries; Web-based user interfaces.

I. Introduction

Search tasks might vary in complexity: they can be relatively simple such as looking for current weather forecast or otherwise tricky and time consuming such as collecting all information required for planning a vacation trip overseas. Over the years, users have become used to retrieve information from the Web, and for that they developed several strategies, which can be summarized as information lookup and exploratory search [16]. Whilst exploratory search requires time for scanning and reading documents, information lookup can be solved by simple factual question-answer interactions.

Regardless the users' need for information, the user interface provided by such information retrieval systems must be simple enough to allow users to formulate queries and understand the results provided by search engines [10]. Quite often, queries start by filling a search box with keywords. Formulating queries in this way is a daunting task that requires users to open a new window, to fill in a form with appropriate keywords, and then scan the list of results until finding the one that corresponds to the goal. Nonetheless, many search tasks can also be accomplished by browsing (or navigating) among interlinked documents. Browsing is often preferred by users because it is cognitively less demanding to recognize a keyword than typing one, but the advantages of browsing is quickly lost if users have to visit many links until reaching the desired information [15].

Currently, the standard way to show search results (obtained by either filling in a form or browsing documents) follows an *inside-out search* approach, which means, results are displayed in a new window/tab and/or replace the current window/tab's contents. We assume that such *inside-out search* approach is only suitable when users want to dissociate the search results from the Web page they are reading. However, some searches that users perform on the web are done just to complement information they are currently reading [12], the so-called *ancillary searches*. For example, users reading an article in a Web page might be curious to know with whom the author of that particular article has published in the past. In this scenario, looking up for co-authors constitutes an ancillary search, which is not meant to divert users' attention from reading the article. For such kind of ancillary-search tasks, using the *inside-out* approach to display results in a new window/tab might be unsuitable since it creates an articulatory distance between the origin of the request and the information display making difficult to users to assess if their goal has been fulfilled or not by the query.

In this paper we propose an alternative approach, called *inside-in* search, where ancillary searches are launched by users from inside the Web page they are currently reading, and the corresponding results are displayed as a contextual help next to the keywords used to formulate query. Giving that the amount of search results from an ancillary search can be huge, we assume that users would benefit from an interactive information visualization technique that provides nested queries (ex. co-authors of a particular co-author), and allows the display of multiple attributes as ancillary data (ex. co-authors, number of joint publications, type of publications in common).

In order to demonstrate the feasibility of the *inside-in* approach, we have developed a supporting tool that embeds an egocentric information visualization technique called IRIS. For the purpose of this paper we present a case study focused on the visualization of co-authors extracted from the DBLP. We present the theoretical background (section II), followed by a task analysis of ancillary searches using an *inside-in* approach (section III). Sections IV and V present

978-1-4799-6953-1/14 $31.00 © 2014 IEEE
DOI 10.1109/LAWeb.2014.21

the implementation of our approach and some preliminary results obtained from a remote user testing experiment. Section VI briefly reviews related work, and section VII concludes our paper, emphasizing the current contribution and providing some insights on future work.

II. THEORETICAL BACKGROUND

In order to better understand the tasks accomplished by users while performing search tasks, Sutcliffe and Ennis [23] propose an information seeking process that encompasses the following steps: i) problem identification, ii) articulation of information need(s), iii) query formulation, and iv) results evaluation. These steps recall Norman's cognitive model [19], which explains the gulfs that exist in the communication between users and systems: the *execution gulf* and the *evaluation gulf.* These concepts were introduced by D. Norman (1986) and became popular through the book [18].

The *execution gulf* is the effort required for a user to express an intention in terms of commands or instructions. In other words, the *gulf of execution* is the difference between the intentions of the users and what the system allows them to do or how well the system supports those actions. For example, if the users find an unknown word whilst navigating the web, they might expect that clicking on a link (on that word) would provide them with the complimentary information required to understand the meaning of the word. In the user's language "click the link" defines the goal for obtaining the word's meaning. However, if the link does not provide the expected results, users have to execute additional actions, such as opening a new window, visiting a search web site, typing the adequate keywords to specify the search, and, finally, browsing the list of results until getting the desired definition.

The *evaluation gulf* refers to the way the results provided by the system are meaningful or understandable by the users, and in accordance with their goals. In other words, the *gulf of evaluation* is the degree to which the system or artifact provides representations that can be directly perceived and interpreted in terms of the user's expectations and intentions. Thus, if the system does not "present itself" in a way that lets the users derive which sequence of actions will lead to the intended goal or system state, or infer whether previous actions have moved them closer to their goal, there is a large *gulf of evaluation*. In our case, the users must spend a considerable amount of effort and significant attentional resources to perform a query in a new window, to identify the answers that correspond to their expectations and, then, to put the word's meaning back in the appropriate context.

Overall, the *gulfs of evaluation and of execution* refer to the mismatch between our internal goals on the one side, and, on the other side, the expectations and the availability of information specifying the state of the world (or an artifact) and how we may change it [18].

Fig. 1 presents a revised version of the execution and evaluation gulfs cognitive model [19] explicitly showing the *articulatory distance* and the *semantic distance* both in terms of user *input* (i.e. when users formulate the query) and in terms of system *output* (i.e. when the system displays the results to be assessed by the user).

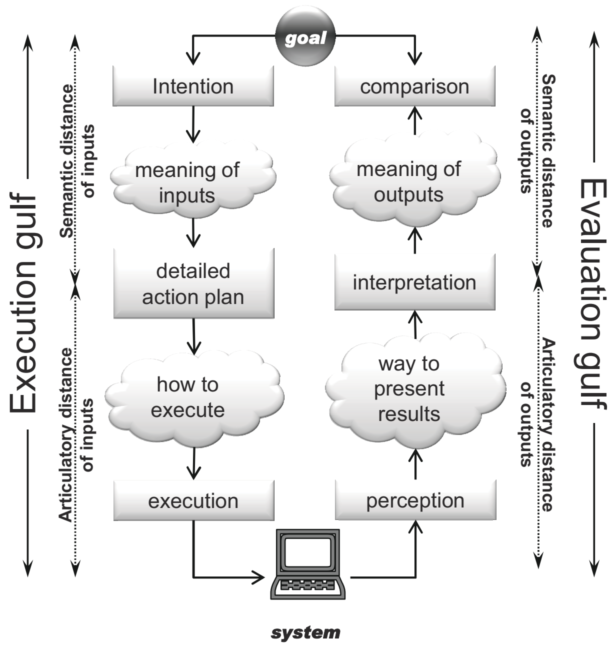


Figure 1. Execution and evaluation gulfs in search tasks, adapted from the Norman's model [19].

Based on this model, we can compare two user interfaces for looking for information on the Web: *browsing* and *filling-in forms*. The *semantic distance of input* can be considered smaller in *browsing* than *filling-in forms* because recognition of words is less cognitive demanding than choosing the appropriate keywords. The *articulatory distance of input* is also smaller for *browsing* because it just requires a click, which is much faster than typing words. However, the *articulatory distance of output* is bigger when users have to navigate/*browse* many Web pages to find the information. This *articulatory distance* is smaller when the list of pages possibly containing the information needed is displayed in a rank. Nonetheless, ranks might increase the *semantic distance of output* as users might not (necessarily) understand how the ranking was created and in which extension the search results are relevant to the information s/he looking for.

III. TASK ANALYSIS OF ANCILLARY-SEARCH TASKS

Some of the difficulties users have to face when using an *inside-out* approach become evident when users perform ancillary search tasks. To illustrate this, we present below a scenario, which ultimately require users to perform ancillary searches for better understanding the contents of an article s/he reading on the Web.

A. Motivating scenario

In order to ground the scenarios around the same application domain, we have chosen to illustrate them with data about co-authorship, as follows:

"John has been recently appointed as expert member of the jury that will evaluate the research of a Graduate Program in Computer Science at an university in the South of France. John received a Web form which contains the list of ~400 researchers for which he has to provide an assessment based on the number of co-authors and relevant publications they have in the field. The number of publications and co-authors is required to calculate two important metrics: the researcher's productivity (accordingly to a formula that takes into account the number of co-authors to estimate the individual effort for the publication) and the size of collaboration network (considered that successful scientific collaborations ultimately lead to joint publications). So, John starts by making a Google search on the Web using the name of the first researcher in the list. Find the right researcher's Web page was not easy as the Google search engine returns many entry points including homonymous and some trash pages. After fixing typos and refining the terms of the query, John found the researcher's Web page where he can count the number of his publications; the mental calculation to accomplish this task is easy. Now, for assessing the size of the research network, things are more complicate. John considers two options: i) to create manually a side-list with the names of co-authors; or ii) look for them in the DBLP web site. John chose the second option so he types the name of the researcher on the search box of the navigation, goes to DBLP web site, scrolls down to reach the zone where co-authors are displayed, and open up the list of co-authors. Now John is ready to fill in the form but then he realizes that the DBLP contents now occupy the window that previously contained the Web form … For the next 399 researchers John decided to create new tabs for keeping the DBLP search apart from the Web form. Then, he found out himself being performing repetitive copy-and-paste between tabs which did not improve his overall performance…"

As we can observe in the described scenario, searching co-authors should be considered an ancillary search that complements the user's main task, which is filling in the Web form. From this scenario we find some issues that make the following users tasks difficult:

- Formulating queries is error-prone (might contain typos) and also time consuming (typing takes time).
- Keywords might be ambiguous and generic search engines will return broad results. Users may have to scan the list of results until finding the one that corresponds to their goals.
- There are many alternative locations for showing results (including new windows/tabs); choosing the best location for displaying results depends on where the results are meant to be used.
- Some queries might be repetitive; so, saving a few seconds in the time required to complete a single task might represent a huge improvement in the overall task performance at the end of the day.

B. Rational for improving ancillary-search tasks

We claim that the issues raised above can be solved (or at least minimized) with an *inside-in* approach including the following mechanisms aimed at supporting ancillary-search tasks:

- Launching queries from words available in the current Web page can reduce typos. Keywords can be selected with mouse click, which is sensibly faster than typing in using a keyboard.
- Ambiguous results are often the result of a broad search. This problem can be reduced by providing specialized queries that operate on specific application domains using user-selected keywords.
- Query results can be shown inside the current page inline to the selected keywords. This is one of the keystones for the *inside-in* approach, but notice that queries should be launched on user's demand. If the system systematically launches queries without user's request, the Web page will become polluted by ancillary results, and the benefits of the *inside-in* approach will be lost.
- Ancillary results should support some kind of interaction to allow users to perform nested queries. This element is important for repetitive tasks, which are often associated to contexts where ancillary searches are required.

The selection of keywords on the text and the use of predefined queries aim at reducing the gulf of execution. This reduction is achieved by minimizing the users' effort in informing keywords to the system (articulatory distance of inputs) and by favoring recognition of keywords and queries rather than imposing the formulation of complete queries (semantic distance of inputs). Predefined queries also help to reduce the evaluation gulf as the results are focused on a particular application domain (semantic distance of outputs). By showing results in the same page and allowing the user to perform nested queries, the *inside-in* approach helps to reduce the *articulatory* distance of outputs.

IV. Implementation of the *Inside In* Approach

In order to support the proposed *inside-in* approach, we have developed a framework whose main principles are briefly illustrated in Fig. 2. Our *inside-in* approach is built upon the concept of Web augmentation [3], which defines strategies for implementing tools that can extend the set of elementary tasks users can do while navigating on the Web. Whilst the full details about the implementation of that framework are out of the scope of this paper, it is interesting to notice that it includes a client-side module and a broker at the server-side.

A. The server side module

The "broker" at the server side (Fig. 2.b) contains a set of preprogramed query functions that are made available to the end users. The query broker was originally conceived to support robust integration of many similarity functions using approximated data instance matching [8]. Nonetheless, the framework is extensible and can accommodate new specialized queries. The number of queries accessible from the client-side can be configured dynamically, but for the purposes of this paper we are just using a specific one, which returns the co-authors of a given researcher. The only thing users have to do is to select keywords from the set of words displayed in the current Web page, and then trigger (by a simple click) the ancillary-search query for co-authors.

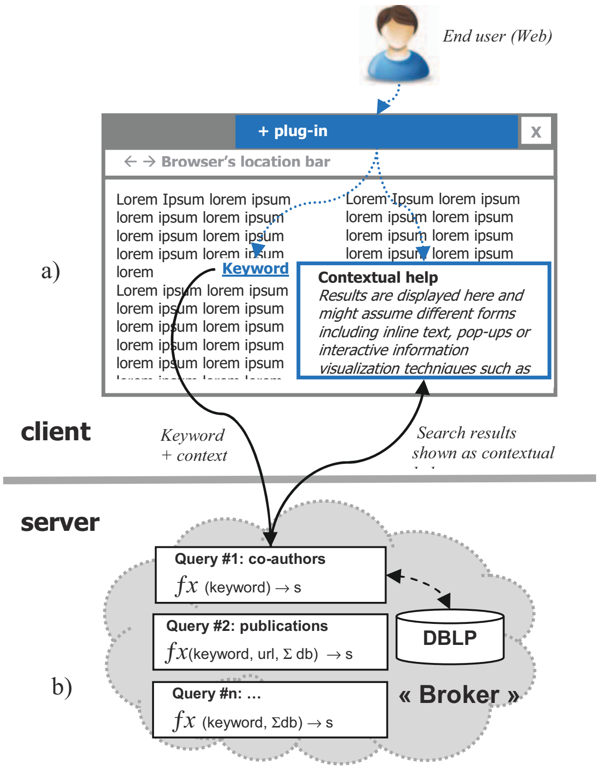


Figure 2. Overview of the framework architecture.

B. The client-side module

This module (Fig. 2.a) allows users to select keywords in the current web page and to trigger the queries for ancillary search available in the broker. Once the broker replies, this module modifies the current Web page to display the search results as a kind of contextual help. For that, the DOM structure of the current Web page is adapted using a set of JavaScript functions, called augmenters [10]. As demonstrated in a previous work [28], adaptations created by augmenters are volatile and do not affect the application running in the Web server. The client-side module can display ancillary data using different interaction techniques, including the information visualization technique IRIS, which is described in the next section.

C. The information visualization technique IRIS

The information visualization technique IRIS (which stands for Investigating Relationships between Indexes of Similarity) was originally conceived to provide visualizations of search results containing some measure of the similarity between the retrieved item and the query terms. Nonetheless, IRIS is a generic visualization technique, and can be used to display many different types of data. The inner data structure supported by IRIS is composed of a set of attributes that can be attached to two main *nodes* and *edges* featuring a radial layout [9]. One of the nodes is defined as the *centroid* to which all the other nodes are linked. Whilst *nodes* and *edges* are fixed and mandatory elements in the visualization, the list of attributes attached to them is arbitrary. Thus, IRIS can be configured to work with a large variety of data sets that could be organized around an egocentric network topology.

In Fig. 3, we illustrate the use of IRIS presenting relationships between an author and his/her n co-authors. The name of the author used as keyword to launch the ancillary search about co-authors is placed at the center of a circle formed by bar graphs. The size of the each bar graph represents the number of co-authored papers between the author and each co-author. The name of the co-author and an ordinal number are also displayed radially. To deal with authors that have a large number of co-authors that would not fit into the representation, we adopted a focus+context approach [5]. The focus is the darker center-right area of the circle, and the context is represented by the two sub-areas in light gray where the font is displayed in decreasing size towards the left side of the circle. This effect is obtained by a fisheye distortion [1]. As we can see, only 47 of 83 co-authors are shown. The ordinal number displayed along the name of the co-author provides information about which subset is being displayed at each moment. Users can move the position of any co-author to the focus area by clicking on his/her name.

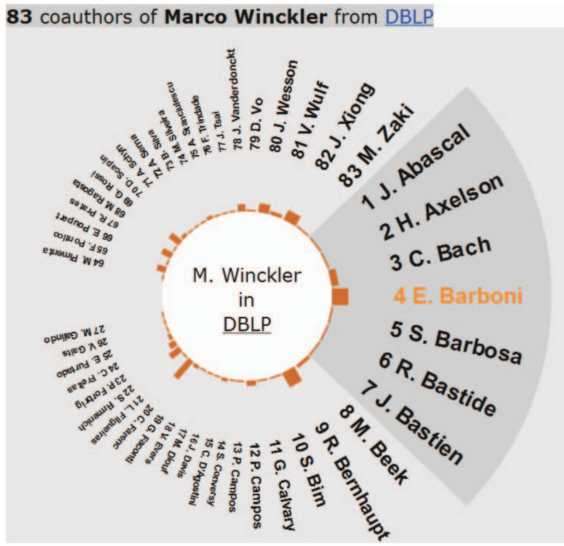


Figure 3. IRIS visualization of co-authorship of information from DBLP.

The implementation of IRIS that is delivered with our framework also embeds several functions for allowing interactive exploration of the dataset. In our case study of co-authorship network, a centroid is made up from the keyword selected by the users on the Web page. The nodes and attributes are obtained as the result of an ancillary-search query.

Other details regarding the attributes can be obtained by activating tooltips which show the complete name of a selected co-author (Fig. 4.a) or the number of papers per category, as provided by DBLP (Fig. 4.b). The link shown below the name of the author gives access to the information in DBLP.

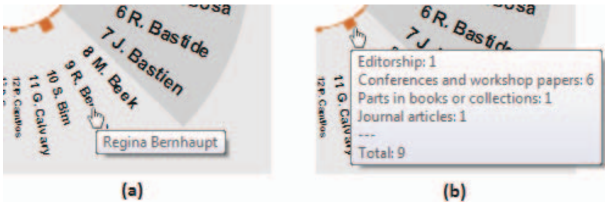


Figure 4. Details of IRIS representation showing tooltips for: a) co-author's name; and b) number of papers per category.

IRIS also provides some interaction beyond the display of tooltips. The user can place a co-author in focus by clicking on it with the left mouse button. This causes the name change its color to orange and move to the center of the focus area. The whole representation is rotated, the direction of the rotation being determined in a way to minimize the angle to be used. The movement is animated to minimize user's disorientation [13][25].

V. Evaluation

In this section we describe a remote user study we have run in order to collect real users' feedback about our tools.

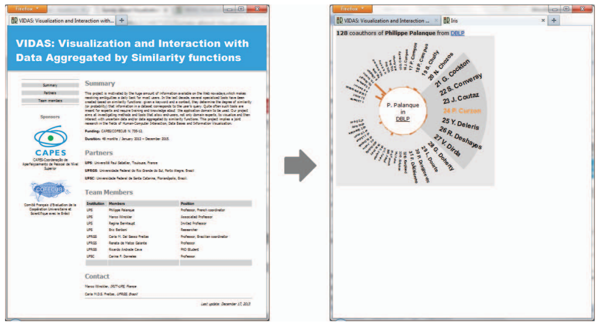
A. *Rationale and scenario used in the evaluation*

The study was focused on a population of users that might have the need (or curiosity) to look up certain type of information while browsing the Web. Moreover, it was important to envisage a scenario where the information needed is easily available by other means than our visualization technique, so that users could compare the different design alternatives. Therefore, we have considered a scenario where there is list of researchers shown in a web page and a user wants to check their co-authors in some bibliographic database, in the present case, DBLP, which is easily accessed at http://www.informatik.uni-trier.de/~ley/db/. So, users could use either DBLP or IRIS to see the same results.

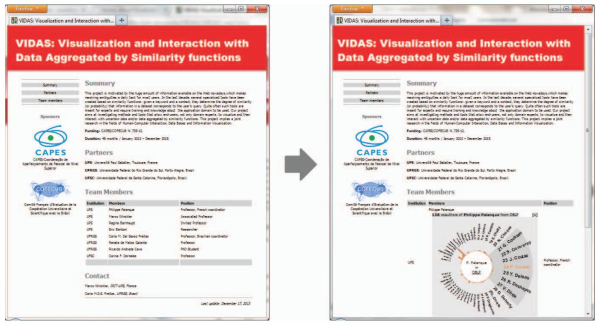
In our scenario, we assume that, given a community of end users, empirical studies can be performed to identify suggestions for ancillary questions that fulfill users' needs in a particular application domain. For the purpose of this paper, we assume that end users are researchers, principal investigators searching for postdocs, deans, experts of funding agencies, etc., for which looking up for co-authors is part of some of their tasks.

For reducing the semantic and articulatory distance of inputs, we devised a browsing mechanism that is based on clicking on names of researchers already available in a Web page. Selecting a name embedded into a Web page launches a predefined search for his/her co-authors so that users do not need to type keywords (as in DBLP or Google). Moreover, the scope of the search is predefined (i.e. co-authors) as well the data source (i.e. DBLP).

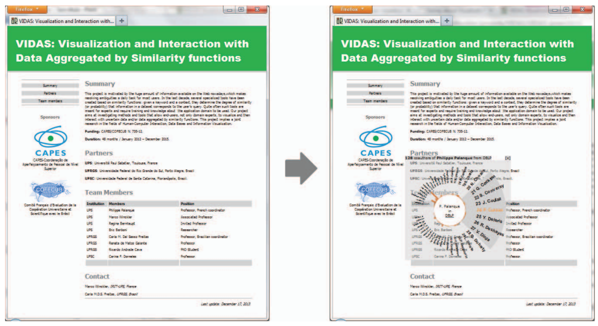
For the purpose of the survey, we used the team members of the VIDAS project's Web page as shown in Fig. 5. However, the design solutions presented herein are generic, and could be used in any Web page using the names it might contain. The list of co-authors is obtained by parsing data from DBLP and displaying it using IRIS. To see the co-authors of a team member one just need to click on (or select) his/her name in the Web page.



a) Results are shown in a new tab on the Web browser.



b) Results are embedded into the Web page changing layout.



c) Results are shown as a semi-transparent layer.

Figure 5. Alternative design options for positioning the search results.

As for the semantic and articulatory distances of outputs, our alternative designs show the search results encoded in an abstract visualization and display them in three ways as illustrated in Fig. 5:

1. *Displayed in another tab/window*. We assume that the articulatory distance of output is greater in this option than in the other ones as results are not shown in the same Web page where the search was triggered.

2. *Embedded into the current pag*e. Notice that this option will change the Web page layout on the client-side to place IRIS next to the position selected by the user. However, users can easily go back to the original page layout just by clicking on the icon [x] available at the top-right site of IRIS panel. Articulatory distance of output is reduced by placing results in the same page, though changing Web layout could be disturbing for some users.

3. *Placing IRIS as a semi-transparent layer over the page*. In this case, the Web page layout does not change, but IRIS can hide some information. Nonetheless, users can freely move IRIS around.

B. Procedures

An online survey was created using the tool SurveyGizmo. The survey was organized in five main chapters including:

- *Presentation of the study:* this step described the task "search co-authors" and how to launch IRIS.
- *About ancillary queries:* using a Likert scale participants were asked to tell if they think whether (or not) "*starting a query directly from a Web page*" by clicking on researchers names was useful, improved performance, and how much they liked it. They were also asked to tell what were the most frequent ancillary queries they triggered while navigating the Web.
- *About IRIS:* using a Likert scale participants were asked if they think that "*showing results of queries on the Web by means of an interactive visualization technique*" was useful, improved performance, and how much they liked it. They were also asked to tell if they prefer a flat HTML page to show query results or IRIS.
- *About alternative locations for displaying search results:* participants were asked to tell which they prefer: ancillary query results in a new tab/window, results embedded into the Web page, the option of results floating over the web page, or none of these.
- *About participants:* the last step addressed the profile of participants including demographic data.

C. Participants

Participants were recruited via mailing lists and social networks of authors involved in this work. The survey was delivered in English regardless the origin of the participants. We have addressed the survey to professors, researchers and students involved in research activities, as that population might feel concerned by the underlying task of looking up information about co-authors as part of their activities. In the short period of 3 weeks, while the survey remained open, we received 150 visitors but only 61 complete responses were used to validate our hypotheses, as described below.

Most of participants were male (77.1%) and, in average 26.1 years old (SD=5.7), being 55.7% between 25-34 years old, 24.6% between 35-54 years old and 19.7% between 18-24 years old. Among the participants 44.3% were students, 39.3% researchers/professors and 16.4% work in the industry. The highest degree included 26.2% of PhD, 44.3% M.Sc. degree, 19.7% Bachelor's degree and 9.8% are still undergraduate students. We have got responses from Argentina, Austria, Brazil, France and Spain. Participants estimated to spend ~5.3 hours (SD=4.1) per week using search engines over the Web, most of which (~4.3 hours per week, SD=3.9) is spent looking for authors and publications. The amount of time participants perform searches related to authors and publications qualify them as typical users of tools with predefined queries for looking for information, such as searching for co-authors. The most frequently reported digital library was GoogleScholar with 88.1%, followed by IEEE Xplore (76.3%), SpringerLink (52.5%), ACM DL (69.5%) and DBLP (35.6%).

D. Results and discussion

The results collected with the survey are quite positive with respect to the *inside-in* approach as implemented by our tools. The results confirm that typing text to formulate a query is less appreciated than selecting keywords, as shown in Table I. Most participants found that selecting a term in a web page for launching a query is useful (N=52; 85.3%) and that it improves performance (N=49; 80,3%). Moreover, only 11 participants (18.1%) prefer typing a text to formulate queries. As mentioned by participant P23, launching a search by selecting keywords on Web pages depends on "…*the amount of text and how it is showed/arranged. The UX won't be the same with large paragraphs...*" Moreover, 47 participants (77%) provided a large list of suggestions for specialized queries that could be launched by simply selecting a term in a Web page. Considering that all open questions in the survey were optional, we might consider that the experience might have somewhat stimulated participants to contribute to the research.

TABLE I. USERS PREFERENCES FOR LAUNCHING QUERIES

	Selecting keywords in the web page						***Typing text***	
	Useful		*Improve performance*		*Prefer*		*Prefer*	
Strong agree	25	**N=52 *85,3%***	18	**N=49 *80,3%***	20	**N=49 *80,3%***	4	**N=11 *18,1%***
Agree	27		31		29		7	
Neutral	1	**N=1 *1,6%***	6	**N=6 *9,8%***	6	**N=6 *9,8%***	22	**N=22 *36,1%***
Disagree	6	**N=8 *13,1%***	6	**N=6 *9,8%***	4	**N=6 *9,6%***	20	**N=28 *45,9%***
Strongly disagree	2		0		2		8	

Regarding the use of a visualization technique as an alternative solution to show the results of ancillary queries, 51 participants (83.6%) said that using IRIS was useful while 44 of them (72.2%) think that it improved performance during ancillary queries. Most of participants (N=50, 82%) prefer to see the results with IRIS. These results, shown in Table II, yet preliminary, validate IRIS as a suitable design solution for the visualization of ancillary queries, at least as far as searching for co-authors are a concern. A significant number of subjects (N=13, 21.3%) are neutral regarding improvement of performance. This can be exemplified by some comments: some users stated that it depends on the application; others posed this dependency on what is provided by the visualization technique – if the technique allows deriving information more easily than a textual list. For example, participant P34 informed to prefer the flat version because "*I am used to HTML pages so I am not sure it would improve my performance*", and participant P111 wrote: "*Only if both solutions allow me to quickly achieve my goal*". These comments suggest that improving user performance is a key factor leading to the adoption of information visualization techniques to display search results.

TABLE II. USERS PREFERENCES FOR THE DISPLAY OF SEARCH RESULTS.

	Search results shown with a visualization technique						***Textual list***	
	Useful		*Improve performance*		*Prefer*		*Prefer*	
Strong agree	23	**N=51** ***83,6%***	12	**N=44** ***72,2%***	20	**N=50** ***82%***	1	**N=7** ***11,5%***
Agree	28		32		30		6	
Neutral	8	**N=8** ***13,1%***	13	**N=13** ***21,3%***	9	**N=9** ***14,8%***	16	**N=16** ***26,2%***
Disagree	2	**N=2** ***3,3%***	4	**N=4** ***6,6%***	1	**N=2** ***3,2%***	27	**N=38** ***62,3%***
Strongly disagree	0		0		1		11	

Considering the alternatives for the location of IRIS, 36 participants (*59%*) said to prefer the option embedding results into the current page while 22 participants (*36,1%*) liked more the design option showing the results floating over the Web page. Most participants prefer to see ancillary results in the same page (total N=58; *95,1%*). Only 3 participants prefer to see ancillary results in a new tab/window (4,9%) as shown in Table III. Interesting enough, among the participants that prefer the floating option, some of them asked to replace transparency by an opaque background, as mentioned by participant P33: "*I like it [floating version] but it would be better if the background wasn't transparent…*".

TABLE III. USERS PREFERENCES FOR THE LOCATION OF RESULTS.

	Preferred location for IRIS
Embedded in the same page	**N=36** | ***59,0%***
Floating over the same page	**N=22** | ***36,1%***
In a new web page	**N=3** | ***4,9%***

It is also interesting to notice that most of the participants clearly pointed out that option Green (floating over the same page) and option Red (embedded in the same page) presented the advantage of reducing the interruption created by search engines when showing the results in a new window/tab. This can be illustrated by participant P67 who wrote: "*Advantage: doesn't disrupt the current page Disadvantage: requires switching tabs*". The frequency on which such disruption was reported in the comments lets us think that participants really notice the articulatory distance created when new windows are open. Moreover, this may also mean a perception of reduced performance as indicated by the comments of participants P2: "*Changing tabs (and losing my thoughts)*" and P33: "*Change of context is annoying.*"

Overall, users did not like the option that shows a new page because of the change of context. This result is compatible with the issues brought by *inside-out* search approaches. However, the few who liked that mentioned the possibility of having more space to display more information. The majority of positive comments were centered on the availability of the additional information right next to the search keyword.

VI. RELATED WORK

Search engines have become an integral part of our daily lives [12], but many users are still struggling to use them to obtain the results they need [11]. Some of the problems users have to face are related to the fact that, given the increasing availability of data in the Web, users should be very precise in the way they formulate their queries. For that, the design of search user interfaces has developed dramatically along the years, from simple keyword search systems to complex combinations of faceted filters and sorting mechanisms.

Wilson [27] claims that the design of the user interface has also an important cognitive impact on tasks performance; thus, search engines should evolve to take into account users' needs. Although these claims are valid, most of research efforts in the area have been focused on two main areas: algorithms for improving the accuracy of search engines with respect to many users concerns and approaches for improving the visualization of Web pages [24]. For example, Schwarz and Morris [21] describe an information visualization approach for augmenting Web pages with data about the credibility of the ranking proposed by the search engine. While such approach does not help users to formulate better queries, it might help users to better select the pages to visit according to the rank of pages proposed by search engines. Capra et al. [4] also proposed to augment the user interface of search results by adding images next to the ranking provided by search engines aiming at helping users to make better decisions. These few examples are illustrative of strategies for improving the design and display of the ranking of results from search engines.

In the last decades, several information visualization approaches have been developed for presenting search results coming either from search engines or widely used databases, such as DBLP, ACM DL, IEEE Xplore, etc. Some search engines with built-in visualization tools have also been developed. The first reports presenting and/or discussing visualization of search results date from the late 90's and early 2000's. Visualization approaches range from 2D plots [14][17], glyph-based techniques [6][20] to 3D designs [2][7][26]. However, although along the years, many different techniques have been evaluated [22] with results favoring visualizations, the majority of web search engines still provide textual lists ordered by some user or tool specified criteria.

It is interesting to notice that current research efforts follow an *inside-out* approach. In fact, most of search user interfaces treated the search task as independent from the rest of the other ongoing user tasks. As far as we know, the *inside-in* approach proposed in this paper is an original contribution that can improve users performance.

VII. Concluding Remarks

This paper proposed a new perspective for looking at the way search user interfaces can be conceived for helping users to perform ancillary-search tasks on the Web. The work typically followed a user-centered design approach. Our initial motivation was to understand what makes searching on the Web so difficult to users. We found out that the predominant approach for searching based on *inside-out* approach is fine when users want to freely explore the information space. However, such approach presents several limitations when the users need to connect the results of search engines with tasks they are performing in another Web page. The proposed *inside-in* approach aims at reducing both execution and evaluation gulfs in the user interaction with search engines. Indeed, one of the key aspects of this approach is to provide a better integration of search results into existing Web pages, where users require complementary information to make their decisions.

Overall the *inside-in* approach is generic and can be implemented using current search engines such as Google or Yahoo! Nonetheless, it can also be implemented using search engines that are suitable to provide more focused and accurate results about data in a specific application domain. Our framework follows this latter approach as illustrated with the implementation of queries for searching co-authors in the DBLP. While looking up for co-authors might be perceived as a very narrow and specific search, it is noteworthy that it is relevant and frequent in the domains of scientific research, and also is a concern to a large population of researchers, students, teachers, and experts from research funding agencies. Moreover, such specialized characteristic can be tuned and adapted according to specific users' needs. Indeed, the main challenge here remains the identification of relevant queries that are suitable to help users to accomplish their tasks.

The tool implementing our approach provided with the framework allows users to: (a) launch a query by selecting a keyword directly in the web page, and (b) display the search results inside the current Web page as a kind of contextual help. In order to support the display of results, we have embedded into the framework an information visualization technique called IRIS. One of the interesting aspects of IRIS is that it is interactive, so users can explore the results and perform nested queries that are meant as ancillary-search tasks.

It is worthy to notice that IRIS is complementary to the *inside-in* approach proposed in the paper but it can also be used as an *inside-out* standalone tool. This aspect about the uses of IRIS became evident during the investigation of design alternatives for displaying the search results of co-authors. Indeed, the alternative shown in Fig 5.a displays the results in another tab of the browser, while the alternatives shown in Fig. 5.b and 5.c adapt the existing Web page to accommodate the results of the ancillary search.

The results obtained by a survey with 61 remote participants confirmed our first hypothesis: most users prefer to launch queries directly from the web page by selecting a keyword. This is not a new finding [16] but indicates that we are in the right path. As for the other three hypotheses, they were confirmed: users also prefer search results being displayed through an interactive visualization technique, located near the search keyword. Regarding location, users expressed to prefer the display of results in a way that does not change their context, this being achieved by two alternatives – displaying the results embedded in the web page, by augmenting it, or displaying them in a floating layer over the same web page.

With these results we also confirm that the semantic and articulatory distances of inputs (*execution gulf*) in the search task are reduced because searching is launched by clicking on a keyword displayed in the Web page. The semantic and articulatory distances of output (*evaluation gulf*) are also reduced in two of our designs (identified as Red and Green) because search results in both are placed in the same page.

Despite the promising results, we know that these are preliminary and there is much work to be done. We would like to measure the distances in the gulfs by performing experiments with direct observation methods. We also intend to proceed with the development of different input and output techniques for performing search tasks since our framework was developed aiming at such studies. The participants of our survey provided a rich set of comments that will allow us to plan further improvements in IRIS as well as to develop techniques targeted for different search tasks.

Future work should include empirical testing with users in a usability laboratory. This step would allow us to assess user performance when performing the tasks and collect more qualitative data via thinking aloud that would better explain the user experience factors that influence the use of information visualization techniques for displaying search results.

Acknowldegments

We are deeply grateful to the remote users that participated in the experiment. We also acknowledge the financial aid from CNPq, CAPES/COFECUB (project VIDAS 735-12) and FAPERGS.

References

[1] Bederson, B. B. Fisheye Menus. Proceedings of ACM Conference on User Interface Software and Technology (UIST 2000), pp. 217-226, ACM Press.

[2] Benford, S., Snowdon, D., Greenhalgh, C., Ingram, R, Knox, I., Brown, C. VR-VIBE: A Virtual Environment for Co-operative Information Retrieval. Computer Graphics Forum, 14(3):349-360, August 1995.

[3] Bouvin, N. O. Unifying Strategies for Web Augmentation. In: Proc. of the 10th ACM Conference on Hypertext and Hypermedia, 1999.

[4] Capra, R., Arguello, J., Scholer, F. 2013. Augmenting web search surrogates with images. In Proceedings of the 22nd ACM international conference on Conference on information & knowledge management (CIKM '13). ACM, New York, NY, USA, 399-408.

[5] Card, Stuart K.; Mackinlay, Jock D.; Shneiderman, Ben. Focus + Context. In: Card, Stuart K.; Mackinlay, Jock D.; Shneiderman, Ben. Readings in Information Visualization: Using Vision to Think. San Francisco, California: Morgan Kaufmann Publishers, 1999. p.307-309.

[6] Chau, M. Visualizing web search results using glyphs: Design and evaluation of a flower metaphor. ACM Trans. Manage. Inf. Syst. 2, 1, Article 2 (March 2011), 27 pages.

[7] Cugini, J., Laskowski, S., Sebrechts, M. Design of 3D visualization of search results: evolution and evaluation. In Visual Data Exploration and Analysis, pages 198-210, IST/SPIE, 2000.

[8] Dorneles, C. F., Gonçalves, R., dos Santos Mello, R. Approximate data instance matching: a survey. Knowledge Information Systems, 27(1): 1-21 (2011).

[9] Draper, G. M., Livnat, Y. and Riesenfeld, R.F. A Survey of Radial Methods for Information Visualization. IEEE Transactions on Visualization and Computer Graphics, 15(5):759-776, 2009.

[10] Firmenich, S., Winckler, M., Rossi, G. A Framework for Concern-Sensitive, Client-Side Adaptation. In Proc. of International Conference on Web Engineering (ICWE 2011), Paphos, Cyprus, June 20-24, 2011. Springer, LNCS 6757, pages 198-213.

[11] Hassan, A., White, R. W., Dumais, S.T., Wang, Y-M. Struggling or exploring?: disambiguating long search sessions. WSDM 2014: 53-62.

[12] Hearst, M. User Interfaces for Search, Chapter 2 of Modern Information Retrieval: The Concepts and Technology behind Search (2nd Edition), Addison Wesley, 2011.

[13] Heer, J., Robertson, G. Animated Transitions in Statistical Data Graphics. IEEE Transactions on Visualization and Computer Graphics, 13(6):1240-1247, 2007.

[14] Mann, T.M. Visualization of WWW-search results. In DEXA Workshop, pages 264-268, 1999.

[15] Marchionini, G. Interfaces for end-user information seeking. Journal of the American Society for Information Science, 43(2):156-163, January 1999.

[16] Marchionini, G. Exploratory Search: From Finding To Understanding. Comm. of the ACM, 49(4):41–49, 2006.

[17] McCrickard, S. and Kehoe, C. Visualizing search results using SQUID. Poster at 6th World Wide Web Conference, April 1997. Available at: http://people.cs.vt.edu/~mccricks/papers/mk97.pdf

[18] Norman, D. The Psychology Of Everyday Things. (1988). Basic Books; 1 edition (June 13, 1988). Basic Books. ISBN 978-0-465-06710-7

[19] Norman, D. A., Draper, S. W. (eds.) (1986): User Centered System Design: New Perspectives on Human-Computer Interaction. Hillsdale, NJ, Lawrence Erlbaum Associates.

[20] Roberts, J.C., Boukhelifa, N., Rodgers, P. Multiform glyph based search results visualization. In Proceedings of Information Visualization 2002, pages 549-554, IEEE, July 2002.

[21] Schwarz, J., Morris, M. R. Augmenting web pages and search results to support credibility assessment. CHI 2011: 1245-1254.

[22] Sebrechts, M.M., Vasilakis, J., Miller, M.S., Cugini, J.V., & Laskowski, S. (1999). Visualization of Search Results: A Comparative Evaluation of Text, 2D and 3D Interfaces. In Hearst, M.A. et al. (Eds.), Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. pp 3-10.

[23] Sutcliffe, A.G., Ennis, M. Towards a cognitive theory of information retrieval. Interacting with Computers, 10:321–351, 1998.

[24] Suzuki, E., Ando, S., Hirose, M., Jumi, M. Intuitive Display for Search Engines Toward Fast Detection of Peculiar WWW Pages. In. WImBI 2006. N. Zhong et al. (Eds.): WImBI 2006, LNAI 4845, pp. 341–352.

[25] Tversky, B., Morrison, J., Betrancourt, M. Animation: Can It Facilitate? Int. J. Human-Computer Studies, 57:247-262, 2002.

[26] Veerasamy, A. Heikes, R. Effectiveness of a graphical display of retrieval results. In Research and Development in information retrieval, pages 236-245, New York, 1997.

[27] Wilson, Max L. Evaluating the Cognitive Impact of Search User Interface Design Decisions. EuroHCIR 2011: 27-30.

[28] Winckler, M., Gaits, V., Vo, D-B., Firmenich, S., Rossi, G. An Approach and Tool Support for Assisting Users to Fill-in Web Forms with Personal Information. In Proc. of the ACM SIGDOC 2011, Pisa, Italy, October 3-5, 2011, ACM, New York, NY, USA, pp. 195-202.